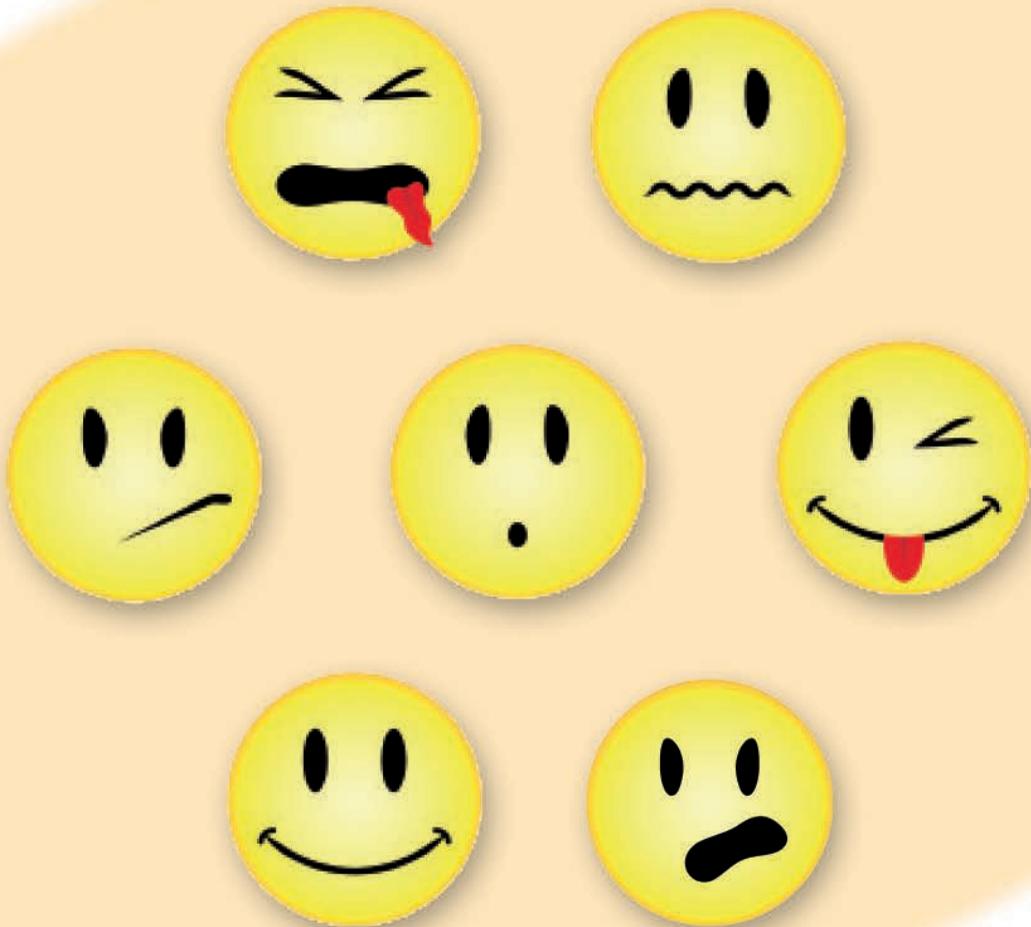


Sensorische Analyse: Methoden- überblick und Einsatzbereiche

Teil 5: Affektive und hedonische Prüfungen



Sensorische Analyse: Methodenüberblick und Einsatzbereiche – Teil 5: Affektive und hedonische Prüfungen

Affektive und hedonische Prüfmethode setzt man ein, um abzuschätzen, wie sehr den Konsumenten ein Produkt, ein Lebensmittelprodukt zusagt. Diese Methoden beruhen nicht auf analytischen, möglichst objektivierten Beurteilungen durch wenige, ausgewählte und gut geschulte sensorische Prüfpersonen, sondern sie erfassen die subjektiven, ganzheitlichen Bewertungen von aus dem Bauch heraus urteilenden Konsumenten. Affektiv leitet sich sprachlich vom lateinischen *affectus*, „Empfindung, Leidenschaft“ ab und bedeutet gefühlsbetont oder emotional. Hedonisch hat eine etwas eingeschränktere Bedeutung, es leitet sich vom griechischen *hēdoné* „Freude, Vergnügen, Lust“ und *hēdonikós* „zum Vergnügen gehörend“ ab, und bezieht sich somit auf eine spezielle Emotion nämlich die Emotion der Lust und der Freude bzw. ihrem Gegenteil der Unlust und dem Ekel. Die affektiven und hedonischen Prüfmethode zielen demnach auf die weitgehend unbewussten, emotionalen und gefühlmäßigen Beurteilungsmechanismen in Konsumenten ab. Sie wurden entwickelt, da sich in vielen Studien zeigte, dass Konsumenten nicht in der Lage sind, sinnvolle Aussagen über die Ursachen zu machen, die ihren positiven, neutralen oder negativen Beurteilungen zu Grunde liegen [1, 2]. Bei affektiven und hedonischen Untersuchungsmethode wird daher nicht nach den vermeintlichen Beweggründe für das Urteil gefragt, da das nur ad hoc Rationalisierungsversuche der weitgehend unbewusst getroffenen Entscheidung zur Folge hätte. Die Qualität und Aussagekraft der erhaltenen Daten ist weitaus besser, wenn Konsumenten sich während der Testsituation in einem intuitiven Beurteilungszustand befinden, als wenn sie in einen analytischen Denkraum gezwungen werden [3, 4].

