

Die Frage nach der Werbewirkung wird nicht einfacher, wenn man sie für verschiedene Plattformen stellt. Ein ungewöhnliches Setting, nämlich eine Experiment hat die australische Wissenschaftlerin Karen Nelson-Field mit der Studie „Not all reach is equal“ im Auftrag des TV-Werbevermarkters Screenforce durchgeführt. Der Markt- und Mediaforscher Dirk Engel hat sich die Vorgehensweise für planung&analyse angeschaut.

**Experimente
helfen**

Screenforce untersucht, wie ein Werbespot in TV und auf digitalen Plattformen wirkt

**Wirkung
zu verstehen**

Rund 25 Milliarden Euro werden in Deutschland jährlich für Werbung ausgegeben. Im Wettbewerb um diese Gelder bemühen sich die Medien, den Werbekunden ihre Stärken zu demonstrieren. Dazu dienen ihnen Studien, die teilweise mit hohem Aufwand durchgeführt werden und bei denen sich verschiedene Medienhäuser zusammenschließen. Screenforce ist ein Zusammenschluss deutscher, österreichischer und Schweizer Fernsehsender, der gerade eine Untersuchung veröffentlicht hat. „Not all reach is equal“ lautet der Titel. Ihr liegt eine Frage zugrunde, die sich viele Werbungtreibende stellen: Wirkt ein Werbespot in gleicher Weise, wenn man ihn im Fernsehen sendet oder auf einer der digitalen Plattformen wie YouTube, Facebook oder Instagram präsentiert? Mehr und mehr Werbebudget wandert vom TV zu diesen Digital-Netzwerken, gerade junge Zielgruppen lassen sich dort gut erreichen. Doch ist die Situation, in denen die Spots gesehen werden, unterschiedlich: der große Flachbildschirm im Wohnzimmer oder das kleine Smartphone-Display, auf das man unterwegs schaut. Unterscheidet sich auch die Wirkung der dort platzierten Werbung? Soviel sei schon mal vorweggenommen: Sie tut es.

Wirkungsnachweis von neutraler Stelle

Den objektiven Nachweis für die unterschiedliche Wirkungsqualität sehen die TV-Vermarkter von Screenforce in ihrer aktuellen Studie. Objektiv deshalb, weil diese Studie von einer Wissenschaftlerin durchgeführt wurde, die bei der Entwicklung der Methode freie Hand hatte. Karen Nelson-Field ist Marketing-Professorin aus Australien, einem Land, das in der Forschung um „evidence-based Marketing“ führend ist. Das methodische Grundgerüst hat sich bereits in Studien in Australien, Großbritannien und den USA bewährt. Nun wurde es für den deutschsprachigen Markt adaptiert. Dabei betonen die Auftraggeber von Screenforce betonen, dass sie keinen Einfluss auf Methode oder gar Ergebnisse genommen haben, obwohl es sich um die bisher teuerste Einzelstudie der TV-Vermarkter handeln soll. Was wurde genau gemacht?

Die Studie ist ein Experiment: Unter kontrollierten Bedingungen wurde der Einfluss bestimmter Variablen überprüft, nämlich der Ausspielung von Werbung auf

verschiedenen Plattformen. Normalerweise werden experimentelle Studien in Wissenschaft und Marktforschung im Labor durchgeführt, also in einer künstlichen Umgebung – etwa wie in einem Marktforschungs-Studio –, in der alle Faktoren kontrolliert werden können. Das Labor wird hier ersetzt durch eine Smartphone-App: Sie dient der Kommunikation mit den Versuchsteilnehmern, liefert die Umfragen, zeigt das Stimulus-Material – die Werbespots im jeweiligen Umfeld – und

nimmt außerdem kontinuierlich Fotos von den Gesichtern der Probanden auf, um ihren Blickverlauf zu analysieren.

