



Hochschule Neubrandenburg
Fachbereich Agrarwirtschaft und Lebensmittelwissenschaften
Studiengang Lebensmitteltechnologie
WS 2017 – SS 2018

Sensorische Charakterisierung von ver- schiedenen Fruchtsäften mit einem ge- schulten und ungeschulten Panel

Bachelorarbeit

Verfasser: Neumann, Alexander

Betreuer: Prof. Dr. Jörg Meier

Neubrandenburg, 23.01.18

urn:nbn:de:gbv:519-thesis2017-0598-2

Abstract

Sensory investigation of different fruit-juices with an untrained and trained panel

Sensory is an important instrument in the food industry or, for example in the automobile industry. The working with different sensory methods make it possible to get, for example a product acceptance or to know if the “new car-odor” is intensive enough. No matter which method is used, humans are the measuring instruments. For product-acceptance it is usual to use an untrained panel of consumers and for more specific information a trained panel is indispensable. The panel performance is very important to get useful data. The perfect measuring instrument needs a high product discrimination-rate and a good repeatability on different testing-days, also when possible a high reproducibility and consistency with the overall panel. The problem is, to get this achievements, a time intensive and expensive training is desired. To avoid this training, it is possible to use an untrained panel. A consumer test with “catch-all-that-apply”-questions, hedonic-scales and “just-about-right”-intensity-questions with an untrained panel gives the usefull data. Also, the investigation of profiles with two trained panels helps to get the informations. The evaluation of the profiling uses the DIN EN ISO 11132. The DIN EN ISO 11132 is an international norm which includes the most significant properties for reaching the most excellent or most efficient or most accurate panel performance. The different panels tested six different branded orange-juices at the University of Applied Science in Neubrandenburg, Germany. With the softwares “Fizz” and “Xlstat” was it possible to check the data and to compare the panels.

The results show that the trained and untrained panel gives similar assessments. So the consumer panel can be used, if no trained panel is available. The panel performance of the trained panel is very high, but the panels has the problem with the attribute “sweet in odor”. Also both panels include sampler with good performance and bad performance, which need more training in different attributes. So the DIN EN ISO 11132 is accomplishes an abridgment in the sensory evaluation of performance. Therefore, the norm is an audited, consistent and acknowledged method to check the performance.

Inhaltsverzeichnis

| | |
|---|----|
| Verwendete Abkürzungen..... | 5 |
| 1 Einleitung..... | 6 |
| 2. Stand der Technik und Wissenschaft | 7 |
| 2.1 Panelleistung in der Sensorik | 7 |
| 2.2 Konventionelle Prüfmethoden für die Panelleistung | 8 |
| 2.3 Prüfung der Panelleistung anhand der DIN EN ISO 11132..... | 10 |
| 2.3.1 Definition Leistungsfähigkeit | 10 |
| 2.3.2 Ziel der ISO 11132..... | 11 |
| 2.4 Rechtliche Grundlagen und Begriffsbestimmung der verwendeten Fruchtsäfte | 11 |
| 2.4.1 Fruchtsaft aus Direktsaft | 12 |
| 2.4.2 Fruchtsaft aus Fruchtsaftkonzentrat | 12 |
| 2.4.3 Nektar..... | 12 |
| 2.4.4 Fruchtsaftgetränke und Fruchtsaftschorlen | 13 |
| 3 Material und Methoden | 13 |
| 3.1 Profilprüfung..... | 13 |
| 3.2 Hedonische Prüfverfahren..... | 19 |
| 3.2.1 Verbrauchertest und 9-Punkt hedonische Skala..... | 19 |
| 3.2.2 Check-All-That-Apply | 20 |
| 3.2.3 Just-About-Right..... | 22 |
| 3.2.3.1 Grundlagen und Aufbau | 22 |
| 3.2.3.2 Penalty Analysis..... | 23 |
| 3.3 Prüfbogen Verbrauchertest | 27 |
| 3.4 Preference Mapping | 28 |
| 3.4.1 Internal Preference Mapping..... | 28 |
| 3.4.2 External Preference Mapping..... | 30 |
| 3.5 Untersuchungsmaterial..... | 33 |
| 3.6 Vorbereitung | 34 |
| 3.6.1 Probenvorbereitung | 34 |
| 3.7 Beschreibung der Probanden | 35 |
| 3.7.1 Geschultes Panel | 35 |
| 3.7.2 Konsumentenpanel..... | 36 |
| 4. Ergebnisse | 37 |
| 4.1 Verbrauchertest | 37 |

| | | |
|-----------------------|--------------------------------------|-----|
| 4.1.1 | Angabe zur Person | 37 |
| 4.1.2 | Beliebtheitsprüfung | 38 |
| 4.1.3 | JAR-Fragen | 40 |
| 4.1.4 | Penalty-Analyse | 43 |
| 4.1.4.1 | Punica | 43 |
| 4.1.4.2 | Tip Nektar | 45 |
| 4.1.4.3 | Amecke plus Zink und Vitamin C | 47 |
| 4.1.4.4 | Hohes C 100% Orangensaft | 50 |
| 4.1.4.5 | Amecke Mandarine und Orange | 53 |
| 4.1.4.6 | Hohes C O-T-L | 55 |
| 4.1.5 | CATA | 58 |
| 4.2 | Profilprüfung | 63 |
| 4.2.1 | Gruppe 1 | 63 |
| 4.2.2 | Gruppe 2 | 70 |
| 4.3 | Preference Mapping | 77 |
| 4.3.1 | Internal Preference Mapping | 77 |
| 4.3.2 | External Preference Mapping | 79 |
| 5. | Diskussion | 84 |
| 6. | Zusammenfassung | 92 |
| Anhang 1.1 | | 94 |
| Anhang 1.2 | | 95 |
| Anhang 1.3 | | 96 |
| Anhang 1.4 | | 97 |
| Anhang 1.5 | | 98 |
| Literaturverzeichnis | | 99 |
| Abbildungsverzeichnis | | 103 |
| Tabellenverzeichnis | | 105 |

Verwendete Abkürzungen

Tabelle 1: Verwendete Abkürzungen

| Abkürzung | Bedeutung |
|-----------|---|
| CATA | Check-All-That-Apply |
| HKA | Hauptkomponentenanalyse |
| JAR | Just-About-Right |
| L.S.D | Least significant difference |
| MDPREF | Multidimensional Preference Analysis (Internal Preference Mapping) |
| MhD | Mindesthaltbarkeitsdatum |
| PCA | Principial Component Analysis |
| PREFMAP | External Preference Mapping |

1 Einleitung

Die Sensorik ist ein sehr wichtiges Instrument in der Industrie geworden. Es wird verwendet um z.B. eine Produktakzeptanz oder sensorische Unterschiede zwischen einer alten und neuen Rezeptur herauszufinden. Von der Sensorik profitiert nicht nur die Lebensmittelindustrie, sondern auch z.B. die Automobilindustrie, wenn es darum geht den passenden Geruch eines Neuwagens zu erhalten oder den Geräuschpegel des Schließens der Tür zu reduzieren. Hierbei gilt der Mensch als das Messinstrument, der zuverlässige und brauchbare Daten liefert. Je nachdem welche Daten gewünscht sind, werden zeitaufwändige, sowie kostspielige Trainings und Schulungen unumgänglich. Das perfekte Messinstrument muss eine hohe Unterscheidungsrate der Produkte aufweisen, sowie eine gute Wiederholpräzision bei mehreren Prüfungssitzungen, aber auch eine möglichst hohe Vergleichspräzision und Konstanz mit dem gesamten Panel.

In dieser Arbeit soll mit Hilfe von verschiedenen sensorischen Verfahren, herausgefunden werden, ob ein ungeschultes Panel aus Konsumenten genauso zuverlässige Daten liefert, wie ein geschultes Panel aus der Profilierung. Des Weiteren soll die Handhabung der DIN EN ISO 11132 beschrieben werden. Dadurch können Kosten, Zeit und die Mitarbeiterbelastung, das von einem Paneltraining ausgeht, eingespart werden. Dafür wird mit Hilfe eines Verbrauchertestes der aus JAR-Intensitätsfragen („Just-About-Right“), CATA („Check-All-That-Apply“) und Beliebtheitsfragen besteht, das ungeschulte Panel geprüft. Parallel wird mit zwei Panels eine Produktschulung, ein Training und die Profilierung in drei Sitzungen durchgeführt, wo es gilt die Intensität von verschiedenen Produktmerkmalseigenschaften zu bestimmen. Für die Auswertung der Profilierung wird die DIN EN ISO 11132 herangezogen, wo es darum geht in wenigen Schritten eine Panel-und Prüferleistung zu erhalten. Das Problem der Prüfung der Panelleistung besteht darin, dass unzählige Methoden zu diesem Thema existieren. Die DIN EN ISO 11132 ist eine internationale Norm, die die wichtigsten Eigenschaften zur Entschlüsselung der Leistung zusammenfasst, beschreibt und somit eine Übersicht in der Bewertung der sensorischen Panelleistung schafft. Dadurch ergibt sich eine geprüfte, einheitliche und anerkannte Methode.