Hedonisch-affektive Prüfmethode finden Anwendung bei der Neuentwicklung von Produkten[5], bei der Modifikation bzw. Optimierung bereits bestehender Produkte [6], aber auch beim kontinuierlichen Monitoring der sensorischen Qualität im Rahmen der Qualitätssicherung (Performance Tracking), das meist verknüpft wird mit einem Benchmarking d. h. einem systematischen Vergleich des eigenen Produkts mit dem erfolgreichsten Konkurrenzprodukt.

Im Wesentlichen gibt es zwei Ansätze im Bereich der affektiven und hedonischen Untersuchungsmethode - den introspektiven und den observationalen Ansatz. Im Folgenden werden diese beiden Ansätze kurz dargestellt.

1. Introspektive Methoden

Der introspektive Ansatz umfasst Tests, bei denen man Selbstberichte subjektiver Erfahrungen von Konsumenten durch Befragung erhält. Die Versuchspersonen versuchen dabei die Art und/oder Intensität ihrer subjektiven Wahrnehmungen auf verschiedene Skalen abzubilden (z. B. Punkte-Skalen oder VAS Visual Analog Scales). Es kann bei solchen Selbstberichten auch die zeitliche Dimension eingeschlossen werden, d. h.

die emotionale Beurteilung eines zeitlich sich erstreckenden Ereignisses, entweder online durch Echtzeit-Beurteilung oder retrospektiv. Es ist klar, dass bei introspektiven Tests nur bedingt ein Maß für die affektive oder hedonische Empfindung gewonnen wird, da diese durch die Bewusstmachung und den Versuch der Beschreibung und Quantifizierung kognitiv überformt und verzerrt wird [7]. Da die introspektiven Tests aber vordergründig leicht und ohne großen apparativen Aufwand durchgeführt werden können, sind sie vor allem in der Marktforschung und der Sensorik weit verbreitet.

1.1 Testmethoden

Man gliedert die introspektiven Methoden in Tests, die der Abschätzung von hedonischen Beliebtheit und in Tests, die der Beurteilung von Präferenzen dienen [8]. Bei den Beliebtheitstests (Akzeptanztests) wird versucht, ein Maß für die intuitive, gesamtheitliche Empfindung bei Konfrontation mit dem Produkt zu gewinnen, indem verschiedene Skalen angeboten werden, auf denen diese Empfindungen abgebildet werden können. Bei Präferenztests dagegen wird ermittelt, welches von zwei oder auch mehreren Produkten gegenüber dem/den anderen bevorzugt wird. Das bevorzugte Produkt muss nicht notwendigerweise eine hohe Beliebtheit aufweisen. Es können auch alle gegeneinander verglichenen Proben schlecht und nur das eine präferierte Produkt weniger schlecht beurteilt werden sollen.

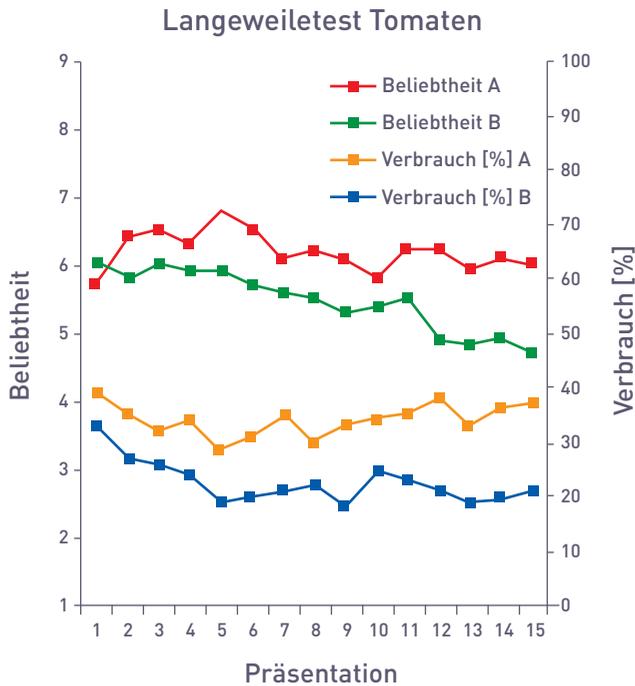
1.1.1 Spontaner Akzeptanztest

Beim spontanen Akzeptanztest wird der Testperson eine relativ geringe Menge der Prüfprobe zur Beurteilung vorgelegt und die Testperson soll sich rasch und intuitiv, aus dem Bauch heraus ein Urteil bilden. Wenn nicht nur die Gesamtakzeptanz der Produkte getestet wird, sondern auch die Akzeptanz einzelner sensorischer Merkmalsgruppen wie Geschmack, Geruch, Aussehen, Textur, dann ist es günstiger die Beurteilung der Gesamtakzeptanz an den Anfang des Tests zu stellen, um zu verhindern, dass die Konsumenten in eine analytische Prüfsituation gedrängt werden. Eine Problematik spontaner Akzeptanztests liegt darin, dass bei ihnen Produkte mit sehr einfachen, leicht zugänglichen, wenig komplexen sensorischen Profilen besser abschneiden als bei realitätsnaher Verzehrsmenge mit größerer Verzehrsmenge. Um dieses Problem, das zu vielen in sensorischer Hinsicht trivialen Produkten geführt hat, zu umgehen, wurden Testmethoden wie der Dauerakzeptanztest konzipiert, die eine realistischere Prüfweise ermöglichen.