Die App als Kernstück für das Experiment

2.500 Menschen nahmen im Januar und Februar dieses Jahres mit Hilfe dieser App an der Studie teil, 80 Prozent davon aus Deutschland, der Rest zu gleichen Teilen aus Österreich und der Schweiz. Für sie verlief die Untersuchung in fünf Schritten, nach der Rekrutierung mussten sie als erstes die App herunterladen. Dann sollten sie die Sozialen Netzwerke (also Facebook, Instagram und YouTube) ganz normal mit ihren eigenen Login nutzen – doch wurde technisch dafür gesorgt, dass nur die zuvor ausgewählten Test-Webespots gesehen wurden. Diese waren so in die Plattformen integriert, dass den Teilnehmern die Manipulation nicht aufgefallen ist. Bei der Fernsicht wurden die Teilnehmer gebeten, ihr TV-Gerät einzuschalten, es mit ihrem Smartphone zu verbinden (ein entsprechender Adapter wurde bereitgestellt) und dann aus einem kleinen Angebot verschiedener Programme auszuwählen. Während die ausgesuchte Sendung angeschaut wurde, sorgte die App dafür, dass die manipulierte Werbung ausgestrahlt wurde. Dieser immense technische Aufwand diente dem Zweck, unter kontrollierten Bedingungen jedem Teilnehmer ein spezielles Set von TV-Spots zu zeigen, entsprechend dem experimentellen Design der Studie. Jeder Proband sah bestimmte Spots und andere nicht. Dadurch erübrigte sich eine separate Kontrollgruppe: Jeder Teilnehmer war in der Experimentalgruppe für die Spots, die er gesehen hat, und in der Kontrollgruppe für jene, die er nicht sah.

Die App zeichnete ebenfalls genau auf, wie die Werbung gesehen wurde – wie lange und wie vollständig war sie auf dem Bildschirm zu sehen? Nach der Rezeption der Werbung wurden die Probanden (wieder mittels App) befragt. Nicht nur die Werbeerinnerung (Recall) wurde erhoben, auch eine Kaufsituation wurde simuliert. In einer Art Shop konnte man zwischen fünf Produkten in verschiedenen Kategorien wählen, darunter jeweils ein beworbenes. So wurde die Kaufabsicht gemessen.

Die Forscher betonen, dass in dem gesamten Prozedere darauf geachtet wurde, eine möglichst natürliche Rezeptionssitua-



Die Wissenschaftlerin Karen Nelson-Field hat die Studie durchgeführt

Not all reach is equal

Fragestellung: Wie wirkt Bewegtbildwerbung auf unterschiedlichen Medienkanälen? Wie wirken Einfach-, Doppel- und Mixkontakte? Wie wirken Werbekontakte langfristig?

Zielgruppe: Erwachsene 18+ Jahre / Teilnehmer mussten Nutzer des jeweiligen Mediums sein.

Stichprobe: n = 2500 / 80 % in Deutschland, je 10 % in Österreich und der Schweiz. Methode: experimentelle in home-Studie mit Online-Befragungen, angesteuert durch eine Testr-App

Feldzeit: Januar bis Februar 2020

Medienkanäle: TV, BVDO, YouTube, Facebook, Instagram Testmaterial: ca. 30.000 View mit 12 Testspots aus unterschiedlichen Branchen, **KPIs:** STATS, Werbeerinnerung, Aufmerksamkeit sowie verschiedene Device-Merkmale wie Viewability (Pixel, Play-Time), Coverage und Sound.