Die Auswertung erfolgt mit Hilfe der Software „Fizz“ (ANOVA, PCA) und „Xlstat“ (JAR, CATA, preference mapping). Als zu testendes Produkt werden diverse Orangensaft verwendet. Orangensaft gehört zu den beliebtesten Fruchtsäften Deutschlands. Im Jahre 2016 wurde ein pro-Kopf-Verbrauch von 7,5 Litern verzeichnet. Selbst der beliebteste Fruchtsaft der Deutschen, der Apfelsaft, verzeichnete im Jahr 2016 den gleichen Verbrauch von 7,5 Litern (Statista, 2017, beruht auf den Verband der Deutschen Fruchtsaftindustrie).

2. Stand der Technik und Wissenschaft

2.1 Panelleistung in der Sensorik

Die Sensorik ist ein wichtiges Instrument, um relevante Produktinformationen für Lebensmittelfirmen zu liefern. Mit diesen Daten können Abteilungen, wie die Produktentwicklung oder das Marketing, neue Anreize, Wege und Möglichkeiten finden, das Produkt zu vermarkten oder zu verbessern. Durch die Sensorik kann z.B. die Produktqualität verbessert werden, aber auch die Produkteigenschaften kann bestimmt werden. Ebenfalls ist es möglich das Produkt mit anderen konkurrierenden Waren zu vergleichen (Latreille, 2006). Um diesen Aufgaben gewachsen zu sein, wird ein Messinstrument benötigt. In der Sensorik gilt der Mensch mit seinen fünf Sinnen als das Instrument schlechthin. Damit jedoch brauchbare und zuverlässige Daten entstehen, wird von dem Panel eine gewisse Professionalität, Leistung und vor allem Zuverlässigkeit benötigt. Die Professionalität kann durch ein auf das Produkt abgestimmtes Training gewährleistet werden. Die Trainingsintensität hängt stark von der Sensorikprüfung ab. So brauchen die meisten deskriptiven Prüfungsverfahren ein intensiveres Training, als ein Unterscheidungstest (Lawless und Heymann, 2010). Für die meisten Menschen ist es leicht, z.B. eine bestimmte Salzkonzentration im destilliertem Wasser herauszuschmecken, jedoch wird es deutlich schwieriger, wenn genau die gleiche Aufgabe an einer Salsa angewendet wird, die außer der Geschmackskomponente „salzig“ auch noch „süß“, „sauer“, „bitter“ und „Schärfe“, aber auch Geruchs- und Geschmacksaromen von Tomaten, diversen Gemüsen und Gewürzen verfügt. Aus diesem Grund muss ein Training erfolgen, das all diese Komponenten in den Hintergrund schiebt und die Salzkomponente in den Vordergrund rückt (Meullenet, 2007). Das Training muss mehrere Sitzungen beinhalten und sollte nach längerer Pause wiederaufgefrischt werden, damit die Leistung nicht abnimmt. Bei einem Verbrauchertest findet keine Schulung statt, denn da ist es nicht von besonderer Wichtigkeit, z.B. eine exakte Intensitätsangabe auf einer 0-100 Punkte-Skala abzugeben, sondern eher eine Beliebtheit oder eine abgestumpfte Variante der Intensitätsmessung mit der JAR-5-Punkte-Methode vorzuweisen. Die Leistung eines gut geschulten Trainings wird anhand von Signifikanzen wiedergespiegelt. Also müssen die Prüfer Produktunterschiede signifikant erkennen, wenn welche vorhanden sind. Außerdem darf keine Wechselwirkung zwischen Prüfer und Produkt vorhanden sein, denn das würde bedeuten, dass einige Prüfer anders bewerten als der Rest des Panels bzw. die Bewertung der Produkte von bestimmten Personen abhängig ist. (Lawless und Heymann, 2010). Es ist also wichtig, dass jeder Proband gleich bzw. ähnlich bewertet wie das Panel, denn sonst würde dies die Gesamtleistung des Panels stark beeinträchtigen (McEwan et al., 2002). Somit ist die Konstante Bewertung bzw. Reproduzierbarkeit des einzel-

nen Probanden und des gesamten Panels, eine Erkennung von Unterschieden und genaue Beschreibung der Unterschiede (z.B. Anhand einer Intensitätsskala) und keine Abweichung eines Einzelnen zu dem gesamten Panel, unumgänglich für zuverlässige Daten und somit einer guten Panelleistung. Ist nur eine Eigenschaft der beschriebenen Faktoren nicht vorhanden, so muss das Panel oder die betreffenden Prüfer erneut geschult werden (Lawless und Heymann, 2010). Meullenent et al. (2007) beschreibt die Panelleistung vereinfacht mit einem Orchester. Die einzelnen Musikinstrumente stellen die Attribute dar, die unterschiedlichen Klanghöhen und- tiefen aufweisen, genauso wie die unterschiedlichen Komponenten eines Produktes. Das Publikum wird dem Panel gleichgestellt. Wird die zarte Flöte als Solo präsentiert, so ist sie leicht zu erkennen. Wird jedoch das gesamte Orchester gefordert, so ist es für den Laien schwierig, die leise Flöte, bzw. die einzelnen Instrumente herauszuhören. Das gleiche Problem tritt ein, wenn ein Produkt zum ersten Mal verkostet wird. Die verschiedenen Komponenten überladen unsere Sinne und es wird ein Geräusch gehört, bzw. etwas geschmeckt, was unmöglich zu differenzieren ist. Somit ist das Verstehen der Sprache der deskriptiven sensorischen Analyse nichts anderes, als das Lernen, wie man richtig Musik hört. Auch hier braucht es Schulung, die einzelnen Instrumente erst kennenzulernen und anschließend in einem Orchester wiederzuerkennen (Meullenent et al., 2007).

2.2 Konventionelle Prüfmethoden für die Panelleistung

Für die Bestimmung der Panelleistung existieren zahlreiche Methoden, die jeweils mit unterschiedlichen Aufwand, sowie diversen statistischen Methoden verbunden sind. Eine gängige und leichte Methode ist die Bewertung eines Panels aus der Profilierung anhand der kostenlosen Software „PanelCheck“. Diese Software wurde zum ersten Mal von Asgeir Nilsen und seinen Kollegen im Jahr 2004 auf dem „Siebten Sensometrischen Treffen“ in Davis/Californien, vorgestellt. Die Auswertung erfolgt graphisch anhand eines sogenannten „Egg-Shell“-Plots. Den Namen erhält die Graphik durch ihr Aussehen, welches an den Schalenrest eines aufgeschlagenen Eies erinnert. Anhand dieser Graphik werden die Rangierungen des gesamten Panels als Konsens und die Einzelbewertungen der Prüfpersonen dargestellt. Somit kann ein Vergleich dargestellt und ermittelt werden, welche Probanden nahe am Konsens liegen und welche stark abweichen. Liegen mehrere Prüfer oder sogar das gesamte Panel stark abweichend vom Konsens, so muss eine neue Schulung erfolgen (Lawless und Heymann, 2010; Meier, 2006; Meullenent et al., 2007).

Eine weitere Methode bestritt McEwan et al. (2002) mit einem Leistungstest (orig.: *Proficiency Testing*), wo 12 Panels einer deskriptiven Profilprüfung von 6 Rotweinen unterzogen wurden. Im

gleichen Jahr wurden ebenfalls von McEwan et al. die Süße von 5 verschiedene Apfelsäften, die ein unterschiedliches Verhältnis von Glukose und Fruktose aufwiesen, von 14 Panels geprüft. Der Leistungstest bzw. die Auswertung des Tests bestand aus der GPA (Generalised Procrustes Analyses), einer ANOVA mit einem Signifikanzlevel von 5%, dem Turkey's HSD-Test ($P < 0,05$), mit dem ein signifikanter Unterschied zwischen zwei Proben festgestellt werden konnte. Außerdem enthielt der Test eine Kalkulation des RV-Koeffizienten, um herauszufinden, wie gut die verschiedenen Panels übereinstimmen. Dadurch konnte die Konstanz und Vergleichbarkeit der Ergebnisse von verschiedenen Versuchen verglichen und kontrolliert werden (McEwan et al., 2002 und Meullenet et al., 2007).

