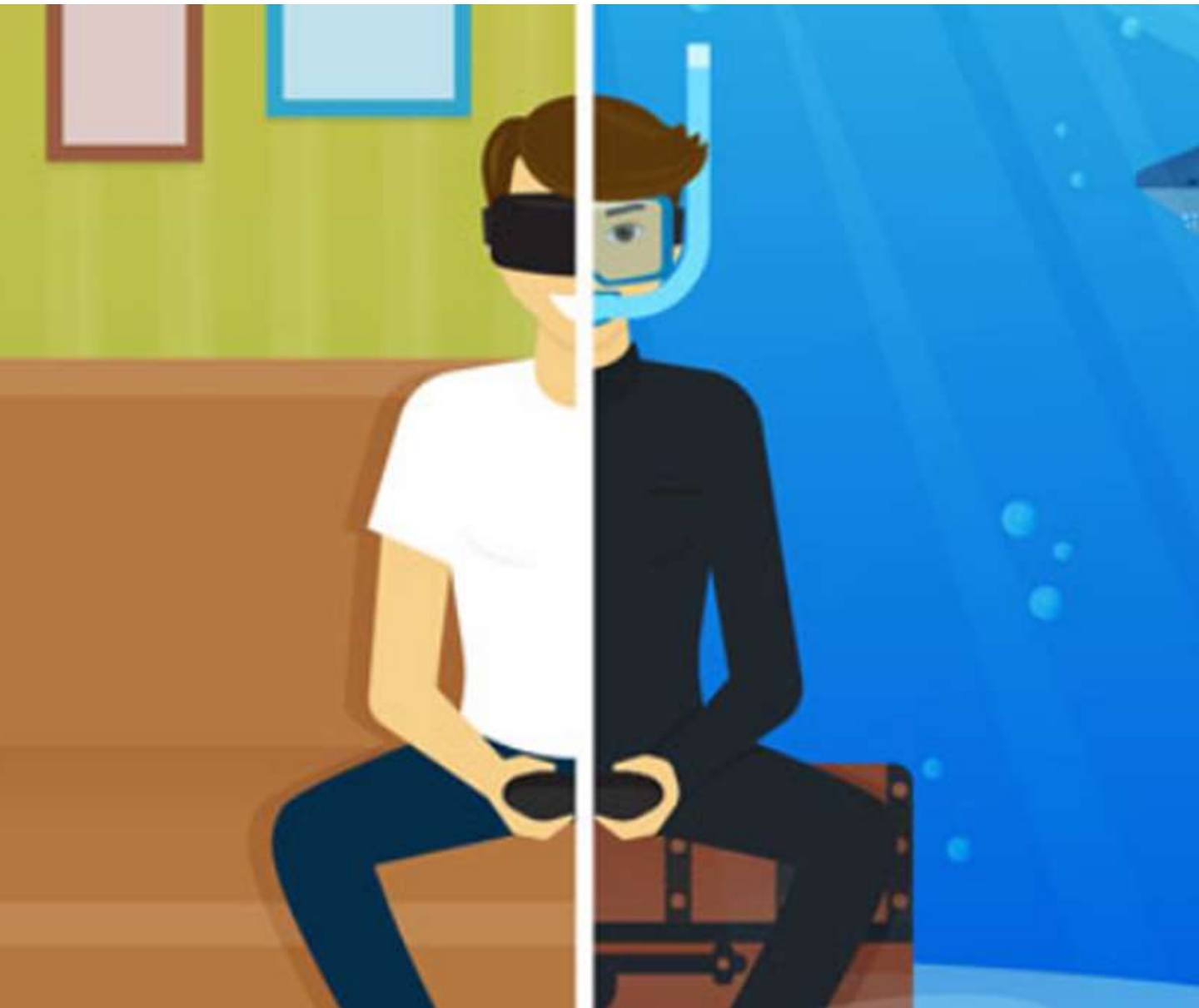


Mit der Technik in die Zukunft?

Apparative Messung



Inhalt

- **03 – Editorial**

Horst Müller-Peters / marktforschung.de

- **05 – How wearables might answer behavioral questions today's Market Researchers might not even think to ask**

Issac Rogers / 20|20 Research, Inc

- **09 – Designforschung mithilfe von Virtual-Reality-Brillen: realitätsnah, effizient und spannend!**

Arne Hellwig / Ipsos

Muriel Girard-Reydet / DB Mobility Logistics AG

Marina Scheckenbach / DB Fernverkehr

- **17– Real Time Response – alte Technologie, die nach wie vor begeistert?**

Dirk Martens / House of Research

- **22 – Car Clinics in virtuellen Welten. Missverständnisse und Potenziale**

Dr. Bastian Zwissler / Spiegel Institut Mannheim

- **25 – Die Vermessung des Unbewussten**

Oliver Spitzer und Markus Küppers / heart.facts

- **34 – Wearables-Technologie in der Marktforschung**

Felix Fischer und Gareth Tuck / eye square

- **38 – Warum sich apparative Forscher nicht unbedingt in der Komfortzone bewegen**

Dr. Beate von Keitz / Institut für Kommunikations-Forschung

- **41 – Sprechen die Daten für sich selbst?**

Dr. Helene Karmasin / Karmasin Behavioural Insights

Editorial



Von Horst Müller-Peters, marktforschung.de

Science Fiction in der Marktforschung

"Apparative" Methoden sind bereits Klassiker in der Marktforschung: Immer interessant, irgendwie cool, sind sie zugleich aber immer eine Nische im Methodenmix geblieben. Sie dienen entweder dazu, die Wahrnehmung, Informationsverarbeitung, Emotion oder Präferenz von Konsumenten unmittelbar zu messen. Die Verfälschungstendenzen durch bewusstes Erinnern, Urteilen und Verbalisieren fallen somit weg. Oder es findet doch eine Befragung statt, die mittels "apparativer" Hilfe Teile der Realität simuliert, die sonst nicht oder nur unter großem Aufwand dargestellt werden könnte.

Von der Schnellgreifbühne in die virtuelle Realität

Mein erstes Erlebnis mit apparativen Methoden war äußerst eindrucksvoll. Die Schnellgreifbühne unseres damaligen Hochschulinstituts diente der unmittelbaren Präferenzmessung. Produktvarianten — beispielsweise verschiedene Bierflaschen — wurden in regalähnlicher Situation für wenige Sekunden exponiert, und die Probanden trafen mittels Zugreifen eine schnelle

Kaufentscheidung. Dieses Gerät machte dem Begriff Apparat alle Ehre: Groß, schwer, komplex und irgendwie furchteinflößend. Als Student sah ich vor meinem geistigen Auge immer abgetrennte Hände — für den Fall, die Mechanik hätte sich schneller geschlossen, als der Proband hätte reagieren können. Etwas weniger bedrückend wirkte ihr kleiner Bruder, das Tachistoskop. Aber auch hier war viel robuste Mechanik am Werk, um die Wirkung von flüchtiger Wahrnehmung von Produkten oder Anzeigen zu simulieren. Sensationell futuristisch erschien uns dagegen die Ankündigung eines unserer Psychologieprofessoren, mittels eines kleinen, batteriegespeisten Kästchens namens Vitaport physiologische Daten wie Blutdruck oder Hautwiderstand mobil und damit außerhalb des Labors messen zu können.

Schnellgreifbühne, Tachistoskop und selbst Vitaport in der damaligen Form sind heute längst in Rente und haben neuen Geräten Platz gemacht. Diese sind zwar meist weniger eindrucksvoll, aber dafür umso intelligenter. Chips, Bits und Bytes haben die aufwändige Mechanik weitestgehend verdrängt. Blickregistrierung (ohne störende Brille), automatisierte Emotionsmessung über die Mimik,

Hirnscans und das Eintauchen in virtuelle Realitäten sind nur Beispiele für den Einsatz neuer Techniken in der Marktforschung. Wie weit wir heute damit sind, bringt Isaac Rogers von 20|20 Research auf den Punkt: "A researcher in 2015 has at his disposal an array of tools that would read almost like science fiction to a researcher from 2005". Und da wir methodisch erst am Anfang der digitalen Welle stehen, dürfen wir wohl davon ausgehen, dass uns die Forschung 2030 aus heutiger Sicht erst recht erscheinen wird wie aus einer anderen Galaxie.

Einblick und Ausblick

In unserem diesjährigen "Adventsdossier", das der Jahreszeit entsprechend etwas weniger umfassend ausfällt als gewohnt, geben wir Ihnen einen Einblick in den derzeitigen Status Quo und zugleich einen Ausblick: Welche Methoden werden heute schon erfolgreich eingesetzt, und was ist in der Pipeline? Und wie werden die fortschreitende Virtualisierung, die kontinuierliche Ortung, die Omnivernetzung durch Wearables (physiologische Messungen inklusive) und das Internet der Dinge bis hin zur Implantation von Chips in den menschlichen Körper nicht nur die Gesellschaft, sondern auch die Marktanalyse in der weiteren Zukunft verändern?

Die "apparative" Messung (auch wenn der Begriff in diesem Kontext antiquiert wirkt) verlässt die Nische und wird omnipräsent. Zugleich verschmelzen aktive und passive Methoden, Marktforschung und Big Data. Frei nach George Orwell: "Market Research is watching you" – nur hoffentlich für einen

guten Zweck und ohne die totalitären Nebenwirkungen.

Ihr Horst Müller-Peters
(Herausgeber)

How wearables might answer behavioral questions today's market researches might not even think to ask



Von Issac Rogers, Chief Innovation Officer bei 20/20 Research, Inc.

When you ask market researchers what technology has had the most dramatic impact on data collection methodologies in the past ten years, most (if not quite all) will be quick to answer "the smartphone". Perhaps no category of device since the Internet-connected computer has changed the way that consumers live their lives as much as our beloved iPhones and Androids. The smartphone era has drastically shifted the way consumers shop, communicate, and entertain themselves; this shift presented incredible new ways to collect behavioral and attitudinal insights never possible through traditional methodologies. A researcher in 2015 has at his disposal an array of tools that would read almost like science fiction to a researcher from 2005; mobile surveys, geofencing, and go-anywhere digital qualitative discussions would all seem like tools from the future to someone looking ahead just ten years ago. If the smartphone has changed so much about our world and the market research industry in the past 10 years, one might ask themselves what

technology will become the next evolutionary leap in the coming 10 years, and how will that impact the way we think about and conduct research in the future?

Just as the smartphone revolutionized our world, many tech companies are betting big on the next evolution in our digital lives: wearables. Leading trend research firms like Gartner and Juniper Research expect this new market segment to grow rapidly over the next few years as consumers embrace the new devices. These wearable technological marvels, devices like the Apple Watch, Microsoft™ Band, NikeFuel, and FitBit, give consumers a simple way to passively collect troves of personal data about their physical activity throughout the day. Some of the more advanced devices allow the user to make phone calls, send text messages, and even hail an Uber — just by raising their wrist and touching the screen.

So what new opportunities will these

wearables present to the market research industry, and how can we best prepare ourselves to take advantage of these exciting new tools to better understand consumer behavior?

Observing passive data: Putting respondent feedback into real-world context

When you stop and consider the way we collect digital research today, we actually don't know much about the context of our consumers as they fill in our online forms or participate in our online qualitative discussions. We might know the time of day they completed the survey, or perhaps what kind of PC or smartphone they were on when they answered the questions, but for any other contextual insights into their experience we must rely on that consumer's self-reported behaviors and their recollection of events and attitudes throughout their day.

How beneficial might it be if we had a glimpse into our respondent's actual human context while they are participating in our research? What if we could capture data about the respondent's daily experiences and use this data to better understand patterns and trends in consumer behavior? What if this data didn't require any significant input from the respondent, and was generated passively and transmitted easily to us as a new level of metadata for research?

This is where wearables and their massive troves of passively-collected activity data might shine new light on actual consumer behaviors. By getting past consumer recall

and instead relying on wearables to track minute actions and events throughout the day, we as researchers may unveil a more complete picture of our respondents.

Envision a scenario where we worked with our research respondents to incentivize them to share their wearable's activity data for the previous 7 days and upload that information after they complete a survey. Using technology that exists today, we could develop an application to capture personal timelines that included physical activity (number of steps walked/ran per day, exercise routines and frequency, percentage of time spent sitting/standing), environmental data (local weather for a particular respondent), and approximate geolocation (perhaps having the respondent identify the address of their work and home on a map).

For a survey to gauge the effectiveness of a new advertising campaign, what if we find a correlation between the concept rating scores and how physically active the consumer had been that day? Or perhaps we might uncover that consumers are slightly more positive towards a new online shopping concept if their average commute to/from work took longer than average? Could we find that certain messaging works better when consumers have had a "low stress" day versus a "high stress" day?

With this data we can now extract new layers of segmentation data, looking for patterns or trends that might influence survey responses or highlight previously unseen trends. What if this new layer of data tells us things about our respondent's actual human contexts that we never

considered before? Once we begin gathering data that puts the consumer in a real-life context, what if we find strong predictors of marketing message receptivity are derived from the precise daily activity of our panelists? For advertisers, this kind of behavioral context might radically change the way they target certain audiences or deliver certain marketing messages.