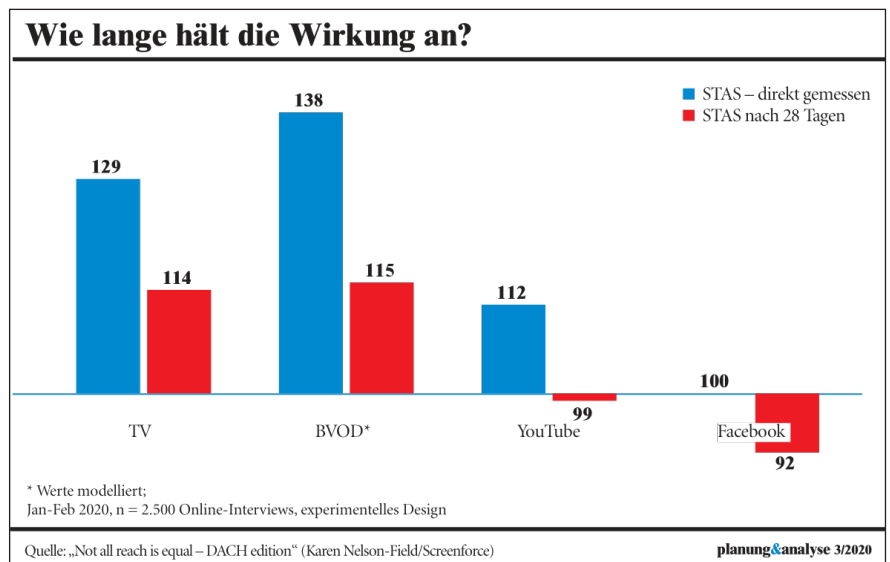
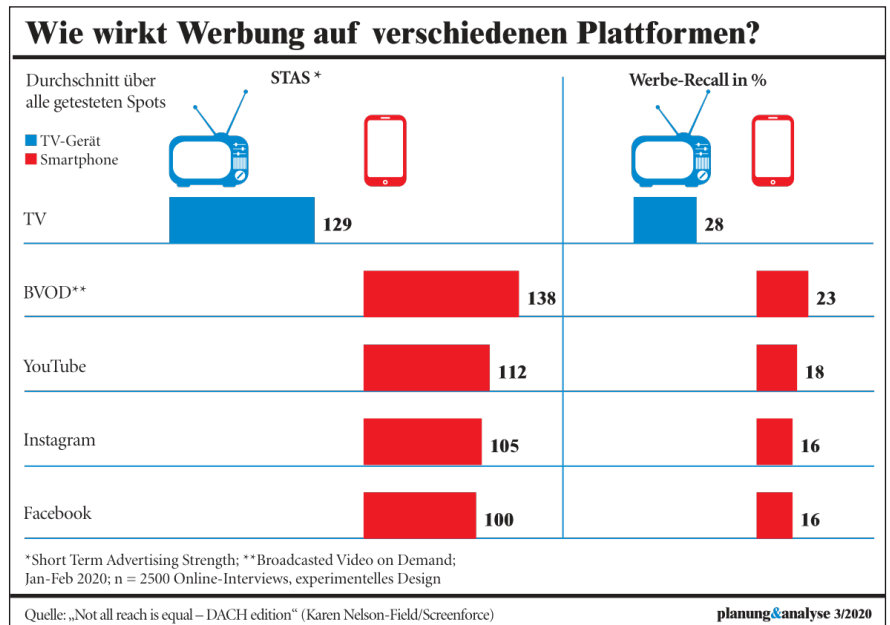
tion zu gewährleisten. Der Proband schaute sich nur solche Plattformen an, die er auch sonst nutzen würde, er sah in Facebook und Instagram seinen gewohnten Newsfeed und schaute die TV-Werbung auf dem heimischen Gerät. Auch der simulierte Shop ist ein Versuch, näher an einen richtigen Kaufakt heran zu kommen, als mit einer bloßen verbale Abfrage der Kaufpräferenzen.

Am schwierigsten war jedoch die Auswahl der untersuchten Spots. Sie mussten Produkte bewerben, die in allen drei deutschsprachigen Ländern bekannt waren. Des Weiteren sollten die werbenden Unternehmen durchaus Kunden der beteiligten Plattformen sein. Auch die Länge der Werbevideos musste variieren, die für die Digital-Medien typischen kurzen Spots inklusive. Zwölf Spots aus unterschiedlichen Produktkategorien wurden verwendet.

Wie hoch ist die Kaufwahrscheinlichkeit?

Kommen wir endlich zu der Analyse und damit zu den eigentlichen Ergebnissen der Studie. Untersucht wurde, ob dieselben Spots auf unterschiedlichen Plattformen verschieden gewirkt haben, das heißt, ob sich Unterschiede bei Werbeerinnerung und Kaufabsicht feststellen ließen. Der entscheidende Indikator ist dabei die Short Term Advertising Strength (STAS), die bereits in den 90er Jahren vom Werbeforscher John Philip Jones eingesetzt wurde. Dabei werden die Käufe, denen ein Werbekontakt vorausging, mit Käufen ohne vorherigen Kontakt in Beziehung gesetzt und indiziert. Ist die STAS über 100, so hat der Werbekontakt zusätzliche Käufe generiert. Während Jones mit echten Kaufdaten aus Verbraucher-Panels arbeitete, handelt es sich bei den „Käufen“ der aktuellen Screenforce-Studie um die in der simulierten Shopping-App getroffene Wahl. Die durchschnittliche STAS für die zwölf Werbespots liegt für das klassische TV bei 129 – anders formuliert: Ein Kontakt mit einem TV-Spot erhöht die Kaufbereitschaft um 29 Prozent. Würde die „normale“ TV-Sendung auf dem Smartphone gesehen (hier spricht man von BVOD - Broadcasted Video on Demand), ist der STAS sogar bei 138. Wurden die Spots bei YouTube angeschaut, ist der Wert mit 112 deutlich geringer. Noch schlechter sieht es bei Instagram (106) und Facebook (100) aus.

Das heißt überspitzt gesagt: Ein Werbekontakt bei Facebook hat die Käufe über-



haupt nicht beeinflusst, und auch bei den anderen Netzwerken passiert nicht viel.