1.1.2 Spontaner Präferenztest

Analog zum spontanen Akzeptanztest wird beim spontanen Präferenztest geprüft, welches von zwei oder mehreren Produkten intuitiv, aus dem Bauch heraus bevorzugt wird. Präferenztests sind als einfache Ausprägungen von Rangordnungstests aufzufassen.

Abb. 1: Darstellung der Daten von Langeweiletests mit zwei unterschiedlichen Tomaten-Sorten



1.1.3 Dauerakzeptanztest

Beim Dauerakzeptanztest wird den Prüfpersonen nicht nur eine kleine Menge des Prüfprodukts vorgelegt, sondern es wird das Produkt von den Prüfpersonen über eine gewisse zeitliche Dauer hin, die von wenigen Tagen bis zu Wochen reichen kann, beurteilt (Beispiel: Halber Liter eines Erfrischungsgetränks jeden Tag über 24 Tage hin). Neben der Beliebtheit kann auch die tatsächliche Verzehrmenge abgefragt und als Parameter der hedonischen Akzeptanz gewertet werden. Um zu vermeiden, dass die Prüfpersonen davon ausgehen, jedes Mal das gleiche Produkt zu beurteilen und damit versucht sein könnten, es unabhängig von ihren Empfindungen immer gleich zu beurteilen, ist es notwendig den Prüfpersonen zu sagen, dass die tatsächlich gleichen Produkte sich „leicht aber eventuell auch stärker“ voneinander unterscheiden können. So ist gewährleistet, dass die Prüfpersonen sich die Möglichkeit geben, die Beliebtheitsurteile schwanken zu lassen, sodass über die Zeit hin eine ab- oder zunehmende Beliebtheit von Produkten abgebildet werden kann. Die abnehmende Beliebtheit kann darin begründet sein, dass das Produkt leichte aversive sensorische Komponenten enthält, die erst langsam wahrgenommen werden und in die Beurteilung einfließen, oder es könnte auch in einer zu geringen sensorischen Komplexität begründet liegen, die über die Zeit hin zu sensorischer Langeweile führt.

In Abb. 1 sind die Ergebnisse eines Tomaten-Langeweiletests grafisch dargestellt und es zeigt sich hier, dass über die Anzahl der Präsentationen hin Produkt B stärker in der Beliebtheit absinkt und auch der Verbrauch deutlich stärker absinkt als bei Produkt A.

1.1.4 Aversionstest

Der Aversionstest ist eine spezielle Ausformung des Dauerakzeptanztests, die sich in vier Phasen gliedert. In der ersten Phase

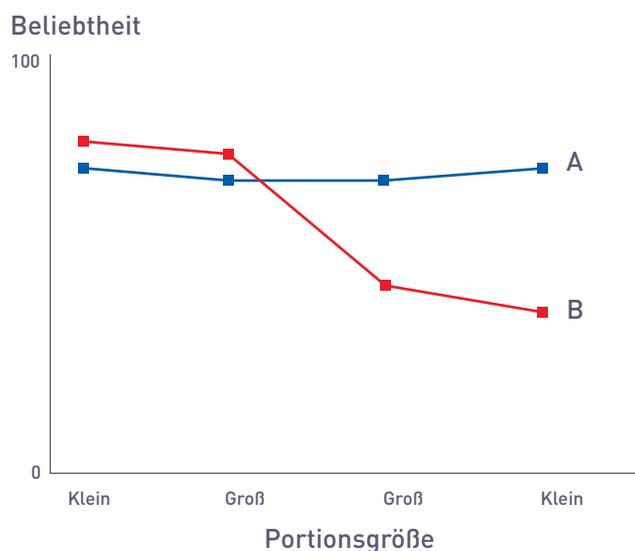
wird ein spontaner Akzeptanztest mit einer kleinen Produktmenge durchgeführt. Dann wird die Testperson „gezwungen“, d. h. gebeten eine große Menge zur Gänze zu verzehren und anschließend hinsichtlich Beliebtheit zu beurteilen; zudem kann die Zeit gemessen werden, die zum Verzehr der großen Menge nötig ist. In der dritten Phase schließlich wird wieder eine große Produktmenge gereicht, aber diesmal wird freigestellt wie viel davon verzehrt wird, bevor anschließend eine Beliebtheitsbeurteilung vorgenommen wird. In der dritten Phase wird neben dem Beliebtheitsurteil auch die Verzehrsmenge festgehalten und die Prüfpersonen werden gebeten abzuschätzen wie viel sie verzehrt haben. Abschließend wird in einer vierten Phase ein spontaner Akzeptanztest mit einer geringen Produktmenge durchgeführt.

Die ermittelten Daten können dann wie in Abb. 2 dargestellt zusammengefasst werden und erlauben ein Urteil darüber, inwiefern bei Konfrontation mit größeren Produktmengen mit Aversionen zu rechnen ist. Ein Absinken der Beliebtheitsurteile, lange Verzehrzeit, geringe Verzehrmenge und großes Verhältnis zwischen geschätzter und tatsächlicher verzehrter Produktmenge sind Indizien für die Entstehung von Aversionen durch die angewandten großen Produktmengen.