Gathering new active data: "microfeedback" from wearables

When researchers think about the impact of wearables, they most often consider the aforementioned "passive" aspect of data collection; however, the ability of these devices to collect active data from respondents might prove to be just as powerful a research tool.

Take, for example, a simple rating app designed for the Apple Watch, that enables focused "microfeedback" survey data on in-home usage of a new prototype athletic shoe. A research firm might recruit a panel of respondents to download the app to their devices and respond to push surveys directly on the watch face. Perhaps once per day, or at certain times of day, the user could provide simple ratings back to the researcher on the current "feel" of the shoe. The user's wrist will vibrate to notify them of a new question and one touch later they've registered their feedback. Simple trigger logic based on consistently low ratings might generate a phone call to the respondent for a quick interview on why the prototype shoe is failing to meet expectations.

Recently, a client challenged us to prototype a real-time behavioral trigger that can gather in-the-moment feedback regarding coffee drinking at home. We were able to leverage off-the-shelf technology to prove you can send a survey to a respondent's phone every time they brewed a new cup of coffee at home. We

developed a method to trigger a mobile survey using a "smart plug" WiFi adapter; the adapter fired off a command to launch a new survey each time the coffee machine was turned on by the user, and this technology could be retrofitted in less than 15 minutes to any common household appliance. While this particular project never made it out of our labs, it showed the promise that activity-based data collection has in store for our collecting insights closer to the point of actual usage.

Our Wearable-enabled Future

Some may brush off the wearables trend as a "fad", or argue that the data has fleeting relevance to market research — the truth is, no one can predict the future and they may very well be correct. However, should the wearable trend continue, it is of little doubt that brands and advertisers will see wearables as another screen on which to market to their customers, and we should prepare ourselves as an industry to understand the impact wearables have on the consumer experience. Others might argue this massive trove of passive data is best left for the "data scientists" in some other discipline who are working to crack the code on big data. That's a somewhat scary proposition, as it means we

relinquish control of some potentially valuable consumer insight to other entities that may one day be able to explain or predict components of consumer behavior better than our own industry.

My hope for the market research industry is that we embrace the potential of wearables and how they might reshape our understanding of actual consumer behavior. We mustn't let this potentially powerful new medium pass us by; we must work together, as technology providers and researchers, to understand how best to integrate this new generation of digital tools into our toolkit.

Designforschung mithilfe von Virtual-Reality-Brillen: realitätsnah, effizient und spannend!



Arne Hellwig ist qualitativer Marktforscher bei [Ipsos](#) in Hamburg. Er hat Sozial- und Kommunikationswissenschaften studiert und ist seit 2012 bei Ipsos beschäftigt. Zu seinen thematischen Schwerpunkten gehören neben der Mobilitäts- und Designforschung auch Kommunikationsforschung und Produktneuentwicklung.



Muriel Girard-Reydet ist Projektleiterin Marktforschung bei der DB Mobility Logistics AG in Frankfurt/Main seit 2001. Zu ihren Schwerpunkten gehören Zug-, Sitz- und Buslabore, UX-Tests digitaler Angebote und Konzepttests für den Personenverkehr.



Marina Scheckenbach ist Projektleiterin im Kundenzufriedenheitsmanagement ICE/IC in Frankfurt/Main und seit 2009 bei der DB Fernverkehr AG tätig. Schwerpunktmäßig kümmert sie sich im Produktmanagement um die Ermittlung der Kundenbedürfnisse hinsichtlich der Zuggestaltungen, -ausstattung und -modernisierung von Fernverkehrszügen.

Virtual-Reality-Brillen (VR-Brillen) kommen einer wirklichkeitsgetreuen Darstellung immer näher. Sie bieten dem Nutzer die Möglichkeit, in virtuelle Welten einzutauchen (Immersion) und diese detailliert zu betrachten. Große Realitätsnähe, bei vergleichsweise geringen Kosten und einem hohen Involvement der Probanden, macht diese Technik für die Designforschung besonders interessant. Im Auftrag der Deutschen Bahn hat Ipsos eine Studie zur zukünftigen Gestaltung der Fernverkehrszüge durchgeführt, bei der erstmalig VR-Brillen zum Einsatz kamen.

Dabei wurde die Oculus-Rift-Brille, ein Vorreiter in der VR-Technik, verwendet. Dieser Beitrag zeigt, wie und für welche Fragestellungen diese Technik eingesetzt werden kann.

Evaluation unterschiedlicher virtueller Designentwürfe

Das Design spielt eine wichtige Rolle für den Reisekomfort, das Reiseerlebnis und die Zufriedenheit von Reisenden. Dementsprechend ist die Erforschung von Kundenbedürfnissen und -anforderungen

zentral für die Entwicklung und Optimierung von Zug-, Bus-, Flug- oder Fahrzeugdesigns. Designforschung bei der Deutschen Bahn erfolgte bisher in der Regel direkt am Objekt, zum Beispiel in einem fahrenden, stehenden Zug oder in einem Mockup, das heißt in einer maßstabsgetreuen Nachbildung eines Zuges beziehungsweise eines Wagens, im Rahmen von sogenannten Clinics oder Laboren. Der verstärkte Einsatz digitaler Innovationen und Angebote im Personenverkehr der Deutschen Bahn (Mobilität 4.0) bewegte das Produktmanagement des Fernverkehrs und die Konzernmarktforschung dazu, Instrumente digitaler Herkunft für Forschungszwecke "auszuprobieren" beziehungsweise zu testen. Vor diesem Hintergrund kam es zum Einsatz von

VR-Brillen, zur Darstellung von Designentwürfen bei der Untersuchung zukünftiger Zugmodernisierungsmaßnahmen. Dabei wurden den Probanden in einem qualitativen Forschungsansatz unterschiedliche virtuelle Entwürfe gezeigt, die unter anderem bezüglich der Anzahl und Anordnung der Sitzplätze, der Größe und Anzahl von Tischen und Ablagemöglichkeiten variierten. Zielsetzung war die Eruierung, welcher Designentwurf den Bedürfnissen der Reisenden am stärksten entsprach und wie das Design insgesamt verbessert werden könnte. Dabei spielten emotionale Aspekte wie Atmosphäre und Raumgefühl, aber auch funktionale Aspekte wie Ausstattung, beispielsweise eine adäquate Tischgröße und Position, eine Rolle.



Abb. 1: Blick durch die VR-Brille (Bild: Ipsos / Deutsche Bahn)



Abb. 2: Blick durch die VR-Brille (Bild: Ipsos / Deutsche Bahn)

VR-Brillen ermöglichen eine realitätsnahe Darstellung von Design-Entwürfen

Eine VR-Brille ist ein visuelles Ausgabegerät in Brillenform, das beim Träger der Brille ein dreidimensionales Bild entstehen lässt. Durch die realitätsnahe 360-Grad-Darstellung hat der Träger einer solchen Brille die Möglichkeit vollständig in die virtuelle Realität einzutauchen und sich durch die Drehung des Kopfes in den virtuellen Zugentwürfen umzusehen. Anhand eines Spiel-Controllers, wie etwa in der X-BOX 360, kann er sich vorwärts, rückwärts und seitwärts innerhalb des virtuellen Zugs bewegen.

Begleitet wurde der Proband von einem qualitativen Marktforscher, der über die Verkabelung der VR-Brille an einem PC-Bildschirm die Kopfbewegungen des Befragten auf dem Bildschirm beobachten und nachvollziehen konnte. Dadurch konnte der Forscher gezielt Fragen zu Design und Zugausstattung stellen. Auf diese Weise wurden die Atmosphäre, der räumliche Eindruck, die Ausstattungskomponenten und die Gesamtattraktivität der jeweiligen Zugentwürfe evaluiert.



Abb. 3: Proband während der Design-Evaluation (Bild: Ipsos / Deutsche Bahn)

Immersion und Involvement

Zu Beginn der Design-Untersuchung wurde der Proband in die Bedienung der VR-Brille mithilfe des Controllers eingewiesen. Durch die unkomplizierte Steuerung via Joystick waren auch Teilnehmer mit wenig oder keiner Erfahrung mit Spiel-Controllern in der Lage, selbstständig durch den Gang und den Großraumwagen des Zuges zu navigieren. Um Stimulus-bedingte Verzerrungen im Hinblick auf die Bewertung der virtuellen Großraumwagen zu vermeiden, begann die Design-Evaluation nicht direkt im Großraumwagen, sondern im Gang des Zuges. Jeder Tester erhielt dort mehrere Minuten Zeit, um sich an das Virtual-Reality-Erlebnis und die Steuerung zu gewöhnen. Anschließend erfolgte die virtuelle Begehung der unterschiedlichen Großraumwagen-Designentwürfe. Die dreidimensionale Darstellung und die Bewegung im Raum wurden als realitätsnah wahrgenommen, was die Reaktionen der Teilnehmer

verdeutlichen:

"Man hat das Gefühl, da richtig drin zu sein, das ist so wie in echt."

"Ich fand, es gibt einem einen wirklich sehr guten Eindruck. So als wäre man im Großraumwagen, und man kann ja auch agieren."

Die Realitätsnähe der Raumwahrnehmung und das Eintauchen in die virtuelle Realität zeigen sich auch darin, dass die Probanden bei ihrer Bewertung nach virtuellen Objekten, etwa den Tischen, griffen oder darauf zeigten. Die realitätsnahe Form der Darstellung und die Möglichkeit, sich im virtuellen Raum zu bewegen, führten zu einer lebhaften und intensiven Auseinandersetzung mit den unterschiedlichen Designentwürfen, bei der ausführliche und detaillierte Designerfahrungen und Informationen zur Wirkung der jeweiligen Großraumwagen sowie konkrete

Optimierungsempfehlungen gewonnen werden konnten.

Virtuelle Auszeiten und selbstständige Navigation, um Motion Sickness zu begegnen

Den positiven VR-Effekten, wie großer Realitätsnähe und hohem Involvement der Probanden, steht das Phänomen der Motion Sickness gegenüber, das Übelkeits- und Schwindelgefühle bei der Nutzung von VR-Brillen beschreibt. Das Auftreten dieses Problems ist bekannt und wird dadurch erklärt, dass die Sinnesorgane unterschiedliche Informationen gleichzeitig verarbeiten müssen: Im Fall dieser Studie erklärt sich Motion Sickness durch die Bewegung in dem virtuellen Raum mithilfe des Controllers, während der Proband selbst bewegungslos auf einem Stuhl sitzt. Beeinflusst wird dies auch durch die Qualität der Auflösung, die Bildwiederholfrequenz und die Latenz beziehungsweise Verzögerung, mit der die (Kopf-)Bewegungen des VR-Brillen-Trägers in die virtuelle Darstellung übertragen werden. Dem Problem der Motion Sickness wurde in der Fernverkehrszug-Studie dadurch entgegengewirkt, dass die Befragten, sobald erste Gefühle des Unbehagens auftraten, die VR-Brille abnahmen und eine kurze virtuelle Auszeit einlegten.

Zusätzlich wurde die virtuelle Begehung, die Wahrnehmung und offene Bewertung der Großraumwagen auf maximal 25 Minuten begrenzt. Die Teilnehmer navigierten selbstständig durch die Großraumwagen, in ihrem eigenen Tempo. Insgesamt setzten sich alle Probanden mit den unterschiedlichen Designentwürfen

auseinander. Tester, die von Motion-Sickness-Effekten betroffen waren, zeigten sich dennoch begeistert über die Realitätsnähe und Bewegungsmöglichkeiten im virtuellen Raum:

"Mir war schon ein bisschen schwindlig zwischendurch, aber trotzdem fand ich super, dass es so echt aussieht und man hin- und hergehen und sich umschauen kann."

Durch die technischen Weiterentwicklungen im Bereich der VR-Brillen, eine exakte Ortung und verzögerungsfreie Wiedergabe aller Bewegungen sollte sich das Problem der Motion Sickness bei zukünftigen Studien weiter verringern lassen, im besten Fall gar nicht mehr auftreten. Auch die Bildschirmauflösung wird zukünftig weiter steigen und die Realitätsnähe erhöhen.

Positive Erfahrungen überwiegen, zudem vergleichsweise geringe Kosten

Insgesamt überwiegen die positiven Erfahrungen beim Einsatz einer VR-Brille. Diese bietet eine wirklichkeitsgetreue Darstellung, bei der die Teilnehmer sich aktiv mit einem virtuellen Designentwurf auseinandersetzen können. VR-Technik lässt sich, wie die durchgeführte Studie zeigt, besonders sinnvoll einsetzen, wenn die Wahrnehmung und Wirkung des Raumes eine Rolle spielt. Im Vergleich zu einem konventionellen Mockup bestehen Zeit-, Raum- und Kostenvorteile. Programmierung und Technik, die für die Evaluation der virtuellen Designentwürfe benötigt werden, sind vergleichsweise günstig.

Zusätzliche Varianten eines Designentwurfs oder Änderungen am Testmaterial können von der Programmierung flexibler, zügiger und günstiger umgesetzt werden. Zudem sind die Kosten für Raummiete, Logistik und Transport bei einer Darstellung mithilfe einer VR-Brille wesentlich geringer, als bei einem physischen Mock Up.