Eine weitere Methode betrifft die Bewertung der Panel-Zuverlässigkeit mit Hilfe der „G-Theorie“ (orig.: *generalizability theory*). Diese Methode wendete Paul Tsalma, aber auch andere Sensoriker vor ihm, an. Die Zuverlässigkeit ist hierbei als Produktvarianz über der Gesamtvarianz definiert. Das Arbeiten mit der G-Theorie erstreckt sich durch das Arbeiten mit diversen Algorithmen, wodurch die Panelleistung und die Leistung der einzelnen Prüfer bestimmt werden können. Es ergeben sich zwei Charakteristiken, mit dem die Panelleistung bewertet werden kann: Koeffizient G und Koeffizient Φ . Mit dem Koeffizienten G wird die Zuverlässigkeit beschrieben. Je höher der Koeffizient ist, desto besser ist die Zuverlässigkeit. Der Koeffizient Φ nimmt alle möglichen Quellen der fehlenden Zuverlässigkeit in die Darstellung. Somit konzentriert sich G auf das Anordnen von Produktpunkten und Φ auf das Anordnen und auf die absoluten Punkte. Durch diese Auswertungsmethode kann die Leistung eines Panels berechnet werden. Zudem können zu jedem Zeitpunkt ein oder mehrere Probanden aus der Auswertung herausgenommen werden und die Koeffizienten G und Φ erneut berechnet werden, um zu schauen, ob das Auslassen des Probandes zu einem besseren Ergebnis geführt hätte. Die Auswertung erfolgt mit Hilfe von MS Office (Talsma, 2015).

Latreille et al. beschäftigten sich 2005 mit einer weiteren Methode, in der die Zuverlässigkeit des Panels und der Prüfer über eine Zeit von mehreren Sitzungen geprüft wird. Des Weiteren wurde die Gesamtleistung und Einzelleistung in Form von Unterscheidung der Produkte, Wiederholbarkeit, Prüferauswirkung und der Wechselwirkung zwischen Prüfer und Produkt, verglichen. Die Auswertung erfolgte mit einem gemischten linearen Modell (orig. *MIXED procedur*) und der Software „SAS“. Diese Methode erlaubt das Testen der Produktauswirkung mit der Darstellung der Abweichung zwischen den Prüfern und in den Prüfern, insbesondere die Korrelation zwischen den Messungen die durch die selben Prüfer in verschiedenen Sitzungen durchgeführt wurden (Latreille et al., 2005).

Eine weitere Methode ist die TDS-Methode (Temporal Dominance of Sensations). Lepage et al. veröffentlichten ein Protokoll mit 7 Indikatoren zur Bewältigung der Frage der Prüferleistung. Die Indikatoren sind aufgeteilt in 3 Indikatoren, die das Panel- und Prüferverhalten in Bestimmung der Attributewahl misst und 4 Indikatoren, die die Unterscheidungsfähigkeit des Panels und der einzelnen Prüfer präsentiert. Der Vorteil dieser Methode ist, dass die Interpretation anhand einer einzigen Tabelle erfolgt. Das Hauptziel der TDS-Methode bei der Prüfung der Panelleistung ist herauszubekommen, ob das Panel bzw. die Prüfer in der Lage sind die Unterschiede in den Produkten zu erkennen (Lepage et al. 2014).

Es kann eindeutig festgestellt werden, dass außer dieser kurz vorgestellten Methoden noch weitere Methoden zur Messung der Panelleistung existieren. Aus diesem Grund wurde die DIN EN ISO 11132 entwickelt, um eine konkrete, simple und weltweit anerkannte Methode, die wichtige Kriterien zur Bewertung der Panelleistung misst, darzustellen.

2.3 Prüfung der Panelleistung anhand der DIN EN ISO 11132

Die DIN EN ISO 11132 (2012) oder auch die sensorische Analyse – Methodologie – Leitlinien zur Überprüfung der Leistungsfähigkeit eines quantitativen sensorischen Panels, ist ein Dokument, welches eine Methode zur Feststellung der Panelleistung eines geschulten und deskriptiven Panels, beschreibt.

2.3.1 Definition Leistungsfähigkeit

Die Leistungsfähigkeit kann als Maß bezeichnet werden, welches die Fähigkeit eines gesamten Panels oder einer einzigen Prüfperson darstellt. Die Fähigkeiten beruhen dabei auf zuverlässigen Bewertungen diverser Produkteigenschaften des zu bewertenden Produktes. Somit muss der Proband in der Lage sein, das Produkt, z.B. bei mehrfachen Sitzungen, gleich zu beurteilen und zu erkennen. Dadurch ist eine Wiederholbarkeit der Ergebnisse gewährleistet. Weitere Punkte der Leistungsfähigkeit sind die Erkennung, Bestimmung und Messung der Merkmalseigenschaften der Produkte. Dazu gehört die ähnliche Anwendung und Nutzung der Merkmalseigenschaften, wie vergleichbare Panels. Des Weiteren gehört zu der Leistungsfähigkeit, die Unterscheidung zwischen Reizen, die richtige Anwendung der Skalen und die Reproduzierbarkeit der Ergebnisse von anderen Panels (ISO11132, 2017).

2.3.2 Ziel der ISO 11132

Mit dieser internationalen Norm wird die Überprüfung und Bewertung der Konsistenz, sowie die Wiederholpräzision, festgestellt. Die Wiederholpräzision beantwortet die Frage, ob die Probanden in der Lage sind, in verschiedenen Sitzungen die gleichen bzw. ähnlichen Ergebnisse zu liefern. Zudem ist ein Ausschließen des systematischen Fehlers, d.h. eine immer gleich falsche Bewertung durch z.B. eine falsche Skalennutzung oder Missverständnis der Attribute und die Unterscheidungsfähigkeit des gesamten Panels oder einer einzelnen Prüfperson, überprüfbar. Bei der Vergleichspräzision sollte der Prüfer gleiche bzw. ähnliche Werte liefern wie das gesamte Panel. Damit wird gewährleistet dass ein einzelner Prüfer die Gesamtleistung nicht herunterzieht. An der Unterscheidungsfähigkeit kann gesehen werden, wie gut die Probanden die einzelnen Proben, durch die Attribute, anhand der Intensitätsmessung, unterscheiden können. Das Ziel ist somit die Leistungsfähigkeit zu bestimmen, Probleme und Nachschulbedarf zu erkennen und gegebenenfalls durch gezielte Maßnahmen zu verbessern. Ebenfalls können bei Schulungen des Panels einzelne Probanden identifiziert werden, deren Leistung nicht mit dem Panel übereinstimmt und somit aus der weiteren Untersuchung ausschließen (ISO 11132, 2017). Die ISO 11132 ist eine einheitliche und definierte Darstellung der Möglichkeit, die Leistung eines deskriptiven Panels in wenigen Schritten, durchzuführen. Somit soll eine einheitliche und anerkannte Prüfmethode, die zahlreichen und verschiedenen Methoden, die oft mit einem hohen statistischen Aufwand verbunden sind, ersetzen. Die DIN EN ISO 11132:2017 wurde ohne Abänderung vom Europäischen Komitee für Normung (CEN-CENELEC) genehmigt.

Die Auswertung der Profildaten erfolgt mit der Software ANOVA. Diese Daten macht sich die DIN EN ISO 11132 zu nutzen. Somit ist hier der Vorteil, dass keine weitere Software benötigt wird, um die Panelleistung zu bewerten.

2.4 Rechtliche Grundlagen und Begriffsbestimmung der verwendeten Fruchtsäfte

Übliche auf dem Markt angebotene Orangensaftes unterscheiden sich in ihrer chemischen aber auch physikalischen Struktur. So können diverse Orangensaftes anhand ihrer charakteristischen Geschmacks- und Aromaeigenschaften deutlich voneinander unterschieden werden (Chambers et al. 2003). Dies hängt von der Orangenart, Beimischungen diverser anderer Citrusfrüchte oder Zusatzstoffe, aber auch von der Verarbeitungsart in der Industrie, ab. Ein wichtiger gesetzlicher Faktor für die Unterscheidung der Säfte ist der Fruchtsaftanteil. Die verwendeten Säfte in dieser Bachelorarbeit sind verschiedene Orangensaftes, die in diesem Kapitel definiert werden.

2.4.1 Fruchtsaft aus Direktsaft

Direktsaft ist ein Fruchtsaft, der direkt aus der frischen Frucht gepresst, filtriert und gegebenenfalls pasteurisiert und abgefüllt ist. Dieser Saft beinhaltet somit 100% Fruchtsaftgehalt und darf nicht extra nachgesüßt werden. Ein Direktsaft, der z.B. als Saft zum direkten Verzehr an einem Verkaufsstand angeboten wird, muss nicht zwingend pasteurisiert werden (DLG Warenkunde). Laut der Definition der Fruchtsaft- und Erfrischungsgetränkeverordnung sind Fruchtsäfte gärfähige, aber nicht gegorene Produkte, die aus einem genießbaren Teil gesunder und reifer Früchte entstehen. Fruchtsäfte können aus einer oder mehreren Früchten gewonnen werden, müssen aber die für die Frucht charakteristische Farbe, das Aroma und den Geschmack, aufweisen. Zusätzlich ist in der FrSaftErfrischGetrV vermerkt, dass bei Zitrusfrüchten der Fruchtsaft aus dem Endokarp (Fruchtfleisch) stammen muss, mit Ausnahme des Limettensaftes, der auch aus ganzen Früchten gewonnen werden kann.