Im Grunde sind diese fünf einfachen Zahlen das wichtigste Resultat des aufwändigen Experiments. Auch die Werbeerinnerung (Recall) zeigt das gleiche Bild: Im Durchschnitt erinnerten sich 28 Prozent an einen Spot, den sie im TV gesehen haben, während es nur 16 bis 18 Prozent bei YouTube, Instagram und Facebook waren. Wenn der TV-Content auf dem Smartphone gesehen wurde, erinnerten sich immer noch im Schnitt 23 Prozent an die dort platzierte Werbung.

Natürlich wurde noch tiefer analysiert. Zum Beispiel wurde neben dem Durchschnitt noch die Varianz der STAS-Werte ausgewertet: Bei YouTube und Instagram gab es nur zwei Spots, die eine STAS über 150 erreicht haben – für das klassische TV waren es immerhin sieben von zwölf.

Beim Mix profitiert Facebook von TV-Werbung

Eine Sonderanalyse konzentrierte sich auf eine Untergruppe der Versuchsteilnehmer, die zwei Kontakte mit den Spots in dem jeweiligen Setting hatten. Ein zweiter Kontakt erhöht nur geringem Maße STAS und Recall, doch TV hat nach wie vor die Nase vorn. In einem Sub-Sample wurde ebenfalls der Mix von TV und YouTube beziehungsweise Facebook analysiert. Hier zeigt sich das bekannte Bild: Der Kontakt via Facebook oder YouTube erhöht die Indikatoren kaum. Beim Mix profitieren Facebook und YouTube allerdings stark, aber TV kaum. Bei einigen Probanden wurde übrigens die Wirkungsmessung (also der Besuch im simulierten Shop) nach 14 und 28 Tagen wiederholt. Auch hier stimmt die Tendenz: Die Halbwertszeit von

TV-Werbung ist auch vier Wochen nach dem Kontakt noch hoch, während sie bei den anderen Plattformen deutlich schneller abnimmt. Die Messungen zur Blickregistrierung und der Sichtbarkeit der Werbung halfen Karen Nelson-Field und ihren Kollegen, diese Ergebnisse zu erklären. Die Dauer der Zuwendung wurde von ihnen als Indikator für Aufmerksamkeit gewertet und dabei haben das herkömmliche Fernsehen und YouTube bessere Werte als Facebook und Instagram. Je größer der Bildschirmanteil des gezeigten Spots ist (also nimmt er den ganzen Bildschirm ein oder gibt es ringsherum noch andere Elemente wie in einem Facebook-Feed), desto besser ist die Wirkung.

Die Ergebnisse der „Not all reach is equal“-Studie sind nicht nur interessant, sie können ökonomische Auswirkungen haben. Werbungtreibende und Mediaplaner könnten sie zum Anlass nehmen, ihre Investitionen in digitale Bewegtbild-Werbung zu überdenken. Selbst wenn aus Reichweiten-Gesichtspunkten die digitalen Plattformen eine gute Leistung bringen, weil sie viele Menschen in bestimmten Zielgruppen erreichen, so könnte die geringere Wirkung ein Nachteil sein.

Wer führt eine Debatte zur Werbewirkung?

„Steht Facebook jetzt eine Debatte über Werbewirkung ins Haus?“ fragt eine Überschrift in der Fachzeitschrift HORIZONT. Screenforce selbst holte einige Aussagen von Media-Experten in einem Video ein, das sie zusammen mit der Präsentation von Karen Nielsen-Field auf einer eigenen Website („What’s On“) veröffentlichten. Die Zitierten begrüßen die Studie und bekräftigen die Kernaussage: Werbung wirkt auf den unterschiedlichen Plattformen nicht in gleicher Weise. Allerdings wirken die Experten auch nicht überrascht – solche Unterschiede seien bereits aus anderer Forschung bekannt und würden schon in den Mediastrategien berücksichtigt.

Eine solch selbstbewusst vorgestellte Studie provoziert natürlich Kritik. Der gängigste Vorwurf: Bei dem Auftraggeber Screenforce sei ja von vornherein klar, was rauskomme. Eine „Prodomo-Forschung“ sei es nicht, betont Screenforce-Geschäftsführer Martin Krapf. Seine Organisation habe die Studie zwar finanziert, aber keinen Einfluss genommen. Die Methode hätten die australischen Wissenschaftler selbst

entwickelt und in anderen Ländern erprobt, für Deutschland seien nur einige Ergänzungen vorgenommen worden, etwa die Analyse des Werbe-Recalls und die Aufnahme von Instagram in die Riege der untersuchten Plattformen.