1.2 Testbedingungen

Die Testbedingungen bei Beliebtheits- und Präferenztests kann man in vielen Dimensionen variieren [9]. Beliebtheit und Präferenz können einmalig und spontan erfasst werden (one-shot) oder immer wieder über einen längeren Zeitraum hinweg durch häufige Konfrontation mit dem Produkt (extended use). Beliebtheits- oder Präferenzprüfungen werden im Sensoriklabor, in Teststudios (CLT Central Location Test) und auch zu Hause durchgeführt (HUT Home-Use-Test). Die Testsituationen im Labor und im Studio sind gut unter Kontrolle des Versuchsleiters,

Abb. 2: Darstellung der Daten eines Aversionstests



	A	B
Zeit [s]	5,3	8
Menge [g]	125	90
Geschätzt/gegessen	1,15	1,57

allerdings sind sie *abiotisch*, denn sie entsprechen weniger den realen Verzehrssituationen von Konsumenten. Home-Use-Tests sind dagegen sehr realitätsnah (*biotisch*), aber die eigentliche Prüfung und das Beurteilungsprozedere durch den Konsumenten entziehen sich weitgehend der Kontrolle durch die Versuchsleitung. Kostenintensive Mischformen sind Studios, die z. B. in Form eines Restaurants oder eines Wohnzimmers eingerichtet sind. Dabei ist die produkt- und konsumenten-gruppenspezifische Verzehrssituation zu berücksichtigen. Bei manchen Konsumentengruppen wie Kindern oder älteren, evtl. gebrechlichen Menschen ist auf eine spezielle Gestaltung des Prüfbereiches zu achten. Kinder sind bevorzugt in einem kinderspezifischen Ambiente zu testen, d. h. es sind geeignete niedrige Stühle und Tische notwendig, eine kindgerechte Kommunikationsweise und auch ein soziales Umfeld nötig, die den Kindern etwaige Ängste und Hemmungen nehmen (z. B. weibliche Testleiterin, Begleitung durch Elternteil bzw. Vertrauensperson). Kindergartenkinder, Schulkinder, Rentner und ähnliche Personengruppen kann man auch in speziell adaptierten Räumen von Kindergärten, Schulen und Rentnerclubs oder -heimen durchführen. Eine kostengünstigere Variante ist ein Test in einem Teststudio (Central Location Test), zu dem die Testpersonen an einen vorgegebenen Prüfort, dem Teststudio, kommen bzw. gebracht werden müssen. Im Teststudio können die Versuchspersonen auch – beispielsweise durch geeignete Bilder, Musik oder eine Erzählung – in eine gewünschte Situation versetzt werden, sodass es zu einer mentalen Simulation der realen Situation kommt. Manche Sensorikinstitute benutzen auch mobile Sensoriklaboratorien (Autobusse), in denen sich definierte Prüfbereiche und eine kleine Küche zur Zubereitung der Prüfproben befinden. Mit solchen mobilen Sensoriklaboratorien kann man in jeden kleinen Ort am Land, vor jeden Supermarkt fahren, um dort Prüfpersonen für die Tests zu rekrutieren, die mit herkömmlichen Rekrutierungsmethoden nur schlecht erreichbar sind.

Mit den Instruktionen an die Versuchspersonen über die Versuchsdurchführung ist sehr vorsichtig umzugehen, da aus vielen Untersuchungen bekannt ist, wie stark Informationen, vor allem emotional wirksame Informationen die affektiven und hedonischen Urteile beeinflussen können. Allein die Art und Weise, wie der Versuchsleiter den Prüfpersonen gegenübertritt (freundlich oder unfreundlich) kann einen Einfluss auf das Ergebnis ausüben. Dem entsprechend ist auch dem Design der Fragebögen große Aufmerksamkeit zu widmen. Wird nur die Frage gestellt „Wie sagt Ihnen das Produkt zu?“ oder werden auch andere Aspekte gefragt? Wie wird die Frage gestellt? Welche Wörter und Begriffe werden verwendet? Wie viele Fragen werden gestellt? Werden die Fragen Face-to-Face gestellt, schriftlich über einen Fragebogen oder über einen Computerbildschirm? Das alles sind Entscheidungen, die das Versuchsergebnis beeinflussen.

Die Angemessenheit der Rekrutierung von Testpersonen ist entscheidend für die Aussagekraft von hedonisch-affektiven Tests. Zuerst ist zu entscheiden, wer zu den Testpersonen zählen soll. In der Regel ist die Grundgesamtheit, aus der eine Stichprobe gezogen werden soll, die Zielgruppe des untersuchten Produkts. Um verallgemeinerbare Aussagen über die Grundgesamtheit machen zu können, ist es notwendig, dass die Stichprobe ein ähnliches Bewertungsverhalten aufweist wie die Grundgesamtheit. Dies ist am besten gewährleistet, wenn die

Stichprobe repräsentativ und ausreichend groß ist. Repräsentativität einer Stichprobe ist dann gegeben, wenn jede Person der Grundgesamtheit die gleiche Wahrscheinlichkeit hat, in die Stichprobe zu gelangen. Da diese Anforderung nur mit großem Zeit- und Kostenaufwand zu erreichen ist, begnügt man sich häufig mit konventionellen Stichproben, die durch die Art und Weise der Rekrutierung definiert ist. Untersuchungsergebnisse resultierend aus konventionellen Stichproben können allerdings wenig aussagekräftig sein, da die Stichprobe wahrscheinlich nicht repräsentativ ist. Marktforschungsinstitute sind in der Lage, repräsentative Stichproben von Konsumenten in Form von Adressdateien basierend auf großen Verbraucherpanels zu liefern.

Eine wichtige Frage ist auch, wie viele Testpersonen für die Prüfung herangezogen werden. Je größer die Grundgesamtheit und je heterogener die Grundgesamtheit ist, desto größer muss auch die Stichprobe sein. Als orientierende Größen werden Zahlen zwischen 80 bis 400 Personen angegeben.