Kunden frühzeitig in den Entwicklungsprozess integrieren

Die Möglichkeit, unterschiedliche Designvarianten zu testen und diese Entwürfe dann vergleichsweise flexibel anzupassen, ist besonders vor dem Hintergrund der Implementierung agiler Design-Thinking-Ansätze bei der Deutschen Bahn von großem Nutzen. Beim Design Thinking stehen die Bedürfnisse und Motivationen der Kunden während des gesamten Produktentwicklungsprozesses im Fokus. Die Reisenden werden möglichst frühzeitig in den Entwicklungsprozess neuer Zugprodukte miteinbezogen, sodass ihre Bedürfnisse und Anregungen in mehreren Entwicklungsschleifen optimal in die Gestaltung der Produkte einfließen können. Dabei bietet die VR-Brille eine einfache Möglichkeit, Designentwürfe realitätsnah zu testen und in einem iterativen Prozess bis zur Markteinführung zu verbessern. Es entsteht ein ausgereiftes, mit Kunden entwickeltes Design, das im Idealfall kurz vor Marktauftritt nicht mehr physisch angepasst werden muss.

Virtuell bleibt virtuell?

Trotz ihrer Vorteile können VR-Brillen physische Tests nicht komplett ersetzen. Zum Beispiel, wenn es um die Haptik und Qualität verbauter Materialien oder die

Handhabung der Ausstattung geht. Noch unklar ist die Frage, in welchem Maße sich die Ergebnisse auf die Realität übertragen lassen: Inwiefern unterscheidet sich die räumliche virtuelle Wahrnehmung im Großraumwagen von der Wahrnehmung in der Realität, zum Beispiel, wenn es um die Beinfreiheit geht? Dabei spielen die wahrgenommenen und geschätzten Größenverhältnisse eine Rolle. Zudem stellt sich auch die Frage, wie sich die virtuelle Realität, in der es beispielsweise keinen Schmutz, keine anderen Fahrgäste und kein Gepäck gibt, auf die Bewertung der Designentwürfe auswirkt? Eine Programmierung derartiger Faktoren, welche die Darstellung noch näher an die Realität rückt, wäre durchaus denkbar. Generell bedarf es daher weiteren Untersuchungen und Methodentests, um die Auswirkungen und die Übertragbarkeit virtuell gewonnener Erkenntnisse in der physischen Realität zu belegen.

Die Zukunft ist VR

Die Erfahrungen aus dieser Pilotstudie und anderen Designstudien der Deutschen Bahn und Ipsos zeigen das Potenzial von VR-Technik für die Designforschung. Der Einsatz von VR-Brillen bietet eine wirklichkeitsgetreue Darstellung von Designentwürfen, bei der die Probanden in die virtuelle Realität eintauchen und sich so ausführlich mit dem jeweiligen Design auseinandersetzen können. Es lassen sich die Designwirkung und Raumwahrnehmung realitätsnah evaluieren und konkrete Handlungsempfehlungen zur bedürfnisgerechten Anpassung von Designentwürfen gewinnen. Unterschiedliche Designvarianten können

im Vergleich zu einem Mockup kostengünstig evaluiert und angepasst werden. Designentwürfe werden so Schritt für Schritt, in enger Zusammenarbeit mit den Kunden, an deren Bedürfnisse angepasst. Diese Technik ist dabei prinzipiell für jede Zielgruppe, auch ohne VR-Erfahrung, nutzbar. Insofern ist davon auszugehen, dass sich die Technik im Rahmen von Prototyping einen festen Platz im Designforschungsprozess sichern wird.

IT'S TIME FOR **SUBSTANCE**

Every day see the world changing faster and faster. Keeping up needs both fast and accurate intelligence. Substance is crucial.

Being at your side, providing you with precise, relevant and accurate information.

At Ipsos we deliver this. We deliver the intelligence that will shed light on the decisions you need to make to move forward. We are Game Changers.



GAME CHANGERS



Real Time Response – eine alte Technologie, die nach wie vor begeistert



Dirk Martens ist Geschäftsführer des Medienforschungsinstituts House of Research in Berlin. Zudem ist er Dozent an der Filmuniversität Babelsberg Konrad Wolf, wo er Digitale Medienkulturen / Nutzung Digitaler Medien lehrt.

"Real Time Response Measurement" (RTR) ist ein apparatives Forschungsverfahren, dessen Ursprünge bis in die 1930er Jahre zurückreichen. Trotz des Alters ist es alles andere als überholt. Mit dieser Methode lassen sich Medieninhalte sekundengenau und zeitgleich zur Rezeption durch Probanden in Fokusgruppen oder größeren Hall-Tests evaluieren.

Paul Felix Lazarsfeld ist vielen Lesern durch seine soziographische Studie "Die Arbeitslosen von Marienthal" (1933) bekannt, in der er die soziologischen Veränderungen nach der Schließung einer Fabrik in einem kleinen österreichischen Ort untersuchte, in der die meisten Bewohner vormals in eben dieser Fabrik arbeiteten und von einem Tag auf den anderen alle arbeitslos wurden.

Mindestens ebenso berühmt ist seine für die Kommunikationsforschung bahnbrechende Studie "The people's choice", in der Wähler eines Countys von Ohio zur bevorstehenden Präsidentschaftswahl 1940 befragt wurden

und mit der die Entstehung persönlicher Wahlentscheidungen untersucht wurde. Er begründete damit das Kommunikationsmodell des "two step flow", nach der die Medien zunächst "opinion leader" erreichen, die dann die ursprüngliche Information gefiltert und gewichtet an ihre „peer group“ weitergeben. Auch wenn das Modell insgesamt durch die Wirkungskraft des damals erst aufkommenden Massenmediums Fernsehens an Bedeutung verloren hat, so spielt der Begriff "Meinungsführer" in der empirischen Forschung nach wie vor eine große Rolle.

Etwas weniger bekannt ist, dass der Österreicher Lazarsfeld ursprünglich Mathematik studierte und dieses Fach in den 20er Jahren als Gymnasiallehrer in Wien unterrichtete, bevor er Anfang der 30er am Psychologischen Institut an der Universität Wien lehrte. Diese Disziplinen waren dann wohl auch maßgeblich für seine technischen Entwicklungen, die er nach seiner Emigration in die USA 1933 vorantrieb. Als Direktor des "Office of

Radio Research" an der Princeton University entwickelte er mit Frank Stanton, dem späteren Präsidenten von CBS, den "Stanton-Lazarsfeld Program Analyzer". Mithilfe einer Apparatur konnten Probanden, während sie ein "Radio Drama" (frühe Form der Soap Opera) hörten, mitteilen, wie sehr sie das Gehörte mögen. In jedem Moment konnte durch Drücken eines grünen Knopfes eine positive, oder eines roten Knopfes eine negative Bewertung abgegeben werden. Jeder Proband erhielt dazu ein entsprechendes Handgerät, das dann elektrische Impulse abgab, die der Program Analyzer, von den Erfindern liebevoll "Little Annie" genannt, auf einer Art Rollenplotter als lange Linie zeichnete. Vergleichbar mit einer EKG-Grafik ließ sich darauf der Grad der Likes bzw. Dislikes auf der Zeitachse darstellen, sodass man im Nachhinein jede Szene des Dramas genau analysieren konnte.



Seit 2001 setzt das House of Research auch das RTR-Verfahren ein. (Bild: House of Research)

Dieses "Real Time Response"-Verfahren (das manche auch "Continuous Response"-Verfahren nennen) erlangte dann in den USA immer mehr Bedeutung.

Wurden anfangs vor allem Radiosendungen untersucht, standen später zunehmend auch Fernsehsendungen im Fokus. Die Handgeräte zur Bewertungsabgabe wurden weiterentwickelt und erhielten einen Drehknopf, mit dem jetzt nicht mehr nur dichotome Antworten wie "yes/no" etc. erfasst werden konnten. Mit dem Fortan als "Dial" bezeichneten Handset konnte nun auch eine Verhältnisskala von 0 bis 100 für eine quasi stufenlose Abfrage verwendet werden, wodurch eine feingliedrige Abstufung der Bewertung für die Messung im Zeitverlauf möglich wurde. Mit der Einführung von drahtlosen Dials, die ihre Messwerte kontinuierlich per Funk an die im Raum befindliche Basis schickten, nahm vor allem die Usability für die Probanden zu, die sich nun freier und entspannter platzieren konnten. Auch wurde der Testaufbau viel einfacher und schneller, da das Verkabeln der Geräte und das Verkleben der stolpergefährlichen Kabel entfielen.

Ein spannendes Feld wurden zum Beispiel die "Presidential Debates", bei denen sich die Anwärter ums Weiße Haus vor der Kamera ein Wortgefecht liefern. Egal ob Ronald Reagan, Bill Clinton, Barack Obama und viele andere: Sie und ihre Gegenspieler mussten sich alle der harten Jury von Forscher-Teams stellen, die zeitgleich mit Teilnehmern größerer Fokusgruppen vor dem Fernseher saßen und Moment für Moment, Wort für Wort und Argument für Argument genau evaluierten. Bewertet wird aber nicht nur das explizit Gesagte. Auch die non-verbale Kommunikation, die Gestik und Mimik, spielt dabei in die Bewertung der Probanden hinein. Diese fehlende Trennung zwischen kognitiven und affektiven Einflüssen auf die

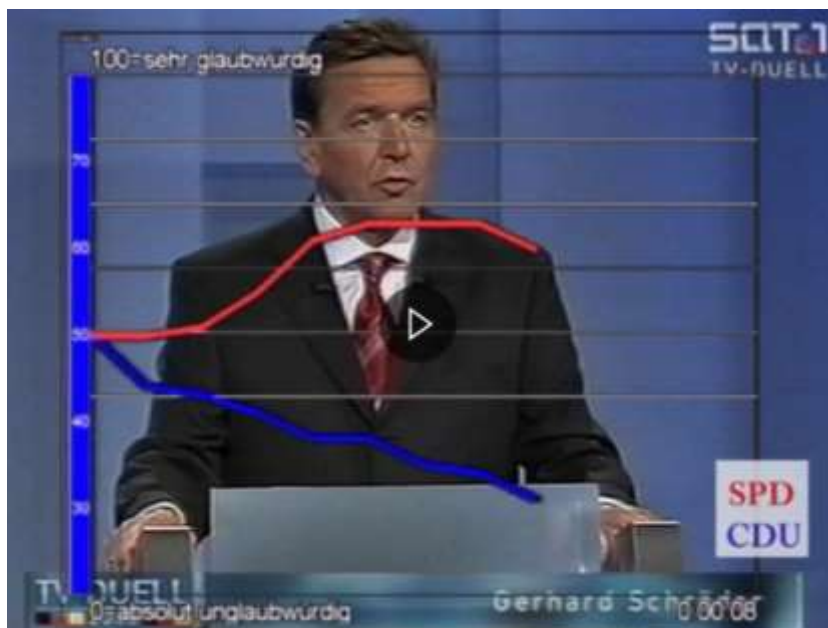
Bewertung ist dabei aber nicht unbedingt ein Nachteil. Ganz nach dem amerikanischen Prinzip des "keep it simple" wird nämlich genau das vermengt, was auch bei der Wahlentscheidung den Ausschlag gibt: eine Mischung aus Kopf und Bauchgefühl, aus Kenntnis konkreter Wahlaussagen und Sympathie oder Antipathie gegenüber den Kandidaten.

Den TV-Debatten wird auch deshalb eine große Rolle für den Wahlausgang zugeschrieben, weil sich unentschlossene Wähler oft erst vor der Wahlurne für eine Partei entscheiden, wo dann für eine Abwägung von Sachargumenten keine Zeit mehr bleibt. Viele entscheiden sich dann für die Partei, von der sie glauben, dass sie die Wahl gewinnen wird, da sie eher zur Gewinnerseite zählen möchten. Dieser "band wagon effect" war übrigens auch ein Ergebnis von Lazarsfelds "The People's Choice"-Studie.

Wie die Erfahrungen des Autors zeigen, lassen sich ganz unterschiedliche

Untersuchungsobjekte mit dem RTR-Verfahren sehr gut in Fokusgruppen oder in größeren Auditorien analysieren. Prinzipiell bieten sich alle audiovisuellen und auditiven Medien dazu an, also zum Beispiel jegliche Art von Fernsehformat, Werbespots, Programmstrecken oder ganze Musikkataloge von Radiosendern bis zu Langfilmen im Kino. Auch Live-Veranstaltungen wie Musicals oder Parteitage können untersucht werden, da die drahtlosen Geräte mit Funkübertragung keine speziellen Testräumlichkeiten erfordern.

Reizvoll an der Technik ist dabei nicht nur, die Likes/Dislikes jeder Sekunde eines Medieninhalts betrachten zu können, sondern auch, die auf- und absteigende Bewertungskurve zeitgleich zu sehen, also während der Rezeption des Medieninhalts durch die Probanden. In der Fokusgruppe hört man zum Beispiel das Gelächter oder sieht die ernsten Gesichter während die Probanden ihre Dials in Negativstellung bringen, was der Interpretation hilft.