2.4.2 Fruchtsaft aus Fruchtsaftkonzentrat

Ein Fruchtsaft aus Fruchtsaftkonzentrat ist ein Saft, der, nachdem er gepresst wurde, größtenteils physikalisch vom Wasser befreit wurde. Dabei wird der Saft um ein Sechstel verdichtet. Diese Methode wird meistens bei Orangensaften verwendet, um Transportkosten aus weiteren Teilen der Welt, von ganzen Früchten bzw. von Fruchtsäften, die nahezu komplett aus Wasser bestehen, zu verringern. Das Konzentrat wird also im Ursprungsort hergestellt und nach Deutschland befördert, wo es dann mit Trinkwasser rückverdünnt, anschließend pasteurisiert und verpackt wird (DLG Warenkunde). Die FrSaftErfrischGetrV gibt an, dass nur schonende Verfahren, d.h. Verfahren, die die physikalischen, chemischen, organoleptischen und nährstoffbezogenen Merkmale eines durchschnittlichen Orangensaftes nicht zerstören, verwendet werden dürfen.

2.4.3 Nektar

Nektare haben im Vergleich zu Direktsaft einen wesentlich reduzierten Fruchtsaftanteil. Der Grund hierfür ist, dass es Obstsorten gibt, die z.B. einen zu hohen Säuregehalt (z.B. Sauerkirschen und Johannisbeeren) oder einen hohen Fruchtfleischanteil (z.B.: Aprikosen und Bananen) verfügen. Somit wäre eine Produktion eines reinen Saftes technologisch nicht möglich. Des Weiteren wäre der Saft ungenießbar (DLG Warenkunde). Ein weiterer Grund für den geringen Fruchtanteil im Nektar, ist das Einsparen von Fruchtsaft und Produktionskosten. Somit können

günstigere Säfte wie z.B. Orangenfruchtnektar, hergestellt werden. Laut der deutschen Fruchtsaftverordnung liegt der Fruchtanteil bei 25-50%. So hat z.B. ein Passionsfrucht- und Johannisbeerfruchtnektar einen Fruchtgehalt von 25%, Brombeernektar 40% und Pfirsichnektar 50%. Des Weiteren dürfen bis zu 20% des Gesamtgewichtes des fertigen Erzeugnisses, diverse hinzugefügte Zuckerarten oder Honig ausmachen (FrSaftErfrischGetrV).

2.4.4 Fruchtsaftgetränke und Fruchtsaftschorlen

Fruchtsaftgetränke zählen zu den Erfrischungsgetränken und können mit oder ohne Kohlensäure angeboten werden. Der Fruchtgehalt liegt dabei bei 6-30% und ist abhängig von der Fruchtsorte. So liegt der Fruchtsaftgehalt bei Zitrusfrüchten bei ca. 6% und Kernobst z.B. Apfel oder Trauben weisen 30% auf. Der Fruchtanteil kann aus frisch gepressten Früchten oder aus einem Konzentrat bestehen, welchem dann Trink- oder Mineralwasser beigefügt werden. Üblicherweise werden den Fruchtsaftgetränken Zucker, Säuren, die nicht unbedingt aus den verwendeten Früchten stammen müssen, und Aromastoffe zur Verstärkung des Geschmacks hinzugefügt. Fruchtsaftschorlen sind ebenfalls Erfrischungsgetränke mit einem Fruchtanteil von mindestens 50%, die mit Kohlensäure versetzt sind. Zu den beliebtesten Fruchtsaftschorlen zählt die Apfelschorle (DLG Warenkunde).

3 Material und Methoden

3.1 Profilprüfung

Die Profilprüfung ist ein deskriptives Verfahren, mit dem die Wahrnehmung des Menschen messbar gemacht wird. Dadurch können unterschiedliche Produkte und Produkteigenschaften identifiziert und beschrieben werden. Somit ist das Ziel der Profilprüfung, die eindeutige Quantifizierung der Produkte durch den Einsatz menschlicher Sinnesorgane. Übliche deskriptive Prüfungen beschreiben durch die Übersetzung der wahrgenommenen Sinneseindrücke des Menschen, welche Produkteigenschaften dazu führen, dass die Produkte eine Akzeptanz oder Ablehnung hervorrufen. Die Produktprofile der Profilprüfung sind wertfrei und geben keine Auskunft über die Akzeptanz der Teilnehmer. Lediglich dienen diese Profile zu einer Beschreibung der Merkmalseigenschaften sowie ihrer Intensität, z.B. die Intensität der Süße eines Orangensaftes. Das Prüfverfahren kann eingesetzt werden, um Marktforschungen, Marketing und Produktentwicklungen anzutreiben, indem die Profildaten mit den Daten der Konsumentenurteile verknüpft

werden. Zusätzlich können chemisch-physikalische Werte eingebunden werden, um ideale Mischungsverhältnisse, Aromen und Rezepturbestandteile für die Produktüberarbeitung und -entwicklung, herauszufinden (C. Rummel, 2002).

Da der Mensch als Messinstrument dient, wird für die Profilprüfungen ein geschultes (deskriptives) und motiviertes Panel benötigt, welches mit dem zu testenden Produkt vertraut ist (M. Nießen, 2007). Das Panel dient dazu, dass die durch das Produkt subjektiv aufgenommenen Sinnesindrücke zu objektiven und produktspezifischen Eigenschaften gebildet werden. Die sensorischen Produkteigenschaften müssen zuverlässig und reproduzierbar sein. Hierfür dient ein naives Panel, welches aus Konsumenten besteht, die das Produkt konsumieren, aber eine Gleichgültigkeit gegenüber den sensorischen Eigenschaften und der Tatsache, wieso das Produkt schmeckt oder nicht schmeckt, aufweisen. Das deskriptive Panel soll aus Personen bestehen, die für mehrere Wiederholungen zeitlich verfügbar sind, sowie ein langfristiges Interesse in der Mitarbeit vorweisen. Zudem dürfen die Probanden keine allergischen Anzeichen auf die Produkte, die analysiert werden sollen, hervorbringen und für das Verstehen und Bewältigen der Aufgabenstellung qualifiziert sein. Weitere Kriterien sind die Fähigkeiten des Probanden, in einer Gruppe zu arbeiten, z.B. bei Panelschulungen, sowie kein völliges Fehlen des Geschmacks- und Geruchssinns und ein gutes Geruchs- und Geschmacksgedächtnis, die zu einer idealen Wiederholbarkeit dienen (C. Rummel, 2002). Werden diese Paneeleigenschaften erfüllt, kann ein Panel aus 7-13 Probanden erstellt werden (D. Paul und K. Schüssler, 2002). Eine Panelgröße von mehr als 12 Probanden ist nicht empfehlenswert, da das Panel kaum handhabbar und Gruppendiskussionen schwer steuerbar sind (C. Rummel, 2002). Nach der DIN 10967 ist eine Mindestanzahl von 6 Probanden notwendig

Der Ablauf einer Profilprüfung kann in fünf Schritten definiert werden. Im ersten Schritt kommt es zu einer Auswahl von geeigneten Prüfpersonen. Die Prüfpersonen müssen unter anderem die eben genannten Voraussetzungen erfüllen, aber auch konzentrationsfähig bei längeren Sitzungen bleiben, sowie ein gutes sensorisches Gedächtnis für die verschiedenen Durchführungstage aufweisen. Der zweite Schritt, ist das Training der Prüfpersonenanwärter. In diesem Teil, werden die Probanden mit den Produkten vertraut gemacht und das Panel beschreibt die Produkte mit Attributen. Diese Attribute werden in der Gruppe diskutiert, es werden Ankerpunkte und Referenzmuster für jedes Attribut festgelegt. So wird von dem Panel z.B. für das Attribut „orangig im Geschmack“ bei dem Produkt Orangensaft, „Wasser“ als ein niedriger Ankerpunkt und „frisch gepresster Organgensaft“ für einen hohen Ankerpunkt, festgelegt. In Gruppendiskussionen werden Attribute, die zu einem erhöhten Rauschen der Daten oder zur Verwirrung bei den Teilneh-

mern führen, unpräzise formuliert oder eine Hedonik aufweisen, eliminiert. Die Anzahl der Merkmalseigenschaften liegt nicht höher als bei 15 Attributen (D. Paul und K. Schüssler, 2002). Es ist darauf zu achten, dass das Panel die Attribute und deren Ankerpunkte versteht und auch die Skalen richtig anwenden kann. Die Anwendung der Skalen muss eindeutig geklärt werden, z.B. wenn ein Attribut nicht auf das Produkt zutrifft, dass dann die Skalenmarkierung auf „0“ gesetzt wird und es zu keiner Angabe von kleinen Werten kommt. Hierbei ist es wichtig, dass sich das gesamte Panel daran beteiligt ihre Ideen und Empfindungen mitzuteilen, sodass genügend Attribute zusammenkommen, die das Produkt beschreiben. Im weiteren Verlauf wird das Panel geschult. Dieser dritte Akt im Aufbau der Profilprüfung ist wichtig, damit das Panel mit ihren Attributen üben kann und somit ein Verständnis aufbaut. Außerdem können in einer letzten Teamsitzung fragwürdige Attribute oder Ankerpunkte diskutiert und verändert werden. Der Testdurchlauf und die gesamte Profilierung soll im Idealfall in Räumen, die nach der ISO 8589 ausgestattet sind und über EDV verfügen, um eine ideale Datenerfassung mittels CASA (Computer Aides Sensory Analysis) oder FIZZ zu gewährleisten (C. Rummel, 2002), durchgeführt werden. Eine Schulung dient dazu, die Genauigkeit der Probanden, d.h. die Trennschärfe zwischen verschiedenen Proben, zu gewährleisten. Außerdem muss eine Vergleichbarkeit mit anderen Panelmitgliedern gegeben sein. Weicht ein Proband stark von dem Gesamtpanel ab, so ist ein Gespräch und evtl. eine weitere Schulung mit dem Probanden durchzuführen. Reproduzierbarkeit, d.h. eine Wiederholbarkeit der Messung von jedem Probanden in verschiedenen Sitzungen, ist ebenfalls ein wichtiger Faktor, die eine erfolgreiche Schulung ausmacht. Somit ist die Leistungsfähigkeit der Probanden für eine erfolgreiche Profilierung essentiell (Meier, 2006).