1.3 Probendarreichung und -bezeichnung

Die Art und Weise, in der die Prüfproben den Testpersonen zur Prüfung gereicht werden, kann das Testergebnis stark beeinflussen und ist daher gut abzuwägen. Bei monadischer Darreichung erhält die Testperson jeweils nur eine Probe zur Prüfung, bei sequentiell monadischen Prüfungen mehrere einzelne Proben hintereinander. Die Testperson vergleicht hier das geprüfte Produkt mit einem durch Erfahrung zuvor bereits erlernten Standard. Monadische Tests sparen Zeit, aber man benötigt viele Prüfpersonen. Zudem erfährt man nichts über die Art und Weise, wie die einzelne Prüfperson die Skale nutzt, sodass es wegen der großen interindividuellen Variabilität möglicherweise zu einer verringerten Aussagekraft der Untersuchung kommt. Weiterhin ist das menschliche Wahrnehmungssystem insgesamt sehr schlecht bei absoluten sensorischen Bewertungen und weitaus besser bei relativen Bewertungen. Bei sequentiell monadischen Tests ist darauf zu achten, dass die erste Probe immer besonders bewertet wird. Das ist auf die Überraschung der Testperson zurückzuführen, die diese Probe auslöst. Die folgenden Proben werden in ihrer Beurteilung auf die erste Probe bzw. auf die zuvor getesteten Proben bezogen, sodass es zu so genannten carry-over-Effekten kommen kann, d. h. die Proben werden nicht mehr unabhängig voneinander bewertet und die Reihenfolge der Prüfung hat einen Einfluss auf das Ergebnis. Durch eine deutliche zeitliche Trennung der Prüfung der dargereichten Proben verringert sich dieser Effekt. Die Beurteilung jeder Proben sollte jeweils sofort nach ihrer Prüfung erfolgen, solange die sensorisch-hedonische Empfindung noch frisch in der Erinnerung der Testperson ist. Die Ordnung der Proben bei sequentiell monadischen Tests folgt üblicherweise einer Randomisation, d. h. einer zufälligen Anordnung der Proben. Allerdings ist es bei einer starken Dominanz eines sensorischen Merkmals sinnvoll, die Probenreihenfolge so zu steuern, dass z. B. die am wenigsten bittere Probe zuerst und die bitterste als letzte Probe bewertet wird, denn nur so ist gewährleistet, dass es zu keinen Adaptationseffekten und zu keinen verzerrten Beurteilungen kommt.

Die Bezeichnung der Prüfproben ist in Abhängigkeit des Untersuchungszieles zu wählen. Wenn der sensorische Kern eines Produktes getestet werden soll, ist es notwendig, seine Markenidentität und Herkunft unkenntlich zu machen. Die Proben

werden blind oder doppelt blind gereicht. Blind bedeutet, dass die Proben mit neutralen Codes (z. B. Ziffern, Zeichen oder Buchstaben) gekennzeichnet werden. Doppeltblind bedeutet, dass auch für den Leiter des Versuchs die Markenidentität nicht erkennbar ist, sodass eine – unwillkürliche – Beeinflussung der Prüfpersonen durch den Prüfleiter auch ausgeschlossen werden kann.

Interessant sind natürlich auch Tests, bei denen die Marke oder andere Produktinformationen gegeben werden. Das Produkt wird gelabeled oder gebranded dargeboten. Bei diesen Versuchen kann zwischen Akzeptanz des sensorischen Kerns und der Akzeptanz der Label- oder Markeninformation nicht unterschieden werden. Darum sind Kombinationen von blind und gebranded durchgeführten Tests sehr interessant, denn so kann man erkennen, welchen Effekt das Branding auf die hedonische Akzeptanz ausübt. In der Mehrzahl der am Markt befindlichen Markenprodukte sind solche Effekte signifikant.

1.4 Skalen in hedonischen und affektiven Tests

Welche Skala für hedonische und affektive Tests die günstigste ist, kann nicht allgemein beantwortet werden. Häufig werden Rangordnungsskalen, Kategorienskalen, Linienskalen und Labelled Magnitude Skalen eingesetzt.

Rangordnungstests können als ausgeweitete gerichtete Unterschiedstests aufgefasst werden. Die Versuchspersonen bringen dabei die Produkte hinsichtlich des Ausmaßes der hedonischen Akzeptanz in eine Rangordnung, ohne die Abstände zwischen den Rängen näher zu quantifizieren. Da diese Methode sehr einfach ist, ist sie gut einsetzbar bei Kindern, Analphabeten oder auch crosskulturellen Untersuchungen. Die Auswertung erfolgt mit Hilfe des Rangsummentests (Basker Tabellen) oder des Friedman Tests. Die Ergebnisse der meisten anderen Skalen (z. B. Linienskalen oder Kategorienskalen) können – unter Informationsverlust – in Rangordnungsergebnisse umgewandelt werden.