In Deutschland fand 2002 das erste Kanzler-TV-Duell zwischen Gerhard Schröder und Edmund Stoiber statt. (Bild: House of Research)

Im August 2002 fanden in Deutschland zum ersten Mal Rededuelle zwischen Kanzlerkandidaten im deutschen Fernsehen statt. Nachdem Helmut Kohl dies immer abgelehnt hatte, wagte sich Gerhard Schröder als erster in den Ring der TV-Journalisten. Während sich Schröder parallel auf Sat.1 und RTL mit seinem Herausforderer Edmund Stoiber verbal duellierte, sahen zeitgleich 100 nach Parteienproporz und soziodemographischen Merkmalen ausgesuchte Probanden die Live-Übertragung in einem Berliner Hotelsaal auf einer Großleinwand und bewerteten währenddessen das Geschehen mit ihrem Dial. Die Bewertungskurve verriet sekundengenau, bei welchen Argumenten Schröder vorne lag, und wo Stoiber punkten konnte. Der Aufriss nach Parteienanhängerschaften kann interessante Dissonanzen aufzeigen, wenn etwa die Zustimmung des gegnerischen Lagers größer ist als des eigenen.

Auch wenn diese Messung den späteren Wahlsieg von Schröder sehr klar abbildete, ist es natürlich kein Prognoseinstrument. Aufgrund der in der Regel auf wenige hundert Teilnehmer begrenzten Fallzahl, aber auch wegen des Fehlens eines repräsentativen Stichprobenansatzes, handelt es sich um ein qualitatives Forschungsverfahren, dessen Ergebnisse sich nicht hochrechnen lassen. Vielmehr liegt der Nutzen des Verfahrens in der qualitativen Analyse der Medienwirkung. Wie werden in obigem Beispiel etwa politische Botschaften inhaltlich, sachlich bewertet? Inwieweit kann die Art der Präsentation durch den Redner, dessen Rhetorik, Gestik und Mimik, dessen Glaubwürdigkeit und zugesprochene

Themenkompetenz überzeugen? Den Einfluss dieser Faktoren auf ihre Medienwirkung wohl wissend, sollen sich in den USA auch schon einige Präsidentschaftskandidaten auf die TV-Debatten mithilfe von RTR vorbereitet haben lassen.

Ganz ähnlich wie mit Politikerdebatten verhält es sich mit Werbespots, die ebenso auf ihre Kommunikationsleistung per RTR untersucht werden können. Hier werden alternative Motive gegeneinander getestet oder auch Storyboards untersucht, bevor sie teuer produziert werden. Teilweise bietet sich auch eine Kombination von RTR mit physiologischen Verfahren an, um der unbewussten (Werbe-) Wirkung näher zu kommen. Nach meiner Meinung empfiehlt sich dies jedoch nicht grundsätzlich, sondern eher bei Grundlagenstudien, die hohen Aufwand für kleine Stichproben rechtfertigen. Auch die affektiven Komponenten können mit RTR nämlich recht gut erfasst werden, wenn die Methode richtig operationalisiert wird.

Gerade Unkenntnis von der richtigen Operationalisierung führt jedoch bisweilen zu einer Methodenskepsis, die oft nicht gerechtfertigt ist. So ist es etwa sehr entscheidend, wie die Endpunkte der Skala bei der RTR-Messung benannt werden. Eine Skala von "mag ich sehr" bis "mag ich überhaupt nicht" macht zum Beispiel bei der Evaluierung eines Dramas überhaupt keinen Sinn. Eine gewalthaltige Szene kann für die Erzählung und auch für die Zuschauer sehr wichtig sein. Doch würde kaum ein normaler Mensch hierbei mit "mag ich" antworten. Dieses Problem hat bekanntlich ja auch Facebook mit dem Daumen, der immer nach oben zeigt und

mal für Zustimmung und mal für das Gegenteil benutzt werden soll. Hier benötigt man dann eher eine Skala für das Involvement der Zuschauer, das den verschiedenen Bedürfnissen und Erwartungen an die Filmnutzung wie Unterhaltung, Spaß, Spannung unter anderem viel besser Rechnung trägt.

Skepsis gibt es teilweise auch bezüglich der Länge des zu untersuchenden Formates, je länger der Test, desto eher würden die Probanden apathisch erstarren und das Evaluieren vergessen. Wenn die Probanden in einem Warm-up richtig instruiert werden, dann lernen sie jedoch schnell, den Dial sehr intuitiv und fast unbewusst zu bedienen, ohne dass es zu nennenswerten Ermüdungserscheinungen kommt. Zeigen lässt sich das sowohl mit einer Betrachtung der hoch bleibenden Volatilität der individuellen Bewertungskurven, als auch mit einer hohen Test-Retest-Reliabilität zwischen der Bewertung eines Inhalts am Testanfang und nochmals derselben am Testende.

Wie die RTR-Kurve eines Langfilms aus dem Arthouse-Bereich zeigt, lässt sich die RTR-Kurve nicht nur live beobachten, sondern auch ex-post für beliebige Subgruppen drucken wobei dann auch die interessierenden Szenen oder markante Wendungen im Handlungsablauf gut darstellen und interpretieren lassen.

Trotz der Informationstiefe, die ein gut geplanter RTR bieten kann, sollte er zumindest bei AV-Medien in anschließende Fokusgruppen eingebunden werden, um auch die Dimensionen beurteilen lassen zu können, die erst nach dem Film möglich sind.

Car Clinics in virtuellen Welten.

Missverständnisse und Potenziale



Der Diplom-Psychologe Dr. Bastian Zwissler ist Project Manager International Market Research beim Spiegel Institut Mannheim

Stellen Sie sich vor, ein Automobilhersteller erwägt, sein gut etabliertes und seit Jahren bewährtes Markengesicht für zukünftige Modelle deutlich zu verändern. Während die Designer diese geplante Neuausrichtung als willkommene Abwechslung und Herausforderung betrachten könnten, wird Entscheidungsträger besonders die Frage umtreiben, ob eine solch tiefgreifende Veränderung auch im Sinne der Stamm- und Neukunden ist. An diesem Punkt wäre die klassische Maßnahme die Durchführung einer Car-Clinic. Konkret hieße das für den Beispielfall: Der Hersteller müsste zur Beantwortung seiner Forschungsfrage mehrere Designentwürfe ausarbeiten, in 1:1- Modelle umsetzen und sie unter höchsten Sicherheitsvorkehrungen auf eine Reise rund um den Globus schicken. Die Kosten und der zeitliche Aufwand wären enorm. Und eine effizientere Lösung vermutlich willkommen.

Eine mögliche Alternative zur klassischen Car-Clinic wird seit geraumer Zeit heiß diskutiert und wurde auch wiederholt erprobt: die virtuelle Car-Clinic. Das physische Modell eines Fahrzeugs wird

hierbei durch eine virtuelle Darstellung ersetzt. Als 'virtuell' werden dabei 2D- als auch 3D-Stimuli verstanden. Ziel dieser Art der Designtesting ist das Erschaffen einer 'Virtuellen Realität' (VR). Derartige computergenerierte Welten sind in ihrer visuellen Qualität – bei gleichzeitig sinkenden Preisen – besonders seit den frühen 2000er-Jahren stets besser geworden. Sowohl in der Grundlagen- als auch in der angewandten Forschung gibt es zahlreiche Hinweise, dass Menschen virtuelle Welten als realitätsnah wahrnehmen und in sie eintauchen (Immersion), selbst dann, wenn keine 3D-Brille (Head-Mounted Display) eingesetzt wird. Kritischer Aspekt für erfolgreiche Immersion ist die Verbindlichkeit der virtuellen Welt. Wird eine Szenerie als verbindlich erlebt, wird der Betrachter Teil des Ganzen. Mit genau dieser Zielvorstellung sollte auch eine virtuelle Clinic konzipiert werden.

Bereits in den 1990er Jahren resümierte das Spiegel Institut im Abschlussbericht zu einer VR-Clinic, dass "nach dem Abwägen aller Für und Wider mehr für den Einsatz virtueller Verfahren zum Zweck einer 'vorgezogenen' (im Sinn von früh, nicht als

Ersatz) Design-Clinic spricht, als dagegen". Der Bericht nannte auch einige wenige Einschränkungen, die sich meist auf die Textur von Oberflächen bezogen ("Lackqualität im Vergleich schwer zu bewerten.", "Reifenprofil wirkt künstlich."). Diese Einschränkungen gelten auch fast 20 Jahre danach noch in ähnlicher Form und sind teilweise im Kern mit der Methode verwachsen. Trotz gewisser Kostenvorteile, sowie höherer zeitlicher und inhaltlicher Flexibilität konnten sich VR-Clinics bisher nicht durchsetzen. Grund dafür sind aber nicht die o.g. Detailprobleme. Es ist das Fehlen eines physischen Objektes und eines haptischen Eindrucks, das vielen Entscheidungsträgern als eine ungünstige Grundlage für eine informierte Designbewertung durch Laien erscheint. Es wird häufig ins Feld geführt, dass das unwägbare Risiko bestehe, ungeschulten Teilnehmern werde eine hohe Abstraktionsleistung abverlangt. Außerdem könnten virtuelle Reize grundsätzlich nie die gleiche Begeisterung entfachen wie ein reales Designmodell, besonders hohe Skalenwerte – offenbar gefragt beim Management - seien also nur in einer 'echten' Clinic zu erzielen.

Aus unserer Sicht ist unter vielen Gesichtspunkten durchaus eine Diskrepanz wahrnehmbar zwischen Forschungsbefunden zur VR in verschiedenen Settings und ihrer Verbreitung und Bewertung in der Praxis. Ein Grund hierfür liegt möglicherweise in falschen Erwartungen an die Methode. Eine VR-Clinic kann kein gleichwertiger Ersatz für eine klassische Clinic sein. Versuche, reale Fahrzeuge und Fahrzeugmodelle schlicht durch deren digitale Variante zu ersetzen, sind

eigentlich von vornherein zum Scheitern verurteilt. Stattdessen müssen die Ansätze unterschiedlich gedacht werden.

Entscheidend ist: Wer erwägt, eine virtuelle Clinic einzusetzen, sollte dies gezielt tun. Eine virtuelle Car-Clinic wird am Ende höchstwahrscheinlich weniger Kosten verursachen als eine klassische Car-Clinic, beispielsweise weil keine 1:1-Claymodelle angefertigt und transportiert werden müssen. Dieser Vorteil ist aber schnell vergessen, wenn das Management des Auftraggebers den erhobenen Daten misstraut. Enttäuschungen dieser Art können aus unserer Sicht am besten durch eine sorgfältige methodische Beratung vermieden werden. Diese sollte nach Möglichkeit früh im Projektablauf einsetzen und zunächst zentrale Erkenntnisinteressen ermitteln. Im Zweifel kann dies dazu führen, dass von einer virtuellen Car-Clinic abgeraten wird und doch eher altbewährte Methoden oder Misch-Ansätze empfohlen werden. Des Weiteren sollten so viele Stakeholder wie möglich einbezogen werden. Dies hilft, Vorbehalte früh zu identifizieren und Erwartungen am Ende auch erfüllen zu können.

Das virtuelle Stimulusmaterial stellt einen weiteren kritischen Aspekt dar. Speziell bei 3D-Darstellungen variieren Qualität und Kosten erheblich, werden 2D-Stimuli (Fotos oder Renderings) eingesetzt, spielen Faktoren wie Bildperspektive, Lichteffekte oder Verzerrungen eine immense Rolle. Auch bei der verwendeten Projektionstechnik gibt es eine Vielzahl von Möglichkeiten. Vielen erscheint die aufwendigste Variante als diejenige, die den Testteilnehmern auch den besten Eindruck ermöglicht. Häufig reichen hier

jedoch einfachere Lösungen völlig aus, beispielsweise, weil das volle Potential eines High-End-Beamers bei dem erforderlichen Projektions- und/oder Bildformat ohnehin nicht ausgereizt werden könnte. Das eigentliche Potential virtueller Clinics liegt in der Möglichkeit, ganz neue Forschungsfragen stellen und beantworten zu können. VR-Clinics sind sehr früh im Entwicklungsprozess und auch kurzfristiger durchführbar. Die Testung einer Vielzahl von Reizkonstellationen ist ohne großen Mehraufwand und ohne Einbußen bei der Standardisierung umsetzbar. Aufgrund des geringeren organisatorischen Aufwands können schwierige Stichproben besser realisiert oder auch längsschnittliche Studiendesigns (beispielsweise mit Pre-Test oder zeitlicher Trennung von qualitativer und quantitativer Befragung) erwogen werden. Neue Locations können erschlossen und so – bei vergleichsweise übersichtlichen Kosten – Premium-Venues gestaltet werden, bei denen nicht grauer Messe-Teppich und weiße Messe-Stellwände hinter Messe-Mietpflanzen versteckt werden.