Nachdem der Schulungsprozess abgeschlossen ist, führen die Probanden in mehreren Sitzungen die Profilprüfung durch. Hierbei werden dem Probanden verschiedene Proben eines Produktes gereicht, die jeweils einer dreistelligen Verschlüsselung unterliegen. Die Verschlüsselung der Proben ist randomisiert. Es sollten nicht mehr als acht Proben gereicht werden, um eine Ermüdung der Sinnesorgane bzw. ein Konzentrationstief zu verhindern. Des Weiteren ist es ratsam, dass sehr intensive Proben, wie auch Proben mit einem starken Nachgeschmack, zum Ende gereicht werden. Der Proband analysiert das Produkt mit seinen Sinnen und trägt die Intensitäten des Produktes in eine Intervallskala ein. Diese Skala ist vorgefertigt mit den zuvor gemeinsam festgelegten Attributen. Somit kann der Proband bei z.B. „sauer im Geschmack“, die Intensität der empfundenen Säure (Intensität nimmt von links nach rechts zu) in die Skala eintragen.

Ist die Analyse der Proben bzw. sind alle Sitzungen abgeschlossen, werden die Rohdaten gesammelt und verarbeitet. Dieser letzte Schritt, ist die Auswertung der gesammelten Daten. An-

hand der Rohdaten und deren Auswertung ist erkennbar, wie das Panel die Proben skaliert hat. Dadurch kristallisiert sich heraus, wie gleich, aber auch unterschiedlich die einzelnen Probanden die Skalen angewendet haben. Ebenfalls kann die Wahrnehmung der einzelnen Merkmalseigenschaften ermittelt werden (D. Plaul und K. Schlüssler, 2002). Durch sogenannte Egg-SHELL-Plots, Intensitäts-Linien-Diagramme, aber auch PCA's, können die Unterschiede graphisch dargestellt werden. Zur Auswertung wird die Software „FIZZ“ herangezogen, wo eine Varianzanalyse (ANOVA= analysis of variance) durchgeführt wird. Hierfür müssen die Rohdaten in einer Excel-Tabelle vorbereitet werden.

Tabelle 1: Rohdatenbeispiel der ANOVA-Tabelle

| Prüfperson | Probe | Wiederholung | Attribut A1 | Attribut A2 | Attribut n+1 |
|-------------------|--------------|---------------------|--------------------|--------------------|---------------------|
| T1 | O-Saft X | 1 | #Wert | #Wert | #Wert |
| T2 | O-Saft X | 1 | #Wert | #Wert | #Wert |
| T3 | O-Saft X | 1 | #Wert | #Wert | #Wert |
| T4 | O-Saft X | 1 | #Wert | #Wert | #Wert |
| T1 | O-Saft X | 2 | #Wert | #Wert | #Wert |
| T2 | O-Saft X | 2 | #Wert | #Wert | #Wert |
| T n+1 | ... | ... | ... | ... | ... |

Die Tab. 1 zeigt eine beispielhafte Datenvorbereitung.

Wenn alle Wiederholungen eines Produktes abgeschlossen sind, wird unter „Probe“ der nächste Saft aufgeführt und weiter, die Wiederholungen dargestellt. Mit diesen Rohdaten und der Software „FIZZ“ können nun verschiedene Diagramme dargestellt werden.

In der Abbildung 1 ist ein Beispiel eines Liniendiagramms dargestellt, in dem die Intensitätsbewertung des Attributes A3 der fünf Proben durch die sieben Prüfpersonen aufgezeigt wird. Hiermit kann die Leistungsfähigkeit einzelner Probanden graphisch dargestellt werden.

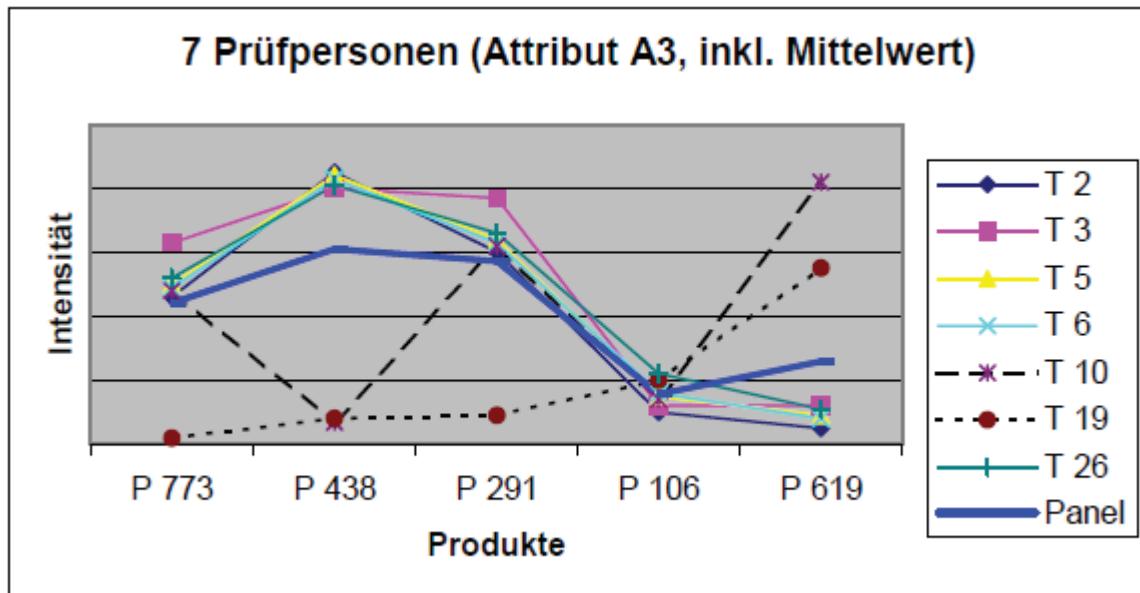


Abbildung 1: Beispiel einer Intensitätsauswertung des gesamten Panels (Meier, 2006)

Alle Probanden, ausgenommen Proband T15 und Proband T19, haben die Intensitäten sehr ähnlich wahrgenommen. Hier ist die Folge, dass die „Ausreißer“ eine andere Wahrnehmung besitzen als der Rest des Panels oder das Attribut und die Ankerpunkte unverständlich erscheinen. Hier wäre ein Gespräch zur Klärung des Missstandes eine Option.

Die Auswertung der Daten anhand der Varianzanalyse des gesamten Panels kann in einer Tabelle, der sogenannten ANOVA-Tabelle, dargestellt werden. Diese Methode eignet sich um einen Überblick über die Messzuverlässigkeit des Panels und die von den Panels wahrgenommene Unterschiede in den Produkten zu verschaffen (Meier, 2006). Mit der ANOVA wird ebenfalls untersucht, ob sich die Werte von den metrischen Variablen unterscheiden (Ellert, 2006).

In der Tabelle 2 ist ein Beispiel einer verkürzten und vereinfachten ANOVA-Tabelle anhand des Attributes „säuerlich Geruch“ zu sehen. In der linken Zeile werden die Prüfer, Produkte sowie die Interaktion zwischen Prüfer und Produkt dargestellt.