1.4.1 Kategorienskalen

Bei Kategorienskalen wird die Intensität der subjektiven Empfindung auf vorgegebenen, diskreten Antwortalternativen abgebildet. Hierfür werden typischerweise 5 bis 15 Kategorien verwendet, die meist numerisch, verbal oder als Bilder dargestellt sind. Gegenüber der Linienskala liegt bei der Kategorienskala also eine weitaus geringere Anzahl an Antwortmöglichkeiten vor. Die Kategorienskala hat das Skalenniveau des Schulnotensystems, das heißt es werden Produkte von 1 bis 5, 1 bis 9 oder 1 bis 11 in ihrer Beliebtheit beurteilt. Den Zahlen werden auch häufig verbale Beschreibungen zugeordnet. 1 bedeutet beispielsweise „mag ich besonders gern“, 5 „mag ich weder/noch“ und 9 „mag ich überhaupt nicht“. Der Vorteil dieser Skalen liegt in ihrem intuitiv leichten Verständnis. Allerdings ist ihrer statistischen Auswertbarkeit aufgrund des Skalenniveaus eine Grenze gesetzt. Es dürfen aus diesen Daten z. B. als Lokationsschätzer

Abb. 3: Beispiel für eine Kategorienskala

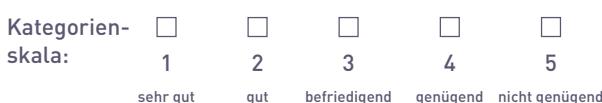
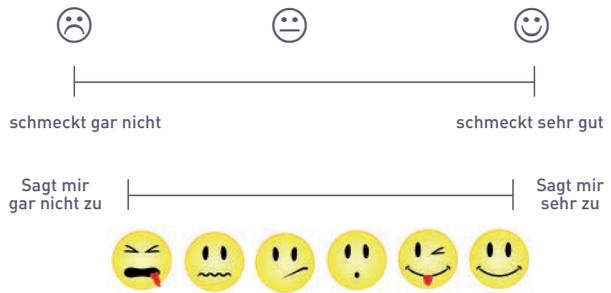


Abb. 4: Beispiele für Linienskalen mit Smilies und Badies



nur Mediane berechnet werden und nicht Mittelwerte. Ein Beispiel für eine Kategorienskala befindet sich in Abb. 3.

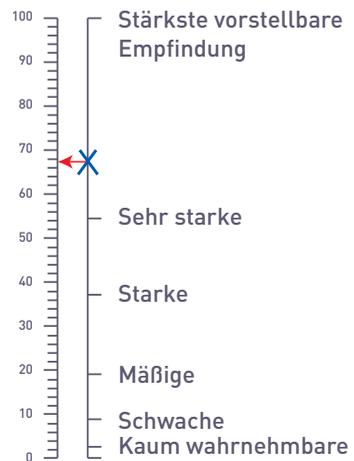
1.4.2 Linienskalen

Linienskalen (vgl. Abbildung 4) bestehen aus einer meist horizontalen Linie, die durch verbal beschriebene Ankerpunkte mehr oder weniger strukturiert werden kann. Auf diese Skalen werden Marken gesetzt, die der wahrgenommenen Beliebtheit des getesteten Produkts entsprechen sollen. Linienskalen haben den Vorteil, dass sie ein höheres Skalenniveau haben als Kategorienskalen, das heißt man darf statistisch betrachtet Mittelwerte und Standardabweichungen berechnen, aber sie sind vor allem für bestimmte Konsumentengruppen schlecht verständlich. Um die Verständlichkeit dieser Skalen zu erhöhen, versieht man häufig Ankerpunkte mit Piktogrammen wie Smilies & Badies oder Comicfiguren, die den Akzeptanzgrad nicht-sprachlich deutlich machen sollen. Bei Verwendung von Badies ist darauf zu achten, dass keine traurigen Gesichter verwendet werden, sondern Gesichter, die tatsächlich Ablehnung bis zu Ekel ausdrücken.

1.4.3 Labelled Magnitude Skala (LMS)

Die LMS (vgl. Abbildung 5) ist eine Linienskala mit quasi logarithmisch angeordneten Ankerpunkten aus Magnitude Estimation-Experimenten. Sie ist noch wenig verbreitet, scheint aber zuverlässigere Bewertungen zu ermöglichen als normale Linienskalen.

Abb. 5: Labelled Magnitude Scale



1.5 Auswertung von Beliebtheitsdaten

Am Anfang der Auswertung von Beliebtheitsdaten steht die Darstellung der Verteilung der Urdaten, denn bei diesen Daten liegt aufgrund der Unterschiedlichkeit der Konsumenten häufig keine Normalverteilung vor. Die Darstellung kann in Form von Häufigkeitshistogrammen oder Density

Traces vorgenommen werden. Mit Hilfe einer Betrachtung der Verteilung der Urdaten kann auch erkannt werden, ob es innerhalb der untersuchten Stichprobe Gruppen gibt – beispielsweise Ablehner und Befürworter des Produkts bzw. neutrale Personen.

Liegen Normalverteilungen vor, kann man zur Charakterisierung der Verteilung den Mittelwert zur Beschreibung ihrer Lage und die Standardabweichung zur Beschreibung ihrer Variabilität heranziehen. Statistische Kennzahlen ermöglichen die Beschreibung der Schiefe und der Wölbung der Verteilung. Es bieten sich t-Tests zur Entscheidungsfindung an, ob signifikante Unterschiede zwischen den Verteilungen der Beliebtheitsdaten zweier Produkte vorliegen oder nicht. Varianzanalysen und Multiple Range Tests zeigen, ob es signifikante Unterschiede zwischen den Beliebtheitsdaten mehrerer Produkte gibt, bzw. welche Produkte sich voneinander signifikant unterscheiden.

Liegen keine Normalverteilungen vor, ist zu klären, ob es sich um zwei- oder mehrgipfelige Verteilungen handelt. Eine zwei- oder mehrgipfelige Verteilung kann auf zwei oder mehrere Konsumentengruppen hinweisen, die man getrennt als normalverteilt betrachten könnte. Ist das nicht der Fall, dann bieten sich statistische Tests an, mit deren Hilfe auf Unterschiedlichkeit von Verteilungen geprüft wird, wie z. B. der Kolmogorov-Smirnov Test.