Das Spiegel Institut Mannheim hat ein Konzept für virtuelle Designtests entwickelt, das den Fokus genau auf die genannten Aspekte richtet: Die experimentelle Pre-Clinic (EPC). EPCs sollen kleine Zukunftswelten erschaffen und auf diesem Wege auch ein Problem angehen, das allen Clinics – virtuellen und realen – gemein ist: Zeit. Clinics müssen ihrer Zeit per definitionem voraus sein. Die meisten Clinic-Teilnehmer sind aber keine Spezialisten für Zukunftstrends oder – designs, für aufkommende Moden oder Technologien. Sie bewerten das Gesehene mit Blick auf die Gegenwart oder

vergangene Erfahrungen. Die Zukunft virtueller Clinics liegt also auch darin, wie gut sie es schaffen, Clinic-Teilnehmer den Blick in die Zukunft richten zu lassen.

Die Vermessung des Unbewussten



Spitzer und Küppers sind Geschäftsführer der heart.facts Gesellschaft für Emotionsforschung in Köln.

Der Psychologe Oliver Spitzer berät und forscht seit seiner Zeit bei Agenturen wie GREY Worldwide oder Changing Minds im Bereich der angewandten wissenschaftlichen Emotionsforschung.

Der gelernte Diplom-Kaufmann Markus Küppers arbeitet seit 20 Jahren zunächst im Agentur-, dann im Marktforschungsumfeld, davon seit 2009 als einer der Geschäftsführer der september Strategie & Forschung GmbH in Köln, der Gründungsgesellschafterin von heart.facts.

Viele unserer alltäglichen Kaufentscheidungen basieren nicht auf rationalen Überlegungen, sondern auf unbewussten Emotionen. In der Marketing-Praxis werden Emotionen vor allem intuitiv eingesetzt und lassen sich mit dem herkömmlichen Instrumentarium der Werbewirkungsforschung kaum objektiv erfassen. Dieser Beitrag zeigt, welche Methoden im Rahmen der modernen Emotionsforschung angewendet werden und welche Ergebnisse sie zum Beispiel beim Testen der emotionalen Performance von TV-Spots liefern.

"Words are merely references to something invisible". Diese Aussage von Tor Nørretranders [1] verdeutlicht, dass die Aufgabe psychologischer Marktforschung nicht nur im Zuhören liegen kann, sondern vor allem im Verstehen von schwer

beziehungsweise gar nicht verbalisierbaren Informationen liegen muss. Inzwischen existieren neue Möglichkeiten, das emotionale Innenleben direkt zu vermessen. Im Folgenden möchten wir zeigen, was apparativ derzeit möglich (und sinnvoll) ist, und wie nah die Forschung, in Kombination mit Tiefenpsychologie an das mystische Unbewusste der Kunden herankommt.

Das Vokabular dazu gründet auf dem Verständnis, dass es zwei seelische "Emotionssysteme" oder Instanzen gibt, die gleichzeitig in uns schlummern: ein langsames, häufig bewusstes "I" und ein schnelles "Me", das vollkommen unbewusst agiert. Unser "Me" trifft in Sekundenbruchteilen zahllose Entscheidungen, die unser tägliches Leben bestimmen – weitgehend automatisch,

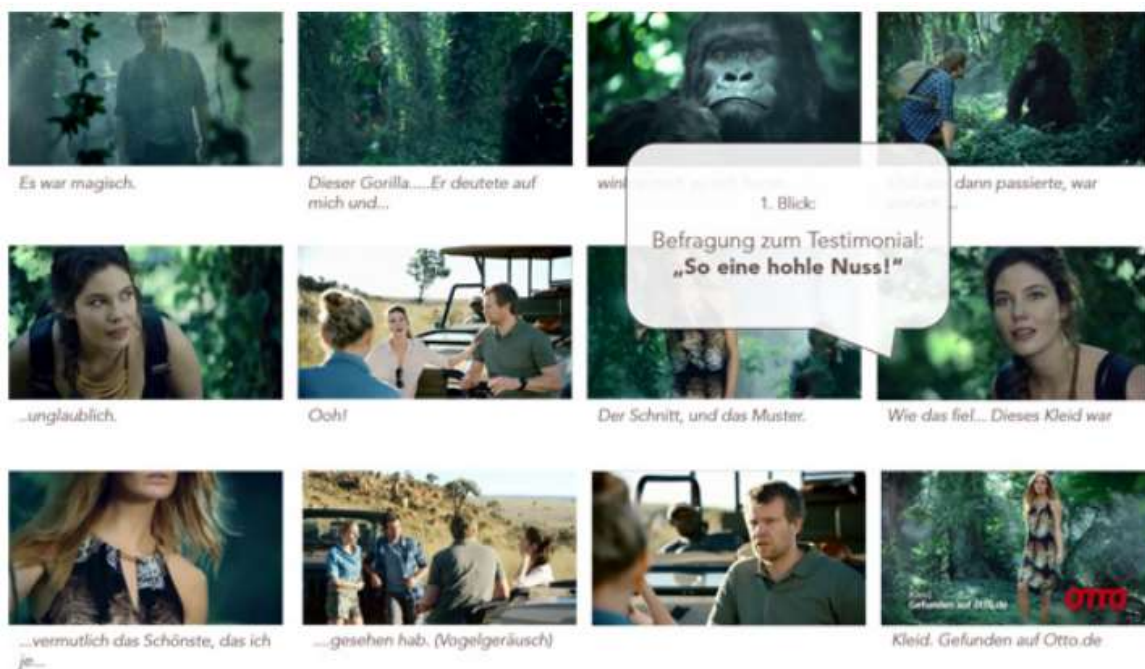
denn nur wenige Emotionen dringen als merkliches Gefühl bis in unser Bewusstsein vor. Auch unsere Kaufentscheidungen basieren – ob wir das wollen oder nicht – auf unbewussten emotionalen Einschätzungen. Das gilt sogar für den Kauf einer neuen Heizungsanlage, und für 'kleinere' Anschaffungen wie zum Beispiel Süßigkeiten oder Mode ohnehin.

Ich Jane, Du Tarzan: warum das Marketing Emotionen liebt, aber kaum versteht.

Nehmen wir an, Sie sind weiblich und sehen abends einen Krimi im Fernsehen.

Sekunden später bewegt sich ein Model in einem atemberaubenden Kleid durch das schwüle Grün. Ihr Kopf sagt jetzt wahrscheinlich: "Typisch Werbung." Aber was sagt Ihr Herz beziehungsweise Ihr "Me" in genau diesem Moment? Vielleicht: "Ich Jane, du Tarzan?" Würden Sie sogar Ihr iPad zücken und online nach dem Kleid suchen?

Die Online-Verkäufe des Kleids können sich sehen lassen, also funktioniert der Spot vermutlich auch gut. Aber warum genau? Werden Sie von den ersten Dschungel-Bildern gefesselt oder irritiert? Finden Sie



TV-Spot "Magic Moments: Kleid" von Otto (Bild: heart.facts)

Kurz bevor der Mörder zuschlägt, kommt eine Werbepause. Eigentlich wollen Sie jetzt den Kanal wechseln, aber da sehen Sie Menschen, die durch den Dschungel schleichen und einem Gorilla begegnen. Irgendetwas an dieser Szenerie weckt Ihr Interesse und Sie bleiben dran. Nur einige

das Kleid echt so toll oder würden Sie viel lieber den süßen Gorilla knuddeln? Identifizieren Sie sich mit dem Testimonial oder halten Sie sie für eine hohle Nuss? Ist die Abschlusszene wirklich witzig oder eher ein Abtörner? Solche emotionalen Prozesse sind für den

Erfolg des Spots entscheidend, den Teilnehmern einer Befragung aber selbst meist gar nicht bewusst. Auch typische Aussagen aus dem Neuromarketing wie "das Belohnungszentrum wurde aktiviert", können diese Fragen nicht beantworten. Für einen direkten Zugang zu unserem unbewussten "Me" müssen wir also einen anderen Weg einschlagen; den der Kombination tiefenpsychologischer und psychophysiologischer Methoden.

Emotionen sprechen Körpersprache

Psychische Aktivitäten beeinflussen gleichzeitig den Blutdruck, das Schwitzen, die Mimik und viele andere physiologische Prozesse. Emotionen lassen sich daher – weitaus einfacher und zuverlässiger als im Gehirn – anhand ihrer körperlichen Auswirkungen nachweisen. Statt eines MRT von der Größe eines Kleinwagens benötigt man dazu nur einen mobilen Bio-Recorder mit einigen Sensoren. Hinter der einfach handhabbaren Technik steht allerdings eine Kombination ausgeklügelter apparativer und statistischer Verfahren.

Unsere Emotionen sind komplexe psychophysiologische Prozesse, die man ganzheitlich und im Zusammenspiel betrachten muss. Nur die Aktivierung messen, nur die Mimik filmen – damit wird man dem vielschichtigen emotionalen Geschehen nicht gerecht. Um emotionale Prozesse eindeutig nachzuweisen, nutzt heart.facts eine Kombination verschiedener Messverfahren. Deren Daten werden in ein übersichtliches Gesamtbild aller körperlichen Reaktionen übersetzt. So entstehen objektive Ergebnisse, die für die Praktiker in Unternehmen und Agenturen direkt verwertbar sind.

Was uns ins Schwitzen bringt

Jeder Mensch weiß, dass man feuchte Hände bekommt, wenn man aufgeregt ist. "Emotionales Schwitzen" findet vor allem an den Händen und Füßen statt, wo es besonders viele Schweißdrüsen gibt. Wenn diese salzigen Schweiß absondern, verändert sich die elektrische Leitfähigkeit der Haut (elektrodermale Aktivität), was mit geringem technischem Aufwand sehr gut gemessen werden kann.

Was unser Herz höher schlagen lässt

Starke Emotionen lassen unser Herz "bis zum Hals pochen". Der enge Zusammenhang von Herzschlag und Emotionen ist eines der bekanntesten Körpersignale in der Emotionsforschung und wird – ähnlich wie beim ärztlichen Gesundheitscheck – mit einem EKG gemessen. Löst ein TV-Commercial zum Beispiel Vertrauen aus, führt das zu einer kurzfristigen Senkung der Herzrate und ungleichmäßigen Abständen zwischen den einzelnen Herzschlägen. Ein gleichmäßig rasender Puls weist dagegen auf eine emotionale Belastung hin. Auch die Stärke der Durchblutung, die mit einem optischen Sensor an der Fingerkuppe gemessen wird, ist in der Werbewirkungsforschung von Bedeutung: Wenn ein Reiz unser Interesse weckt, weil er uns begeistert (oder abschreckt), dann schwankt das Pulsvolumen stärker.



Die implizite Emotionsmessung erfasst und erkennt Muster aus über 20 psychophysiologischen Datenströmen. (Bild: heart.facts)

Hautleitfähigkeit oder Puls(volumen) sind wichtige Indikatoren für die Erregung eines Probanden, lassen alleine aber noch keine Aussage zu, ob diese Reaktion positiv oder negativ zu bewerten ist. Ekel und Begeisterung aktivieren ähnlich stark und sehen zum Beispiel ziemlich gleich aus, wenn man nur die Hautleitfähigkeit betrachtet. Es braucht also mehr, um eine Emotion – und nicht nur körperliche Aktivierung – zu messen.

Das innere Lächeln

Im Alltag können wir Stimmungen wie Wut und Angst aus der Mimik unseres Gegenübers gut erkennen, wenn sie stark genug sind. In der TV-Serie "Lie to Me" liest Dr. Cal Lightman den Verbrechern die verräterischen Emotionen förmlich an der Nasenspitze ab. Doch im Marketingkontext

hilft das wenig, weil Werbemittel wie zum Beispiel Online-Banner in der Regel auch für das geschulteste Auge oder die schnellste Kamera keine sichtbaren mimischen Reaktionen hervorrufen. heart.facts bedient sich der nicht sichtbaren Mikroexpressionen im Gesicht unserer Probanden. Diese spiegeln Emotionen wider, die sich unterhalb der Wahrnehmungsschwelle abspielen, weil Sensoren bereits die elektrische Spannungsveränderung der Gesichtsmuskeln messen. So lässt sich vor allem die Richtung (Valenz) einer emotionalen Reaktion einordnen, also ob wir ein Testimonial innerlich anlächeln, bei einem Packshot die innere Nase rümpfen oder bei einer Aussage die innere Stirn runzeln.

Was uns grübeln lässt

Weckt ein radikal neues Kreativkonzept unser Interesse und müssen wir über eine Werbebotschaft erst mal nachdenken, steigt unsere Gehirnaktivität im Bereich der schnellen Betawellen. Solche EEG-Messungen liefern Hinweise auf anstrengende mentale Verarbeitungsprozesse. Diese empfinden wir als unangenehmen emotionalen Stress. Folgt in diesem kritischen emotionalen Zustand nicht schnell eine positive Aufladung, sind wir kaum mehr bereit, uns mit den Inhalten eines Werbemittels zu befassen.

Welche Emotionen im Marketing wichtig sind

Die Basisemotionen nach Paul Ekman sind allgemein bekannt. Doch dieses Konzept stammt aus der klinischen Psychologie und ist für die Marktforschung wenig hilfreich. Denn wer will mit seiner Werbung schon Angst, Wut oder Ekel auslösen? Speziell für den Marketingkontext hat heart.facts daher aus vielen tausenden gemessenen Daten von Werbemittel-, Produkt- und Markentests ein präzise differenziertes Emotions-Spektrum mit sieben sogenannten emotionalen KPI abgeleitet und verifiziert.