Tabelle 2: Beispiel einer Varianzanalyse

| Variation | F-Wert | Signifikanz |
|-------------|--------|----------------|
| Prüfer | 11,8 | <0,0001 *** |
| Produkt | 8,39 | <0,0001 *** |
| Interaktion | 3,35 | <0,0001 *** |

*Signifikant (5%), **Hoch-Signifikant (1%), ***Höchst-Signifikant (0,1%)

Der F-Wert ist die Fehlerwahrscheinlichkeit. Je höher der F-Wert z.B. bei dem Produkt ist, desto unterschiedlicher sind die Produkte anhand des Attributes. In der rechten Zeile wird das Signifikanzniveau dargestellt. Die Prüfer sind mit der Höchst-Signifikanz belegt, dies bedeutet, dass die Prüfer anhand des Attributes „säuerlich Geruch“ unterschiedlich bewerten. Dies kann daran liegen, dass die Prüfer die Skalenbereiche unterschiedlich nutzen, jedoch rangfolgetechnisch die gleiche Aussage haben. Bei einem geschulten Panel müsste eine weitere Schulung durchgeführt werden, damit die Signifikanz abnimmt und das Panel stimmig anhand des Attributes bewertet. Die Produkte sind anhand des Attributes höchstsignifikant unterschiedlich. Dieses ist bei unterschiedlichen Produkten zu erwarten. Schlecht wäre eine Höchst-Signifikanz, wenn die Produkte sehr ähnlich bzw. wenn gleiche Produkte mit dabei wären. Anhand dieses Beispiels, kann die Behauptung aufgestellt werden, dass die Produkte anhand des Attributes „säuerlich Geruch“ deutlich voneinander unterschieden werden können. Eine Höchst-Signifikanz in der Interaktion hat die Aussagekraft, dass die Intensitätsmessung der Produkte bzw. Produkteigenschaften abhängig von den Prüfern ist. Somit besteht eine Wechselwirkung zwischen Produkt und Prüfern, d.h. es gibt Personen im Panel, die regelmäßig anders bewertet haben. Hier besteht die Aufgabe des Prüfungsleiters darin, sich die Rohdaten genau anzuschauen und herauszufinden, welche Prüfer für diese Interaktion verantwortlich sind. Idealerweise ist bei einem geschulten Panel mit unterschiedlichen Produkten keine Signifikanz bei den Prüfern und der Interaktion, aber eine Höchst-Signifikanz bei den Produkten, erwünscht. Dies hätte die Aussage, dass die Prüfer alle

gleich bewerten, die Produkte unterschiedlich anhand des Attributes sind und es gäbe keine Wechselwirkungen zwischen dem Prüfer und den Produkten (Meier, 2006).

3.2 Hedonische Prüfverfahren

3.2.1 Verbrauchertest und 9-Punkt hedonische Skala

Im Rahmen der Bachelorarbeit wird ein Verbrauchertest durchgeführt, um die Daten mit den Ergebnissen aus der Profilprüfung zu vergleichen. Die Methode des Verbrauchertests beschäftigt sich direkt mit dem Konsumenten. Bei einer Beliebtheitsprüfung (hedonische Prüfung) werden die Konsumenten, die keiner sensorischen Schulung unterliegen, herangezogen. Durch den Einsatz des Verbrauchers werden Rückschlüsse auf eine Akzeptanz des Produktes gewonnen. Die Durchführung eines Verbrauchertestes kann auf verschiedene Weise stattfinden. So existieren sogenannte „Home-Use-Tests“, wo die Konsumenten das Produkt, meistens über einen Zeitraum von 1-4 Wochen, zugeschickt bekommen und dieses in einer gewohnten Umgebung und Atmosphäre testen. Dabei werden Fragebögen über das Produkt ausgefüllt. Dieser Test ist realitätsnah, da der Konsument üblicherweise diverse gekaufte Artikel ebenfalls zu Hause konsumiert. Die Teilnehmer repräsentieren dabei die Zielgruppe. So wird z.B. ein zu testendes Nahrungsergänzungsmittel für Schwangere nicht an alleinstehende Männer oder Senioren gesandt. Ein Nachteil der Methode ist, dass die Probanden das Produkt unterschiedlich testen und deutliche Ablenkungsfaktoren vorliegen. Um den Ablenkungsfaktoren entgegenzuwirken, kann ein Verbrauchertest in einem Labor bzw. Sensoriklabor durchgeführt werden. Hier kann die Produktvorbereitung und -präsentation besser gesteuert und kontrolliert werden. Üblicherweise werden Verbrauchertests mit einer bestimmten Zielgruppe, wie z.B. Schulkindern, vor einer Schule durchgeführt. Produkttests mit Säften, die keiner bestimmten Zielgruppe unterliegen, können z.B. vor Kaufhäusern und Menschen durchgeführt werden. Probleme stellen ebenfalls die Ablenkung der Prüfer durch Umgebungsumstände, wie z.B. diverse Gerüche und die Geräuschkulisse, dar (Carr et al., 2015).

Bei den drei genannten Verbrauchertests wird die subjektive Akzeptanz des Probanden des jeweiligen getesteten Produktes auf eine sogenannte 9-Punkt hedonische Skala bzw. Akzeptanzskala notiert. Diese Skala wurde von Jones et al. (1955), Peryam und Pilgrim (1957) entwickelt (Ellert, 2006).

Tabelle 3: 9-Punkt hedonische Skala

| Einstufung der Beliebtheit | | 9-Punkt hedonische Skala |
|-------------------------------|------------------------|------------------------------|
| 9 | Gefallen- Bereich | gefährlich außerordentlich |
| 8 | | gefährlich sehr |
| 7 | | gefährlich ziemlich |
| 6 | | gefährlich etwas |
| 5 | Neutraler-Bereich | weder gefällt noch missfällt |
| 4 | Missfallen- Bereich | missfällt etwas |
| 3 | | missfällt ziemlich |
| 2 | | missfällt sehr |
| 1 | | missfällt außerordentlich |

Die Skala bzw. Tabelle 3 enthält 9 Stufen, die in zwei Notenbereichen unterteilt ist. (9-6 und 4-1). Die Notenbereiche 9-6 geben den Bereich „gefährlich“ und die Notenbereiche 4-1 den Bereich „missfällt“, an. Der Bereich „5“ ist der Wendepunkt der Skala und gibt an, dass das Produkt dem Probanden weder gefällt noch missfällt. Um die Aussagekraft eines akzeptierten Produktes ausführen zu dürfen, benötigt es mindestens 80% der Urteile im „gefährlich“-Bereich (9-6). (Fliedner, I. und Wilhelmi, F., 1995). Die Daten werden mit der einfaktoriellen Varianzanalyse, unter Zuhilfenahme der Software „Fizz“, ausgewertet.

3.2.2 Check-All-That-Apply

Ein Verbrauchertestbogen kann auf verschiedene Weise gestaltet werden. Zu einem Beliebtheits-test mit einer 9-Punkt hedonischen Skala werden für die Problemstellung dieser Bachelorarbeit „CATA“ (Check-All-That-Apply) und JAR-Intensitäts-Fragen (Just-About-Right) herangezogen. CATA dient der sensorischen Produktbeschreibung. Die Beschreibung erfolgt bei wenig komplexen Lebensmitteln oder bei Vergleich von mehreren ähnlichen Produkten mit verschiedenen Rezepturen. Diese Methode ist eine rein beschreibende Methode, sodass eine Wertung oder In-

tensitätsbeschreibung nicht stattfindet. Dadurch ist die CATA-Methode eine „verbal-based“-Methode (Derndorfer und Schneider-Häder, 2016). Die Information, die der Test zum Vorschein bringt, besteht aus der produktspezifischen und hedonischen Wahrnehmung der Konsumenten. Bei CATA stellt sich die Frage, welche Erwartungen bzw. Vorstellungen der Konsument von einem Lebensmittel hat und welche Eigenschaften für eine bessere Produktakzeptanz sorgen. Diese Fragestellung wird mit der Zuhilfenahme eines hedonischen Tests mit einer 9-Punkt hedonischen Skala beantwortet. Lebensmittelproduzenten schätzen die Produktwahrnehmung der Verbraucher, da diese essentiellen Informationen für z.B. die Produktentwicklung und Produktüberarbeitung, sowie Marketing, die Produktionsprozesse und die Qualitätssicherung, hervorbringt (Ares et al., 2010). Das Testverfahren kann als Schnellmethode klassifiziert werden und eignet sich, um kostenintensive und zeitaufwändige Methoden zu umgehen. Hierfür erhält der Proband, ähnlich wie in der Abbildung 2, zu seinen verschlüsselten Proben einen Verbrauchertestbogen mit diversen vorgegebenen Attributen (z.B. süß, sauer, etc.).