Google und Facebook kritisieren Schlüsse aus der Studie

Tatsächlich wird die Methode auch nicht von den deutschen Forschern bei Google und Facebook angegriffen. „Grundsätzlich halte ich experimentelle Forschung mit kontrollierten Bedingungen für gut“, meint etwa Torsten Müller-Klockmann, als „Marketing Science Lead DACH“ bei Facebook zuständig für alle Forschungsaktivitäten im deutschsprachigen Raum. Er kritisiert aber die Schlussfolgerung: Die Studie zeige nicht, dass TV besser als Facebook sei, sondern dass ein TV-Spot besser im TV wirke. Das liege aber daran, dass ein solcher Spot eben für das Medium optimiert wurde. Genauso müsse man Spots für Facebook optimieren: vertikal statt horizontal (damit sie besser im Feed des Smartphones erscheinen), sie sollten ohne Ton funktionieren, gegebenenfalls Untertitel verwenden und schnell auf den Punkt kommen. Während in einem Erzählmedium wie TV durchaus ein Spannungsbogen aufgebaut werden könne, müssten Aufmerksamkeit und Relevanz im Feed bei Facebook und Instagram nach wenigen Sekunden erregt werden. Müller-Klockmann ist sich sicher, dass solche optimierten Spots auf Facebook eine gute Wirkung erzielen, aber wahrscheinlich im klassischen TV schwächer sind. Die Operationalisierung beim Thema Mix-Kontakte hält er nicht für ausreichend: „In unseren Studien, etwa mit Consumer-Panel-Daten, sehen wir deutlich stärkere Synergie-Effekte“.

Ähnlich denkt auch Daniel Meyer, Director Market Insights bei Google: „Es ist schwierig, die Stärken von YouTube in einem kontrollierten Experiment abzubilden.“ Dazu zählen Faktoren wie präzises Zielgruppen-Targeting, Frequency Capping, die Möglichkeit für Nutzer, Werbung zu überspringen und der Mix von unterschiedlichen Spot-Längen innerhalb einer Kampagne. So bietet YouTube sogenannte Bumper Ads an, die gerade einmal sechs Sekunden dauern. Bei der Screenforce-Studie wurden hingegen nur für YouTube relativ lange Formate von 12

bis 30 Sekunden verwendet. Werbung auf YouTube sei nicht nur, einen Spot zu platzieren, sondern eine Kampagne zu entwickeln, die alle verfügbaren Stellgrößen nutze und deren Kreation für die Plattform optimiert ist. Meyer findet es auch schade, dass die Probanden in der Screenforce-Studie nur das klassische TV-Programm auf einem großen Screen gesehen haben. Denn 20 Prozent der Nutzungszeit von YouTube geschehe ebenfalls auf einem TV-Gerät. Meyer vermutet, dass dann die Wirkungsunterschiede weitgehend verschwinden würden.

Einen direkten Vergleich mit TV-Werbung scheue YouTube nicht. In der viel beachteten „Medienäquivalenz-Studie“ arbeitete Google mit den Forschern vom Institut Facit und dem TV-Vermarkter SevenOneMedia sowie der Media-Agentur Mediaplus zusammen. Die Ergebnisse zeichnen ein differenzierteres Bild des Verhältnisses von TV und digitaler Bewegtbild-Werbung. Skeptisch ist der Google-Mann auch bei der Messung der Kaufbereitschaft im simulierten Shop – eine solche künstliche Situation überschätze Effekte. Das bemängelt auch Kirsten Latour, Media-Chefin bei MCM Klosterfrau und im Vorstand des Verbandes der Werbungtreibenden OWM, in einem Screenforce-Video. Die STAS sei nur ein schwacher Indikator im Vergleich zum Langzeit-Return-on-Invest, der aus in Panels gemessenen Umsätzen berechnet wird.

Einig sind sich Kritiker und Freunde der Studie über die Forscherin: Karen Nelson-Field wird für ihren Ideenreichtum und ihre Präsentation gelobt, die jeder in eigens von Screenforce produzierten Videos abrufen kann. Vor dem Hintergrund der Forderung des OWM nach einer medienübergreifenden Wirkungsforschung, die hohen Qualitätsstandards entsprechen soll, ist die Studie „Not all reach is equal“ sicherlich ein gutes Beispiel für eine Studie einer Gattungsmarketing-Organisation. Dass es sich dabei um ein sozialwissenschaftliches Experiment handelt, ist ebenfalls ein gutes Zeichen. Obwohl diese Methode die gängige in der Wissenschaft ist, findet man sie in der angewandten Mediaforschung selten. Hier überwiegen Case Studies, Tracking-Studien und statistische Modellings. Experimente helfen hingegen, die Mechanismen der Werbewirkung besser zu verstehen. Dabei gibt es noch vieles zu Erforschen – genug Stoff für zukünftige Forschungsaktivitäten. ■