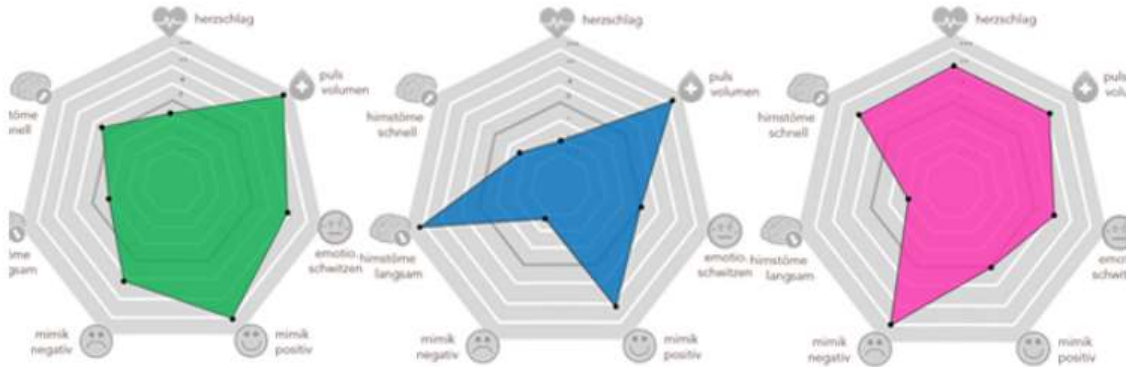
1. Vertrauen
2. Attraktion
3. Nähe
4. Sympathie
5. Relevanz
6. Skepsis
7. Stress

Diese Prozesse laufen gleichzeitig ab – der

Mensch empfindet zu einem gegebenen Zeitpunkt nicht eine Emotion, sondern mehrere, und dies schon in der allerersten Sekunde. So können wir eine neue Bekanntschaft attraktiv (zum Beispiel sexy) finden, würden aber gleichzeitig einen längeren Urlaub mit ihr ausschließen, weil wir vielleicht nicht wüssten, was wir mit ihr zu reden hätten (geringe Relevanz, hohe Skepsis, geringe Nähe) – obwohl wir diesem Menschen grundsätzlich vertrauen würden, weil wir ihn für zuverlässig halten (kein negatives Vertrauen). All diese Facetten können wir in der gleichen Sekunde erleben, und sie widersprechen sich nicht. Wenn wir uns aufgrund des spontanen Eindrucks dagegen entscheiden, mit diesem Menschen ein Bier zu trinken, dann ohne Überprüfung unserer Grundannahmen ("langweilig"). Der erste Eindruck ist, gerade im Konsumalltag, der bleibende.

Dieser blitzschnelle und hochkomplexe Prozess lässt sich mit der psychophysiologischen Methode detailliert abbilden.

Beispiele für die psychophysiologischen Muster der emotionalen KPI:



Beispiele: Attraktion, Vertrauen und Skepsis (heart.facts)

Wenn zwei sich streiten: Wie wir unsere inneren Konflikte bewältigen

Die psychophysiologischen Messungen liefern ein klares Bild der emotionalen Reaktionen einer Testperson. Sie zeigen, welche Elemente eines Spots Vertrauen schaffen, sympathisch sind und welche für Stress oder Skepsis sorgen. Doch eine zentrale Frage ist noch unbeantwortet: Warum kommt es zu diesen gemessenen Mustern? Was geschieht in unserem Unbewussten, das unseren Körper so reagieren lässt und nicht anders? Mit einem Fragebogen lässt sich das natürlich nicht so ohne weiteres beantworten.

Die für die Messungen kausale "Story" analysiert heart.facts durch tiefenpsychologische Interviews, die im Anschluss an die Messungen mit denselben Teilnehmern stattfinden. Unsere Analysen zielen darauf ab, die typischen inneren Konflikte der Probanden ans Licht zu bringen und zu prüfen, wie gut die Marke

oder die Kommunikation die Gegensätze zwischen unserem triebhaften, schnellen Kern ("Me") und unserer äußeren, sozialisierten und langsameren Benutzeroberfläche ("I") überbrücken kann. Denn erst wenn eine Marke nicht nur einen relevanten seelischen Konflikt zwischen "Me" und "I" aktiviert, sondern auch zu behandeln verspricht, liefert sie einen psychischen Mehrwert. So werden die gemessenen psychophysiologischen Reaktionen der Probanden tiefenpsychologisch erklärbar – ein unverzichtbarer Bestandteil der Beurteilung und Optimierung von Werbemitteln.

So kann die psychophysiologische Messung eines TV-Spots zum Beispiel einen negativen Wendepunkt in der emotionalen Performance eindeutig nachweisen. Doch die Zahlen können nichts darüber aussagen, was man an dieser Stelle ändern müsste. Reicht es, wenn man etwas am Text ändert? Wird das falsche Thema, der falsche innere Konflikt angesprochen. Sollte

man besser ganz auf den Zwischenschnitt verzichten? Erst die Interviews klären auf, was genau die Probanden irritiert. Nun lässt sich wesentlich besser entscheiden, wo und wie die emotionale Dramaturgie optimiert werden kann. Die psychophysiologische Messung braucht eine erklärende Komponente. Die tiefenpsychologische Analyse braucht objektive Kennzahlen. Gemeinsam entfalten die Methoden ihr volles Potenzial.

Praxisbeispiel: Kleid statt Gorilla – emotionale-Analyse eines TV-Spots von Otto

Zurück zu unserem Otto-TV-Spot. Der Hamburger Versandhändler wirbt seit 2013 erfolgreich mit der Leitidee "magic moments" in verschiedenen Variationen: Eine Frau erlebt eine denkwürdige Situation, hat dann aber nur Augen für ein Kleidungsstück von Otto. Der von heart.facts untersuchte Film aus dieser Reihe zeigt die Mitglieder einer Dschungel-Safari, die mitten im Regenwald einem stattlichen Gorilla begegnen – bis plötzlich das Kleid einer im Urwald auftauchenden Frau die Protagonistin in seinen Bann zieht.

Die Probanden wurden zur Messung ihrer impliziten Reaktionen auf den Spot und anschließende Tiefeninterviews eingeladen. Dabei wurde zunächst die emotionale Aktivierung im Normalzustand gemessen, die von Mensch zu Mensch unterschiedlich ist. Während der Darbietung des Spots auf einem Monitor wurden dann mittels mehrerer Sensoren fortlaufend die feinen impliziten Reaktionen der Probanden gemessen und im Nachgang zu den einzelnen emotionalen

KPI verdichtet.

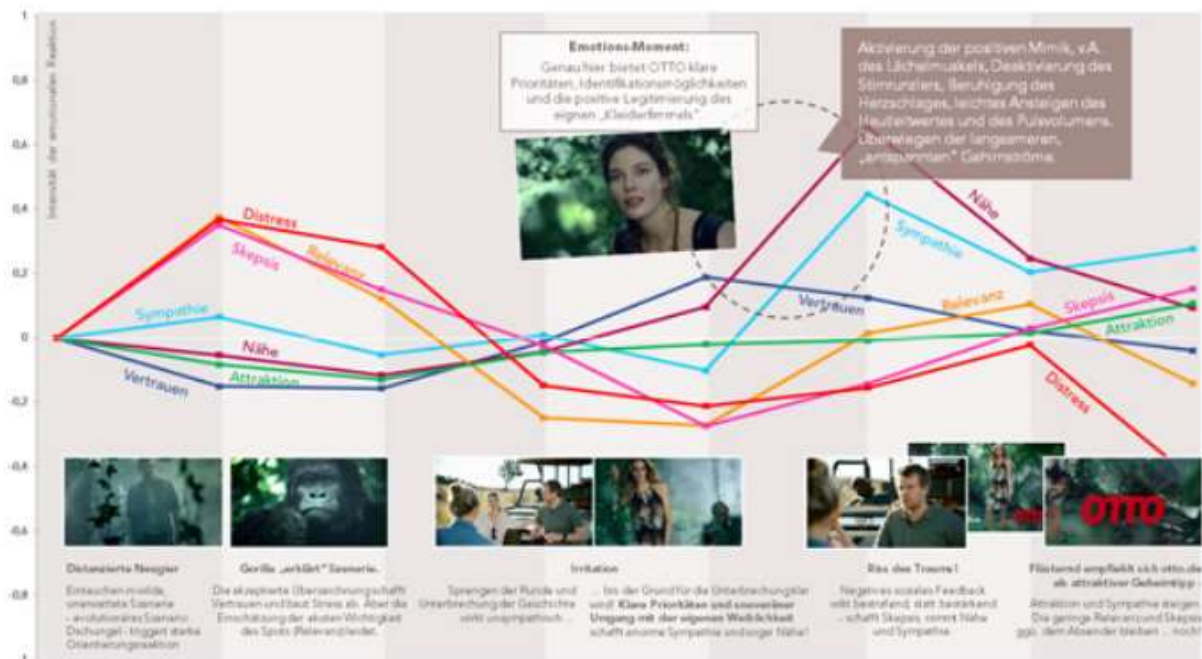
Zunächst löst der Spot mit seiner unerwarteten Dschungel-Szenerie steigende Werte für Stress, Relevanz und Skepsis aus, die im weiteren Verlauf schnell aufgefangen und positiv gewendet werden. Der Gorilla ist friedlich und stellt eine positive Beziehung zu den Protagonisten her. Diese zunächst starke Orientierungsreaktion, die gut in die emotionale Dramaturgie eingebunden wird, ist ein gelungener Einstieg, der nachweislich Neugier weckt. Der Spot hat gute Chancen, beachtet zu werden.

Jetzt wechselt der Handlungsort weg vom Dschungel zur realen Erzählebene: Die Unterbrechung der Geschichte führt zu einem Abfall aller emotionalen KPI in den negativen Bereich. Im Gespräch empfinden die Probanden diesen dramaturgischen Wechsel als unsympathisch. Mit der Inszenierung des Kleides in der zweiten Hälfte des Spots gehen die emotionalen KPI, insbesondere Nähe und Sympathie, dann weit in den positiven Bereich. Auch die Relevanz steigt wieder, denn man weiß jetzt, worum es geht und kann das Geschehen einordnen.

Es folgt der emotionale Höhepunkt: die Protagonistin wendet ihre Aufmerksamkeit nur noch dem Kleid zu. Der Dschungel als Catwalk – was rational eher Verwirrung stiftet, funktioniert emotional offensichtlich hervorragend. Die Stimme der Vernunft stellt der Heldin an dieser Stelle ein zuerst schlechtes Zeugnis aus: Wie kann man in so einem epochalen Moment an Kleider denken! An dieser Stelle haken die Tiefeninterviews nach. Es zeigt sich: Die weiblichen Zuschauer

können sich gut mit der Heldin identifizieren. Sie ist eine moderne "Jane", eine starke Dschungel-Amazona, die provozierend zu ihrem "Kleiderfimmel" steht und ihre eigenen Prioritäten selbstbewusst durchsetzt.

Das negative soziale Feedback der männlichen Safari-Teilnehmer stoppen den Traum abrupt. In der Schluss-Sequenz erscheint nun erstmals die Marke Otto, der es gelingt, Sympathie und Attraktion wieder zu steigern. Doch es bleibt auch



Damit stellt sie eine attraktive Projektionsfläche für Frauen dar, diese fühlen sich entlastet, weil so auch ihr eigener "Kleiderfimmel" Absolution erfährt. Während solche Wendungen in rationalen Befragungen gerne als klischeehaft zerrissen werden, entfalten sie – wie die Emotionsmessung und die Tiefeninterviews belegen – im Unbewussten doch die beabsichtigte Wirkung: Sympathische Nähe und Relevanz entstehen, also gute Voraussetzungen für den Kauf des Produkts.

Dann folgt der erneute Szenenwechsel: Die bis dahin erfolgreiche, emotionale Performance des Spots fällt deutlich ab. Die Tiefeninterviews erklären warum:

eine gewisse Skepsis zurück. Wie ein solcher dramaturgischer und emotionaler Bruch vermieden werden kann, zeigt Otto im aktuellen Spot "Mantel", dem es hervorragend gelingt, seine Heldin auch am Ende als die emotionale Siegerin zu inszenieren.

Fazit

Jede Kaufhandlung basiert auf unbewussten, emotionalen Bewertungen und (Vor-) Entscheidungen. Je mehr wir diese verstehen, desto besser können wir den Erfolg von Kampagnen beurteilen und systematisch optimieren. Die Kombination aus hoch differenzierten psychophysiologischen und bewährten

tiefenpsychologischen Verfahren bietet ein schnelles, äußerst valides und praktikables Sensorium für unser Unbewusstes – egal ob es um TV-Spots, Online-Banner, Anzeigen, Neuprodukte, Verpackungen und vieles mehr geht.