Bitte kreuzen Sie alle sensorischen Eigenschaften an, die auf den jeweiligen Tomatensaft zutreffen:

| | Probe 378 | Probe 565 | Probe 410 |
|--------------|--------------------------|--------------------------|--------------------------|
| süß | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| sauer | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| umami | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| fruchtig | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| grün, grasig | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| würzig | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| muffig | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| ... | | | |

Abbildung 2: Beispiel eines CATA-Prüfformulars (Derndorfer und Schneider-Häder, 2016)

Der Verbraucher kann nun die Attribute ankreuzen, die auf das Produkt zutreffend sind (=„Check All That Apply“). Dabei wird die Reihenfolge der Attribute nach der Verkostungsreihenfolge (Aussehen, Geruch, Geschmack, Textur) durch den Versuchsleiter festgelegt (Beeren und Schneider-Häder, 2015). Im Gegensatz zu den JAR-Fragen, kommt es hier zu keiner Beschreibung der Intensitäten. Die Wahl der Attribute erfolgt durch das geschulte Panel aus der Profilierung (Meullenet et al. 2009). Somit arbeitet das Konsumentenpanel mit den gleichen At-

tributen wie das geschulte Panel. Zusätzlich werden hedonische Begriffe wie, „harmonisch“, „lecker“ und „erfrischend“, beigefügt. Es ist wichtig dem Panel Attribute zur Verfügung zu stellen, denn würde jeder Proband mit seinen eigenen Attributen arbeiten, so wäre die Auswertung sehr langwierig, da jedes Attribut separat interpretiert und mit ähnlichen Attributen verglichen, sowie zusammengefasst werden müsste (Carr, Civille und Meilgaard, 2015). Dies würde die gesamte Auswertung unüberschaubar gestalten. Die Wahl der Attribute durch das geschulte Panel hat den Vorteil, dass die Attribute umfassend und genau beschrieben sind (Meullenet et al. 2009). Des Weiteren kann in der Industrie eine Attributewahl z.B. durch eine Rezepturveränderung, vorgenommen werden. So kann z.B. bei einer Reduzierung der Säure, das Attribut „sauer“, verwendet werden. Bei der Wahl der Attribute sollten Begriffe gewählt werden, die der Konsument versteht. Attribute wie adstringierend versteht der Verbraucher nicht. Dadurch bilden die CATA-Fragen eine deutlich niedrigere Auswirkung auf die Konsumentenwahrnehmung und die Gesamtbeliebtheit, als die JAR-Methode und diverse Intensitätsskalen (Ares et al., 2010). In der Auswertung, die mit Hilfe der Software Xlstat Version 19.7 2017 durchgeführt wurde, wird deutlich, wie viele Konsumenten jedes Attribut als Produktbeschreibung nutzen (Beeren und Schneider-Häder, 2015). Diese Daten können zusammen mit z.B. den Daten aus der Profilprüfung weiterverwendet werden, um eine weitere Methode, z.B. „Preference Mapping“, durchzuführen. Mit Zuhilfenahme von „Preference Mapping“ wird die Frage beantwortet, warum der Konsument ein bestimmtes Produkt bevorzugt (Ares et al., 2010).

3.2.3 Just-About-Right

3.2.3.1 Grundlagen und Aufbau

Das Arbeiten mit den JAR-Fragen (Just-About-Right) ist eine weitere Methode, die es erlaubt eine Aussage über die Intensität der Attribute und der hedonischen Wahrnehmung des Probanden zu geben. Zum ersten Mal wurde die JAR-Methode von Shepherd et al. im Jahr 1989, zur Beurteilung von Suppen angewendet, die es ermöglicht die Kundenakzeptanz zu überprüfen (Heymann und Lawless, 2010). Hier werden von dem Probanden drei Bereiche gleichzeitig durchgeführt: Bewertung des Produktes, Vergleich mit seinem vertrauten Produkt und Klarstellung, welche Attribute bzw. Merkmalseigenschaften zu optimieren sind (Schröder, 2008). Die Aussage ist leicht zu interpretieren und gibt z.B. direkte Anzeichen, welche Eigenschaft an einem Produkt verändert werden soll. Bei der JAR-Methode wird üblicherweise mit einer bipolaren 5-Punkt-Skala gearbeitet, die in 3 Kategorien eingeteilt ist. Diese Skala bzw. Tabelle 4 ist

einem Attribut, wie z.B. „sauer“ untergeordnet und besitzt zwei Enden, die mit z.B. „etwas zu schwach“, sowie „viel zu schwach“ und gegenüberliegend mit „etwas zu intensiv“, sowie „viel zu intensiv“ gekennzeichnet ist. Der mittlere Punkt oder auch Zentralpunkt genannt, ist als „genau richtig“, bzw. „just about right“ gekennzeichnet (Heymann und Lawless, 2010).

Tabelle 4: JAR-Kategorien anhand des Attributes „sauer“

| Geschmack: sauer | Kategorie |
|-------------------|---------------|
| viel zu intensiv | zu viel |
| etwas zu intensiv | |
| genau richtig | genau richtig |
| etwas zu schwach | zu wenig |
| viel zu schwach | |

Die Attribute sollen so gewählt werden, dass ein ungeschultes Panel diese Merkmalseigenschaften versteht und somit die Skala nachvollziehen kann, z.B.: salzig, süß, sauer, bitter, usw. Üblicherweise wird eine Merkmalseigenschaft benotet, es können jedoch auch gegensätzliche Attribute (z.B. feucht-knusprig) benutzt werden. Die Verwendung von gegensätzlichen Attributen ist störanfälliger, da diese zu einer Fehlinterpretation bzw. Skalenveränderung führen kann. Außerdem sollte, um Verwirrungen zu verhindern, eine einheitliche Struktur bei den JAR-Fragen durchgeführt werden, d.h. entweder gegensätzliche Begriffe oder ein Attribut an beiden Seiten der Skala. Es empfiehlt sich gleiche Begriffe für die Skala zu benutzen, da es schwierig ist, für Geschmack und Geruch gegensätzliche Begriffe zu finden (Moskowitz et al., 2003). Die Auswertung der JAR-Fragen erfolgt in Kombination mit den Beliebtheitsdaten des Panels. Mit 70% der Nennungen in der JAR-Kategorie, gilt das Produkt als akzeptiert (Meullenet et al., 2007). Bei 50% JAR, gilt das Produkt als ausreichend akzeptiert (Schröder, 2008). Ebenfalls kann erst bei mindestens 20% der Nennungen im Bereich „zu viel“ oder „zu wenig“, eine signifikante Aussage getroffen werden und somit ein Eingreifen in das Produkt (Meullenet et al., 2007). Die Auswertung der JAR-Fragen erfolgt mit Hilfe der „FIZZ“- und „Xlstat“-Software.

3.2.3.2 Penalty Analysis

In der Auswertung der JAR-Fragen ist es von besonderer Bedeutung, die Penalty Analyse (deutsch: *Analyse der Strafe/Strafpunkte*) durchzuführen. Diese Methode gibt Auskunft über die Strafen der verschiedenen Produkteigenschaften, mit der die Attribute herauskristallisiert werden

und deshalb verändert werden müssen. Dies ist notwendig um die Gesamtbeliebtheit des Produktes zu verbessern. Somit werden die Beliebtheitsdaten der hedonischen Prüfung mit den Daten der JAR-Fragen verknüpft. Im Prinzip erfolgt eine Klarstellung, um welchen Wert die Produktbeliebtheit reduziert ist, wenn ein Produkt nicht „gerade richtig“ (just-about-right), also entweder „zu viel“ oder „zu wenig“ bewertet wurde. Aus diesen Informationen wird die sogenannte „Strafe“ berechnet, woraus ersichtlich wird, wie sich nicht ideale Produkteigenschaften auf die Beliebtheit der Produkte auswirken. Die Strafen weisen verschiedene Wichtigkeitsstufen auf, wo die Richtung der Abweichungen (ist „zu viel“ oder „zu wenig“ wichtiger?) identifiziert werden muss. Mit Hilfe eines Vergleiches zwischen den Beliebtheitsbewertungen der Konsumenten, die das Produkt als entweder „zu viel“ oder „zu wenig“ und die Konsumentenbeurteilungen, die das Produkt als „gerade richtig“ empfunden haben, ist diese Identifizierung möglich (Ellert, 2006). Ein weiterer Vorteil, der mit der Berechnung der Strafpunkte zusammenhängt, ist das bessere Verständnis von den Vorlieben und Abneigungen der Konsumenten. Die hierbei zu klärende Fragestellung wäre, ob eine Beeinträchtigung der Gesamtbeliebtheit vorliegt, wenn ein Abweichung einer Produkteigenschaft vom Ideal vorhanden ist. So werden Konsumenten bei einem Produkt, wo z.B. die Intensität der Süße gefragt ist, eine süße Probe mit dem Eindruck „ungesund“ verbinden. Mit Hilfe dieser Analyse kann ermittelt werden, dass Konsumenten, die das Produkt als „zu süß“ bewertet haben, keine negativen Interpretationen hervorrufen, so wie die Konsumenten, die die Süße „genau richtig“ finden. Das Ergebnis wäre, dass nichts an der Produkteigenschaft „Süße“ verändert werden soll (Ellert, 2006).

Um herauszufinden welche Produkteigenschaften verändert werden sollen, können die berechneten „Mean Drops“ (Mittelwerts-Abfälle) und die „Penalties“ (Strafpunkte) als Maß genommen werden. Durch die berechneten Mean Drops in der Kategorie „zu schwach“ und „zu intensiv“ wird deutlich, wie viele Punktabzüge in der Beliebtheit verzeichnet werden, wenn das Produkt in einer der Attribute diese zwei Kategorien hervorbringt. Bei den „Penalties“ bzw. Strafpunkten wird noch einmal deutlich, wie Abzüge in der Beliebtheit resultieren, wenn ein Attribut nicht „JAR“ bzw. „genau richtig“ bewertet wird.