Als multivariate statistische Methoden der Darstellung von Beliebtheitsdaten werden die so genannten Preference Mappings häufig genutzt. Bei ihnen werden Hauptkomponentenanalysen auf die Beliebtheitsdaten der Konsumenten angewendet, so dass man eine zweidimensionale Projektion der Urdaten erhält, in der sowohl die Produkte als auch die Urteile der einzelnen Personen sichtbar sind. Aus diesen Darstellungen ist gut erkennbar, welche Produkte von welcher Personengruppe hoch bzw. niedrig bewertet werden. Man kann gut erkennen, ob Gruppen innerhalb der untersuchten Versuchspersonen vorliegen und welche Produkte in ihrer Beliebtheit relativ ähnlich sind.

2. Observationale Methoden

Da die moderne Psychologie die Möglichkeit zur Introspektion und zum Bericht über eigene mentale Prozesse insgesamt als sehr eingeschränkt beurteilt [10], versucht man, in den relativ jungen observationalen Testmethoden die Bewusstmachung und Abfragung der affektiv-hedonischen Empfindungen zu umgehen, indem man Parameter beobachtet und misst, von denen man sicherer auf die affektiv-hedonischen Empfindungen schließen kann. Man vermeidet mit solchen Methoden also die Verzerrungen durch den gesamten Komplex der Befragung. Im Folgenden wird ein kurzer Abriss über die Methoden gegeben, die bisher entwickelt wurden. In der Untersuchungspraxis von Unternehmen sind die observationalen Methoden noch nicht so stark etabliert wie die introspektiven Methoden.

2.1 Bewertung durch Beobachter

Diese Methode wurde früh bereits bei Produkten für Babies eingesetzt, denn Babies können sprachlich noch nicht Auskunft über ihre Empfindungen geben. Die Mütter der

Babies sind dagegen sehr gut in der Lage, die emotionale Situation ihres eigenen Babies zu beurteilen und wie sehr diesem bestimmte Lebensmittel zusagen. In den letzten Jahren versucht man mit wechselndem Erfolg, die Vermessung und Interpretation der mimischen und gustofazialen Reaktionen mit videobasierten Auswertungen durchzuführen. Diese Auswertungen sollen schlussendlich Rückschlüsse auf die hedonische Akzeptanz der getesteten Lebensmittel zulassen.

2.2 Autonome Maße

Unwillkürliche, der bewussten Steuerung nicht zugängliche Reaktionen des autonomen Nervensystems können ebenfalls als Maße für emotionale, nicht bewusste Regungen herangezogen werden. Zu diesen Reaktionen gehören Hautwiderstand, Atemfrequenz, Blutdruck, Herzfrequenz, Pupillenweite, Kaubewegungen, Speichelfluss etc. Die Messung dieser Parameter ist häufig ein kleineres Problem als ihre Interpretation in Hinsicht auf Art und Intensität einer dahinterliegenden Empfindung bzw. Emotion.

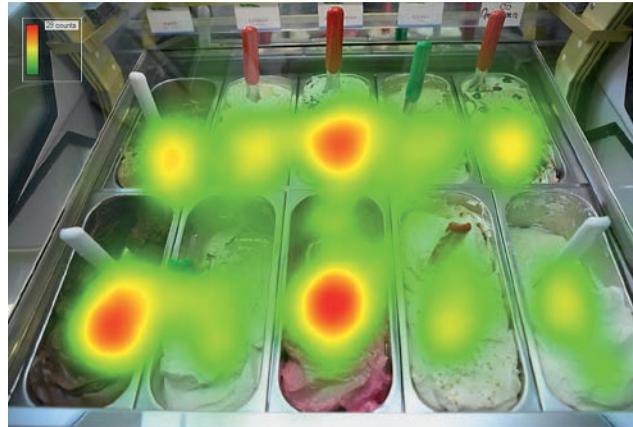
2.3 Gehirnbasierte Maße

Der direkte Blick in die Aktivitäten des Gehirns wird durch bildgebende Verfahren der Gehirnforschung wie Kernspintomografie oder Elektroenzephalographie zunehmend präziser möglich und erlaubt auch die Bestimmung von mentalen Zuständen, derer sich die Versuchsperson nicht bewusst sind. So zeigt die Aktivierung bestimmter Gehirnregionen Lust und Wohlbefinden an. Andere Bereiche des Gehirns hingegen zeigen, dass Empfindungen wie Appetit oder Begehren vorhanden sind und wieder andere aktivierte Bereiche weisen auf Ekel- oder Unlustgefühle hin. Das Ausmaß dieser Empfindungen kann über das Ausmaß der Aktivität der entsprechenden Regionen abgeschätzt werden.