[1] Tor Nørretranders "The User Illusion"

Wearables-Technologie in der Marktforschung



Dipl.-Sozialwissenschaftler Felix Fischer ist seit 2005 für [eye square](#) tätig. Als Leiter des Team Project Services obliegt ihm die Koordinierung der Analysen von Eye-Tracking-Studien und deren Qualitätssicherung. Sein Team verantwortet die Beschaffung, Aufarbeitung und Analyse des Datenmaterials und unterstützt die Projektteams bei statistischen und grafischen Auswertungen.

Psychologe Gareth Tuck ist Director des UK Office bei [eye square](#). Sein Team führt globale Werbewirkungs- und User Experience-Studien für Kunden der Branchen FMCG, Consumer Electronics und E-Commerce durch.

Die Wearables-Technologie ist zweifelsohne eine der neuen, großen Trends. Aktuell wird der Consumer-Markt von den kleinen mobilen Messgeräten für den Privatgebrauch, die etwa die Leistung beim Joggen analysieren, geradezu überflutet. "Fitnesstracker" oder "Smartwatches" für das Handgelenk fanden im Weihnachtsgeschäft reißenden Absatz. Auch in der Telemedizin kommen Wearables zur Eigenbeobachtung und

Übermittlung medizinischer Daten zunehmend zur Anwendung. Nützliche, kleine Geräte, die mittlerweile in großer Stückzahl und zu erschwinglichen Preisen für den breiten Markt produziert werden. Den Technologieunternehmen ist es gelungen, Wearables durch Miniaturisierung, Verlängerung der Betriebszeit, intuitive Bedienbarkeit und die Bereitstellung meist kabelloser Schnittstellen hochattraktiv für die Nutzer zu gestalten.



Eye-Tracker verschiedener Generationen (Bild: eye square)

Wearables-Technologie in der Marktforschung

Doch können diese mobilen Messgeräte auch in der Marktforschung genutzt werden? Um dieser Frage nachzugehen, nahmen Kantar und eye square 2014 im Rahmen einer Studie sowohl das Google Glass und den digitalen Fotoapparat Autographer von OMG Life, der, getragen an einem Halsband, automatisch alle zehn Sekunden ein Umgebungsbild erstellt, als auch die Eye-Tracking-Brille Pupil genauer unter die Lupe. Im Fokus stand dabei die Frage, wie sehr das per Außenkamera aufgenommene Videomaterial der drei Devices für die Marktforschung tauglich ist und welches die besten Resultate liefert.

Der Studienaufbau

eye square passte die Forschungsfragen und das Studiendesign jeweils an die

spezifischen Möglichkeiten der Geräte an und ließ Testpersonen mit dem Ziel, Smartphones und Tablets in Geschäften zu begutachten, durch eine Shopping-Mall in Großbritannien laufen. Konkret sollte geprüft werden, inwiefern sich folgende Forschungsfragen mit den drei Devices beantworten lassen:

- Wie viel Zeit wird in welchem Laden verbracht?
- Kontaktzeiten mit bestimmten Geräten (Smartphones/Tablets) und Marken
- Kontaktzeit mit Point-of-Sale-Materialien
- Erfassung von Interaktionen mit Produkten und Verkäufern

Die Ergebnisse

Bei der Analyse des Bildmaterials wurden folgende Unterschiede zwischen den Geräten deutlich:

- Bildqualität:

Pupil und Google Glass liefern dank HD-Kameras eine deutlich höhere Bildqualität als die Aufnahmen des Autographers. In puncto Informationsgehalt des Bildmaterials ist das Ergebnis hingegen weniger eindeutig. Aufgrund der Positionierung in Brusthöhe nehmen Smartphones und Tablets bei den vom Autographen erstellten Bildern öfter eine „zentrierte“ Position ein. Verglichen dazu liefern Google Glass und Pupil aufgrund der auf Augenhöhe angebrachten Kamera öfter „abgeschnittene“ Produktbilder. Allerdings ist die auf Augenhöhe montierte Kamera von Vorteil, wenn es um die Sichtbarkeit der Produktmarken geht

- Vergleich Pupil mit Devices ohne Eye-Tracking-Daten

Ob Probanden ein Produkt „anschauen“, wird beim vom Autographen und Google Glass aufgezeichneten Material (beide ohne Eye-Tracking-Daten) anhand der Produktposition im Bild bestimmt. Ist zum Beispiel ein Tablet zentriert im Bild zu sehen, kann davon ausgegangen werden, dass es sich auch im Blickfeld des Probanden befindet und demnach angeschaut wird. Um herauszufinden, inwieweit das aufgezeichnete Eye Tracking dazu beiträgt, genauer festzustellen, ob Probanden ein Produkt tatsächlich anschauen, wurde das vom

Pupil aufgezeichnete Video sowohl mit als auch ohne die überlagerten Eye-Tracking-Daten ausgewertet. Die Analyse ergibt eine erstaunlich hohe Trefferquote. Zu beachten ist jedoch, dass die Auswertung mit Eye-Tracking wesentlich genauere Schlüsse darauf zulässt, was genau an welchem Gerät wann angeschaut wurde

- Vergleich Google Glass und Autographen

Google Glass und Autographen liefern in Bezug auf die Produktwahrnehmung im Laden leicht unterschiedliche Ergebnisse. Ursache dafür sind die Aufnahmeperspektiven (aufgrund unterschiedlicher Kamerapositionen der Geräte), damit verbundene Fehler in der Codierung sowie Unterschiede in der Interpretation

- Post-Shopping-Interview

Bei der Bewertung der drei Devices für ihre Tauglichkeit bei einem Media-Assisted-Interview erweist sich Pupil als Sieger, da die Eye-Tracking-Brille nicht nur ein Video aufnimmt, sondern insbesondere auch den Blickverlauf anzeigt. Google Glass schneidet mit seinen Videoaufzeichnungen am zweitbesten ab. Die alle zehn Sekunden geschossenen Bilder des Autographen können beim Post-Shopping-Interview auch sehr wertvolle Aufschlüsse liefern, sind den kontinuierlichen Videoaufzeichnungen von Pupil und Google Glass aber unterlegen.

Das Fazit

Implizite Messmethoden wie FET (Facial Emotional Tracking), GSR (Hautleitwertmessung), EEG (Herzfrequenz) und insbesondere das Eye-Tracking (Blickaufmerksamkeitsmessung) sind in der Praxis mittlerweile fest etabliert. Die Marktforschung profitiert dabei davon, dass Messgeräte immer mobiler, vernetzter, schneller, kleiner und anwendungsfreundlicher werden. Sie ermöglichen es der Marktforschung, niederschwellige und unbefangene Testsituationen zu schaffen. Für den breiten Consumer-Markt produzierte Wearables bieten vor diesem Hintergrund enorme Chancen für die Marktforschung. Ihr großer Vorteil sind insbesondere die hohen Stückzahlen und somit erschwinglichen Preise. Wearables sind zwar nicht explizit auf die Durchführung von Studien im Bereich der Marktforschung zugeschnitten, können aber durchaus dafür verwendet werden. Im Test ermöglichen alle drei Devices Aufschlüsse und Analysen über Aufenthaltsdauer im Laden, Kontaktpunkte der Probanden mit Produktkategorien und Produkten sowie Sichtfelder von Marken und Point-of-Sale-Materialien. Interaktionen mit Verkäufern und Geräten werden ebenso festgehalten. Die Testergebnisse zeigen, dass die Wearables-Technologie eine interessante Erweiterung für die Marktforschung darstellen kann.

Weiterführende Infos:

Alex Johnson, Kantar Operations/UK: „What do Wearable Devices Bring to Market Research? Are weareable devices like Google Glass viable alternatives to mobile handsets for market research?“ von der ESOMAR 2014 als Best Paper ausgezeichnet.

Warum sich apparative Forscher nicht unbedingt in der Komfortzone bewegen



Dr. Beate von Keitz gründete 1983 das Institut für Kommunikations-Forschung in Saarbrücken. Im Gespräch mit marktforschung.de beschreibt sie die technischen Entwicklungen apparativer Forschungsmethoden und warum sie eine Kombination mit Interviews präferiert.

marktforschung.de: Frau von Keitz, wann sind Sie zum ersten Mal mit apparativen Forschungsmethoden in Berührung gekommen?

Beate von Keitz: Prof. Kroeber-Riel, Gründer des Instituts für Konsum- und Verhaltensforschung an der Universität des Saarlandes, ist ja ganz früh in mein Leben gekommen. Auf der Suche nach einer Möglichkeit, Konsumentenverhalten zu studieren, war das der einzige Weg – es gab sonst weder bei den Psychologen noch bei den Betriebswirten dieses Lehrangebot, das war komplett neu. Und bei Prof. Kroeber-Riel war ich nicht nur in den Veranstaltungen, sondern ich durfte ein Praktikum am Lehrstuhl machen. Ja, und da waren Methoden wie das Eye-Tracking, der Aktivierungstest über Hautwiderstandsmessungen oder der Tachistoskop-Test im täglichen Einsatz. Ich war auch in Forschungsprojekte involviert wie das legendäre HOBA-Experiment zur "emotionalen Konditionierung". Es ging da um den Nachweis, dass Verbraucher durch emotionale Werbung beeinflusst werden. Das war damals eine ganz neue Theorie in der BWL und heiß umstritten ...

marktforschung.de: Gegen welche Widerstände mussten Sie seinerzeit ankämpfen?

Beate von Keitz: Als ich mit dem Institut startete, da war bei vielen Auftraggebern schon ein hohes Interesse an den apparativen Verfahren vorhanden und auch eine Offenheit. Ich war ja auch die erste, die die ganzen Verfahren in einem Institut anbot und die die Methoden bei Prof. Kroeber-Riel quasi "gelernt" hatte. Aber ich hörte schon öfter mal den Vorwurf, dass die apparativen Methoden mit ihren Geräten das Verbraucherverhalten verfälschen würden, zum Beispiel die Brille beim Eye-Tracking. Schon damals lernten wir in Validierungsstudien – Dr. Lachmann von Philips war hier ein toller Sparringspartner – dass selbst die großen Brillen, die wie Skibrillen aussahen, das Wahrnehmungsverhalten nicht verändern. Aber da existierten öfters irrationale Vorurteile, mancher Auftraggeber "mochte" die Geräte einfach nicht. Dazu kam: Die Möglichkeit, für apparative Tests die Geräte bei den Herstellern zu mieten, gab es damals noch nicht, und der Kaufpreis der Geräte war immens hoch.

Da musste ich dann auch noch die Bank überzeugen, dass auch eine Mutter mit zwei Kleinkindern erfolgreich eine Firma gründen kann und dass sie kreditwürdig ist...

Von großen Geräten, die immer kleiner wurden

marktforschung.de: Sie haben Ihr Institut 1983 gegründet. Was waren die wesentlichen Milestones der Firmenhistorie? Welche technologischen Entwicklungen haben Ihr Unternehmen maßgeblich geprägt?

Beate von Keitz: Als ich angefangen habe, war das noch die Zeit der Großrechner. Wir mussten eigene Software für die Auswertung der apparativen Tests programmieren lassen, und vieles musste auch manuell ausgewertet werden. Insofern habe ich ganz tolle Dinge erlebt: Die PCs wurden immer leistungsfähiger und preisgünstiger, das Angebot an Standardsoftware wurde immer umfangreicher, die Auswertung wurde immer leichter, und vieles konnte man zumindest halbautomatisieren. Es gab eigentlich ständig Neues. Auch das Problem der großen und auffälligen Geräte hat sich in Wohlgefallen aufgelöst. Es kamen berührungslose Systeme auf, die der Verbraucher praktisch nicht mehr sieht. Eines ist für mich aber ganz witzig: Jetzt folgt wieder ein Revival der "Glasses" – also der Eye-Tracker, die der Verbraucher aufsetzen muss. Aber mit denen kann er sich frei bewegen. Das ist zum Beispiel in POS-Studien oder in Mailing-Tests wichtig. Die Diskussion ist in jedem Fall viel sachlicher geworden. Und das ist ein fortlaufender Prozess:

die Geräte werden immer kleiner, leichter und komfortabler, und die Unterstützung durch die Software wird immer umfassender. Heute gerade haben wir ein nettes kleines Päckchen aufgemacht – es enthielt unseren Eye-Tracker, den wir neu gekauft haben.

marktforschung.de: Apparative Techniken haben eher den Ruf eines marktforscherschen "Nischenprodukts". Wie verorten Sie Eye-Tracking, Tachistoskop und Co. im Methodenmix?

Beate von Keitz: Ich persönlich bin ein absoluter Fan des Interviews – mir ist aber auch sehr bewusst, dass man viele Dinge nicht abfragen kann. Das heißt: man kann natürlich immer fragen, und man bekommt auch meist Antworten. Aber diese Antworten können schlichtweg falsch sein. Immer wenn es um Fragen der Aufmerksamkeit, Wahrnehmung und Zuwendung geht, sind die apparativen Verfahren meilenweit überlegen.