2.4 Verhaltensbeobachtung

Die Beobachtung bestimmter Verhaltensaspekte lässt ebenfalls Rückschlüsse auf die hedonische und affektive Bewertung von Produkten zu. Typische beobachtete Parameter im Zusammenhang mit Lebensmitteln sind Auswahlverhalten, Verzehrsmenge, Verzehrsdauer, Verzehrsfrequenz oder auch die Reaktionszeit bei Beurteilungen. Das Auswahlverhalten ist ein Indiz dafür, welche der vorliegenden Produktalternativen in der untersuchten Situation bevorzugt werden. Eine große Verzehrsmenge zeigt an, dass ein Lebensmittel beliebt ist und eine rasche Verzehrsgeschwindigkeit sowie hohe Verzehrsfrequenz werden als Indikatoren für eine positive Akzeptanz des untersuchten Produkts interpretiert. Rasche Reaktionszeiten zeigen an, dass ein eindeutiges, stark ausgeprägtes Urteil vorliegt und negative Bewertungen werden in der Regel rascher vorgenommen als positive. Die Verhaltensbeobachtung kann durch das Untersuchungspersonal direkt erfolgen oder über Videokameras, mit denen Verzehrs- oder auch Kaufsituationen aufgezeichnet werden. Die Auswertung dieser Daten in größerem Ausmaß ist kostenintensiv, da Auswertalgorithmen wie man sie in der Tierforschung verwendet hat, offensichtlich nicht ausreichen, das komplexe Verhalten von Menschen zu analysieren. Daher sind nach wie vor Personen für die Auswertung einzustellen.

Abb. 6: Hot Spot Darstellung des Blickverhaltens bei einer Eisdiele (rot: lange Verweildauer, grün: kurze Verweildauer)



Ein spezieller Aspekt der Verhaltensbeobachtung ist – unter dem populären Motto „Augen sind die Fenster der Seele“ – das Beobachten des Blicks von Konsumenten. Das Blickverhalten, das man mit Hilfe von Eyetracking Techniken aufzeichnen kann, zeigt welche Aspekte eines Produktes, eines Regales oder einer Verpackung die höchste visuelle Aufmerksamkeit auf sich ziehen, auf welche Teile als erstes geschaut wird oder auf welche Teile immer wieder. In Abb. 6 sind die Hot Spots der visuellen Aufmerksamkeit bei einer Eisdiele dargestellt. Man kann erkennen, dass die Eissorten Schokolade, Himbeere und Stracciatella die intensivste visuelle Aufmerksamkeit erfahren und dann diejenigen Produkte folgen, die gewisse visuelle Kontraste in sich tragen wie Kirschen mit roter Sauce oder Kaffeeeis mit Kaffeebohnen. Die Eisprodukte ohne Färbung oder Kontrastierung werden am wenigsten beachtet. Sicher spielt hier auch die Anordnung der Eissorten eine Rolle, Produkte in der Mitte oder am Rand ziehen leichter visuelle Aufmerksamkeit auf sich als Produkte, die dazwischen angeordnet sind.

Mit manchen Eyetracking Anlagen kann man gleichzeitig auch die Pupillenweite messen und das mimische Verhalten aufzeichnen, sodass man weitere Indizien dafür gewinnt, welche emotionalen Beweggründe für das Blickverhalten vorliegen. Auch Speichelfluss, Hautleitfähigkeit und andere unwillkürliche Reaktionen können gleichzeitig mit der Aufzeichnung des Blickverhaltens gemessen werden, sodass ein vollständigeres Bild entsteht. Zwischen der Attraktivität eines visuellen Objektes und der Blickdauer besteht häufig eine Art U-förmiger Zusammenhang. Das heißt Konsumenten richten ihre visuelle Aufmerksamkeit lange auf sehr schöne, attraktive Bildelemente und auch auf solche, die sehr hässlich, sehr ungewöhnlich, ganz neu, unbekannt oder sonst wie irritierend sind. Diejenigen Bildelemente, die zwischen diesen beiden Polen liegen, werden eher kurz betrachtet.

Autor: Prof. Dipl.-Ing. Dr. nat. techn. Klaus Dürrschmid, BOKU Wien, Österreich

Detaillierte Informationen zum Thema:

1. Köster, E. P., *Diversity in the determinants of food choice: A psychological perspective*. Food Quality and Preference, 2009. 20(2): p. 70–82.
2. Frewer, L. and H. v. Trijp, *Understanding consumers of food products*. 2007, Cambridge: Woodhead Publishing Ltd.
3. Frandsen, L.W., et al., *Subtle differences in milk: comparison of an analytical and an affective test*. Food Quality and Preference, 2003. 14(5-6): p. 515–526.
4. Frandsen, L.W., et al., *Feelings as a basis for discrimination: Comparison of a modified authenticity test with the same-different test for slightly different types of milk*. Food Quality and Preference, 2007. 18(1): p. 97–105.
5. MacFie, H., *Consumer-led food product development*. 2007, Cambridge: Woodhead Publishing Ltd.
6. Moskowitz, H.R., *Product optimization: approaches and applications*, in *Measurement of Food Preferences*, H. J. H. MacFie and D. M. H. Thomson, Editors. 1994, Blackie Academic & Professionals: London. p. 40.
7. Wilson, T. D. and J. W. Schooler, *Thinking Too Much: Introspection Can Reduce the Quality of Preferences and Decisions*. Journal of Personality and Social Psychology, 1991. 60(2): p. 181-192.
8. Lawless, H. T. and H. Heymann, *Sensory Evaluation of Food. Principles and Practices*. 1999, Gaithersburg: Chapman&Hall.
9. Lill, F. and E. Köhn, *Methoden, Anwendungen und Analysen der Sensorik in der Marktforschung*, in *Praxishandbuch Sensorik*, M. Busch-Stockfisch, Editor. 2006, B. Behr's Verlag: Hamburg.
10. Nisbett, R. E. and T. D. Wilson, *Telling more than we can know: Verbal reports on mental processes*. Psychological Review, 1977. 84(3): p. 231-259.