Sie reichern das Interview in seiner Aussagefähigkeit deutlich an. Sie machen zum Beispiel die Bewertung von Werbemitteln sicherer, indem sie weitere Indikatoren der Werbewirkung in den Test integrieren: Aufmerksamkeit, Zuwendung, emotionale Reaktionen. Und sie sind absolut nützlich, wenn es um die Diagnostik und um die Optimierung geht. Können die Benefits nicht überzeugen, weil sie in der Werbung gar nicht beachtet werden oder weil sich die Verbraucher nicht intensiv genug damit beschäftigen oder weil sie sie nicht verstehen oder weil die Argumente nicht glaubwürdig wirken? Die Kombination "Interview + apparativ" bringt nach unserer Erfahrung einen

signifikanten Mehrwert, weil die Evaluation und die Diagnostik viel präziser werden.

Unsere Verortung der Methoden im täglichen Einsatz:

- Eye Trackings: Tests für visuelle Kommunikation (Print, Digital), Packungen, POS
- Aktivierungs-Tests über elektrodermale Reaktionen: Werbung, speziell TV-Spots, Visuals, Claims
- Tachistoskop-Tests: Plakat, Erscheinungsbilder, Packungen – gern auch in Kombination mit einem Eye Tracking

"Apparativ" bedeutet aber immer auch einen Mehraufwand. Es sind Geräte nötig, es geht um Investitionen und um Organisation, man braucht qualifizierte Manpower, und man hat oft auch mit technischen Störungen zu kämpfen. Kein Wunder, dass es viele Forscher vorziehen, sich auf Interviews zu beschränken. Der Einsatz von apparativen Verfahren ist anspruchsvoll, und man bewegt sich als apparativer Forscher im Institut nicht immer in der Komfortzone, gerade wenn man es gut machen möchte. Insofern sind die apparativen Verfahren vielleicht ein Nischenprodukt, aber ihr Einsatz steigt doch kontinuierlich an.

Warum Digitalisierung und Big Data eine Chance sein kann

marktforschung.de: Was ist für die Aussagekraft apparativer Methoden wichtig?

Beate von Keitz: Extrem wichtig ist der Test mit Umfeldern. Wenn man ohne Umfeld testet, dann verändert das die Ergebnisse. Verbraucher nehmen Kommunikation ja

immer auch in einem Umfeld wahr. Wichtig ist etwas, was ich als "strukturell realistisches Setting" bezeichne. Ein Beispiel dazu: Eye-Trackings am Regal sind eine spannende Sache, um die Suchprozesse der Verbraucher aufzuzeigen. Nun möchte ich vielleicht auch wissen, was den Verbrauchern dabei durch den Kopf geht. Wenn ich die Probanden nun auffordere, parallel zur Suche ihre Gedanken zu verbalisieren, dann verändere ich damit gleichzeitig den Ablauf des Suchprozesses. Menschen sprechen langsamer als sie gucken; mit dem Sprechen wird der Wahrnehmungsprozess verlangsamt und verändert. Das Eye-Tracking zeigt dann also nicht mehr den natürlichen Wahrnehmungsprozess. Solche Dinge muss man wissen, um zu validen Ergebnissen zu kommen. Auch Tests in den "falschen" Medien, zum Beispiel der Printtest auf dem Bildschirm, sollte man immer hinterfragen. Und eine solide Stichprobe ist natürlich ein Must.

Wir sind sehr froh, dass sich bei unserer Methodik immer wieder positive Korrelationen zwischen den apparativen Ergebnissen und den Reaktionen im Markt zeigen – bis hin zu hohen Korrelationen mit den Absatzzahlen.

marktforschung.de: Die fortschreitende Digitalisierung bringt nicht nur eine schier unfassbare Menge an Daten hervor, sondern auch Produkte wie Wearables, denen man eine hohe Bedeutung im Hinblick auf die zukünftige Gewinnung von Informationen attestieren kann. Wie sehen Sie diese Entwicklung?

Beate von Keitz: Wenn man mit apparativen Verfahren arbeitet, dann ist man es gewohnt, dass man mit hohen Datenmengen und mit ständig neuen Technologien umgehen muss und umgehen darf. Auch hier gilt wie bei Big Data sonst auch: Es sind intelligente Lösungen und kluge Köpfe gefragt, um aus Daten Erkenntnisse und Learnings zu machen.

Das Interview führte Claas Lübbert

Sprechen die Daten für sich selbst?



Dr. Helene Karmasin leitet das Institut Karmasin Behavioural Insights mit Sitz in Wien. Sie ist auf qualitative Marktforschung und semiotische Analysen spezialisiert und berät internationale Markenartikel- und Dienstleistungsunternehmen bei Strategie, Umsetzung und empirischen Tests.

Die Ergebnisse von apparativen Verfahren besitzen einen besonderen Reiz: Sie signalisieren, dass hier etwas objektiv gemessen wird und zwar eine Reaktion des Kunden, die unabhängig davon ist, was dieser explizit berichtet.

Dies sind zwei starke Argumente, eigentlich auch zwei Träume der empirischen Forschung: So zu messen, wie das Naturwissenschaften tun und eine Möglichkeit zu haben, zu erkunden, was Konsumenten geheim antreibt, ohne dass ihnen dies explizit bewusst ist. Das erste wird stark durch die Art der Darstellung gefördert: Kurven, Prozessverläufe, Steigungsquotienten, Punktwolken, farbige Bilder, wie bei Messungen der Gehirnaktivität. Die Frage, was diese Daten genau aussagen, wodurch sie bedingt sind, wie ihre Intensität einzuschätzen ist, wird dabei selten gestellt. Sie geben sicher eine Auskunft über das, was Befragte nicht explizit erzählen, aber worüber genau? Und wenn sie unbefriedigend sind: Wie kann man sie optimieren?

Wir sind der Meinung, dass diese Art der Messung nur dann sinnvoll ist, wenn sie von einem Modell der Verhaltenssteuerung begleitet ist und wenn sie Möglichkeiten bietet, das Gemessene zu interpretieren und zu optimieren. Wir kombinieren bei unseren Forschungsprojekten apparative Verfahren, die wir je nach Projekt auswählen, mit Ansätzen der Behavioural Economics und der Semiotik, die die notwendigen theoretischen Modelle liefern.

Dazu ein Beispiel:

Behavioural Economics wird bekanntlich inzwischen in vielen politischen und sozialen Fragestellungen eingesetzt, zunehmend auch in Fragen der Unternehmensführung und des Marketings. In der empirischen Erforschung von Kundenreaktionen, also im unmittelbaren Bereich der Marktforschung sind dagegen nur Ansätze erkennbar.

Wir haben daher eine Reihe von Methoden entwickelt, um hier gezielt zu messen. Bekanntlich geht Behavioural Economics davon aus, dass Menschen keineswegs immer rational handeln und entscheiden, dass sie ihr eigenes Verhalten schlecht prognostizieren können, dass sie mehr als sie denken von Situationen und den verbalen und visuellen Rahmungen einer Botschaft abhängig sind.

Sie legen in vielen alltäglichen Entscheidungen das von Kahnemann so benannte System 1 zugrunde: Ein schnelles, intuitiv, anstrengungslos vorgehendes Entscheidungssystem, das stark von situativen Reizen und von den zeichenhaften Rahmungen einer Botschaft beeinflusst ist.

Dieses System müssen wir messen. Wenn wir Befragten alle Zeit der Welt geben, um über ihre Entscheidungen zu brüten und sie zu begründen, wenden wir uns nur an ihr rationales Entscheidungssystem. Dies ist speziell im Bereich von Pretests wichtig.

Das folgende Beispiel stammt aus einem Pretest für Packungen:

Eine Marke und zwar eine Preiseinstiegsmarke einer großen Lebensmittelkette, die nahezu in allen Produktkategorien Produkte anbietet, wollte ihre Packungen einem Relaunch unterziehen. Wie immer stellt sich damit die Frage: Wie stark oder schwach entfernt vom derzeitigen Auftritt soll so ein Relaunch sein? Ist er zu weit, erkennen die Kunden ihre Marke nicht, ist er zu schwach, fragt man sich, warum man überhaupt eine Änderung vorgenommen hat. Es geht also darum zu erkennen:

- Bei welchem Entwurf behält die Marke ihre Anmutung bei, sodass sie mühelos wiedererkannt wird?
- Welcher Entwurf reichert die Bedeutung der Marke an?
- Welcher Entwurf sagt aus, dass diese Marke in dieser Produktkategorie ein hervorragender Vertreter ist?

Marken kommunizieren ihre Bedeutung bekanntlich, indem sie sie in visuelle und verbale Codes übersetzen, sie bauen also einen semiotischen Unterschied auf. Semiotisch gesehen heißt die Frage also: Wie weit kann der Markencode gedehnt werden? Wie gut behauptet sich eine Packung im Code der Produktgattung? Gelingt es ihr den Rang eines Prototyps zu erreichen, also zu kommunizieren, dass sie ein hervorragender Vertreter der Produktkategorie ist? (Siehe dazu: Produkte als Botschaften [1])

Wie kann man diese Entscheidung messen?

Trügerisch ist es in jedem Fall, sich auf Gefallens- und explizite Bewertungsurteile zu verlassen. Apparative Verfahren stellen in diesem Fall eine gute Testmöglichkeit dar. Nur welche sollen es sein? Nicht wirklich zielführend sind Verfahren, die die emotionale Aktivierung messen – eine Packung kann hoch aktivierend sein, aber nicht an die Markeninhalte anschließen.

Nehmen wir die Erkenntnisse der Behavioural Economics ernst, so ist klar, dass die Packung am besten funktionieren wird, die am schnellsten von System 1 erfasst werden kann. Wir müssen also die schnellen Reaktionen messen, noch bevor System 2, das rational vorgeht, wirksam wird.

Dafür gibt es mehrere Möglichkeiten. Sehr gut bewährt hat sich die Darbietung über ein Onlinetachistoskop.

Dieses misst die erste Phase der Wahrnehmung, beginnend mit einer Expositionszeit von 1/100 Sekunde bis zu einer 1/30 Sekunde, der Zeit, in dem die bewusste Wahrnehmung einsetzt. Wir exponierten die einzelnen Entwürfe in diesem Zeitraum und stellten jeweils fest:

- Von welcher Marke war die Packung?
- Welche Produktgattung war das?
- Wie hat sie gewirkt?

Als Merkmale wurde hier die wesentliche Dimension der Testmarke gewählt. Durch dieses Verfahren zeigte sich: Ganz schnell wiedererkannt wurde der Entwurf, der besonders nahe an dem bisherigen Erscheinungsbild der Packung lag – rein von den Daten her hätte man also diesen Entwurf nehmen müssen.

Von den kommunizierten Merkmalen her wies er allerdings keinerlei Fortschritt auf und es bestand auch eine gewisse Schwäche, was die schnelle Identifikation der Produktgattung betraf. Dies war vor allem bei der Milch der Fall – eines der wichtigsten Produkte der Marke.

Nun kann man einwenden, das das irrelevant ist: Im realen Supermarkt steht die Packung natürlich neben den anderen Milchprodukten. Rational weiß also jeder, dass dies eine Milch ist. Nimmt man jedoch das Konzept der Prototypen ernst, so ist diese langsame Übermittlung der Produktkategorie (die man in diesem Fall

erst nach Lektüre der verbalen Inschrift erkannte) ein Manko.

In jeder Produktkategorie existieren bestimmt Farb-, Form- und Größencodes, visuelle Programme, Ästhetiken, die in Sekundenschnelle kommunizieren: Das ist ein hervorragender Vertreter, der Prototyp der Produktkategorie – sie inszenieren gewissermaßen das "Milchhafte" und das eben in den ersten Stadien der Wahrnehmung. (Siehe dazu: Verpackung ist Verführung [2])

Der Hersteller hatte also zwei Optionen:

Entweder er entschied sich für den Packungsentwurf, der sehr schnell die Marke vermittelte - dann aber musste die Packung in Bezug auf ihre produktspezifischen Codes überarbeitet werden oder er entschied sich für den zweiten Sieger – eine Packung, die ausgezeichnet in der Vermittlung der Markenmerkmale war, die ganz schnell die Produktgattung kommuniziert, damit also nahe an dem Prototyp war, die aber nicht ganz so schnell die Marke erkennen ließ. Dies konnte man entweder durch eine Überarbeitung aussteuern oder man entschied sich bewusst dafür und propagierte die erneuerte Marke durch massive Darstellungen in der Werbung.

[1] Karmasin, Helene: Produkte als Botschaften. mi-Fachverlag. Landsberg am Lech: 2007

[2] Karmasin, Helene. Verpackung ist Verführung. Die Entschlüsselung des Packungscodes. Haufe Verlag. Freiburg 2015

Impressum

marktforschung.de

Smart News Fachverlag GmbH
Max-Ernst-Str. 4
D-50354 Hürth (bei Köln)

Tel.: +49 (0)2233 - 71 004-0
Fax: +49 (0)2233 - 71 004-99

E-Mail: [info\(at\)marktforschung.de](mailto:info(at)marktforschung.de)

Registergericht: Amtsgericht Köln
Registernummer: HRB 60274
USt.-IDNr.: DE257625438

Geschäftsführung:

Dipl.-Kfm. Lars-O. Gayk

Inhaltlich verantwortlich nach § 10 Absatz 3 MDStV: Dipl.-Kfm. Lars-O. Gayk

Titelbild: © Julia Tim - fotolia.com