Die Darstellung und Auswertung wird einmal anhand des Beispiels eines Kartoffelchipsverbrauchertests, wo die Attribute „Salzigkeit“, „Süße“, „Säure“ und „Knusprigkeit“ untersucht wurden, gezeigt. Dieses Beispiel stammt von der Homepage Xlstat und beschreibt ein Tutorial zur Auswertung der „Penalty Analyse“.

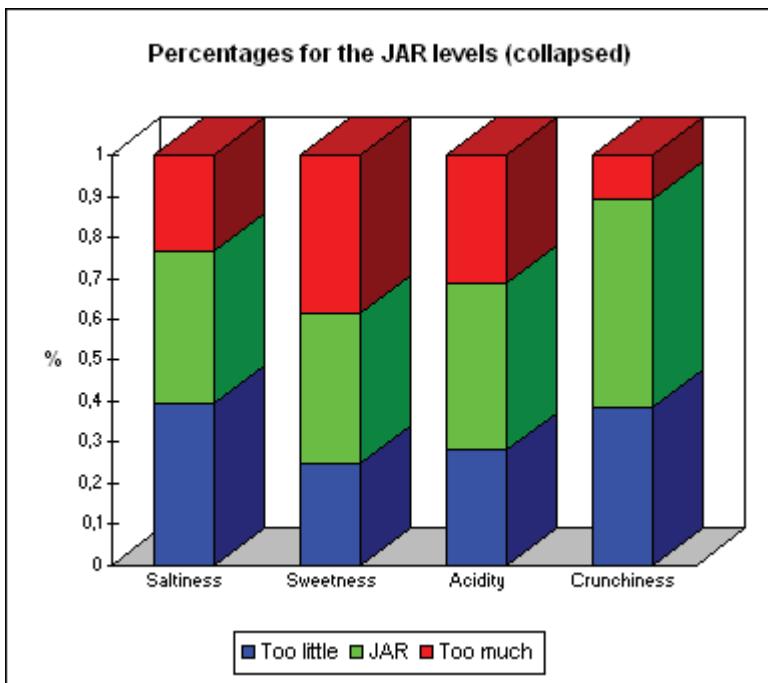


Abbildung 3: Beispiel JAR-Bewertungen (Xlstat,2017)

In der Abbildung 3 ist ein Beispiel zu sehen, wo im ersten Schritt die fünf JAR-Kategorien zu drei Kategorien („zu wenig“, „genau richtig“ und „zu viel“) zusammengefasst wurden. Aus dieser Abbildung geht hervor, wie viel Prozent der Konsumenten für welche JAR-Kategorie abgestimmt haben. Das Programm Xlstat, gibt auch die Tabelle mit den Daten vor, wo die genauen Prozentzahlen der einzelnen Kategorien einzusehen sind.

Die Mean Drops in der Abbildung 4 zeigen genau, in welcher der Attribute und der Kategorie, die höchsten Punktabzüge in der Beliebtheit vorhanden sind.

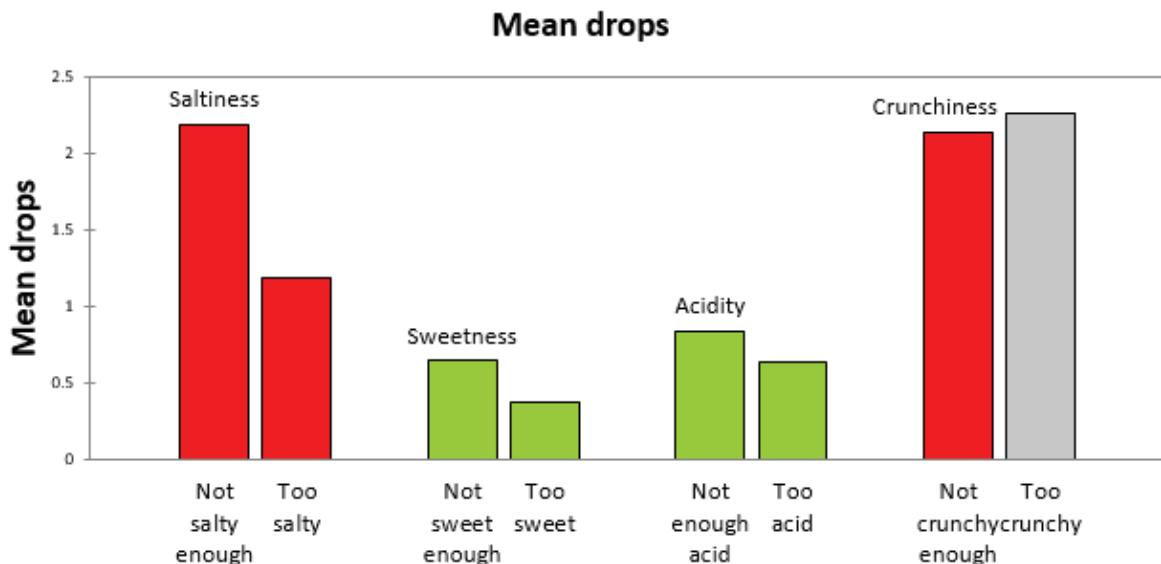


Abbildung 4: Beispiel der Mean Drops (Xlstat, 2017)

So kann anhand von diesem Beispiel verdeutlicht werden, dass, wenn die Produkte nicht salzig genug oder nicht knusprig genug erscheinen, dann mit einer Signifikanz von 5%, Punktabzüge im hohen Bereich von über zwei Punkten anzunehmen sind. Die Bereiche „Süß“ und „Säure“ beinhalten niedrige Punktabzüge, aber keine Signifikanz auf einem Niveau von 5%. Die Knusprigkeit in der Kategorie „zu knusprig“ hat ebenfalls einen hohen Mean Drop, jedoch wurde hierfür keine Signifikanz berechnet, da die Anzahl der Probanden, die das Produkt „zu knusprig“ fanden, unter dem eingestellten Schwellenwert (meistens 20% der Probanden) liegt. Der Schwellenwert wird höher, je weniger Konsumenten am Panel teilgenommen haben. Dadurch kann keine Sicherheit über die Aussage „zu knusprig“ gegeben werden.

In der Abbildung 5 werden die Penalties, also die Strafpunkte des Kartoffelchips-Beispiels, dargestellt. Hier ist zu sehen, dass die höchsten Strafpunkte in der nicht Erfüllung der „JAR“- bzw. „genau richtig“-Kategorie, in der „Knusprigkeit“ und „Salzigkeit“ vorhanden sind.

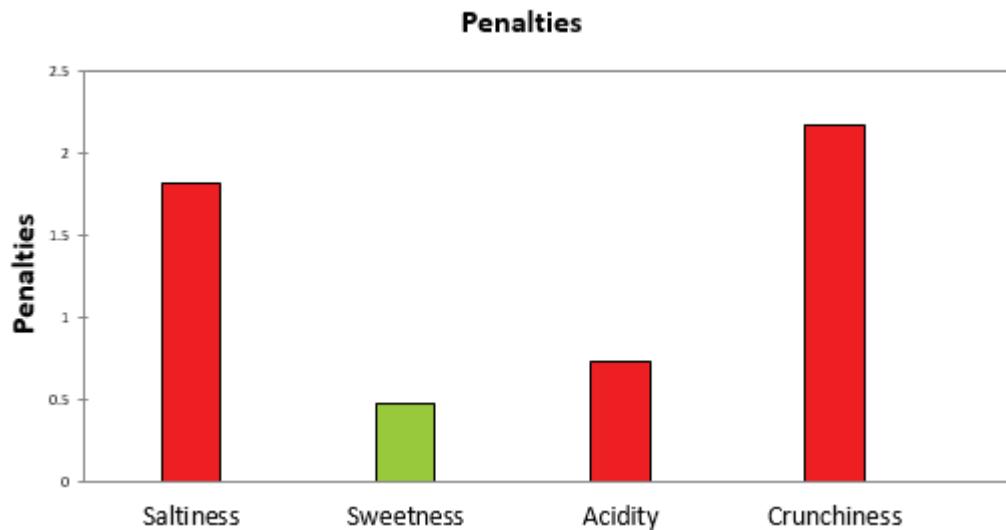


Abbildung 5: Beispiel Penalties (Xlstat, 2017)

Somit kann die Aussage getroffen werden, dass dem Konsumenten besonders wichtig ist, dass das Produkt nicht zu hart, aber auch nicht zu weich ist, bzw. das Produkt eine perfekte Knusprigkeit aufweist. Zusätzlich darf das Produkt nicht fade, aber auch nicht übersalzen sein. Eine genaue Abstimmung ist hier wichtig, um die Beliebtheit des Produktes zu verstärken. Durch die drei Abbildung kann das Resümee gezogen werden, dass das Produkt in die Richtung mehr Salzigkeit und Knusprigkeit verbessert werden sollte, damit die Kundenakzeptanz verbessert wird (Bongartz et al., 2011 und Xlstat, 2017).

3.3 Prüfbogen Verbrauchertest

Für den Verbrauchertest wurde über die Software „FIZZ“ ein Prüfbogen vorbereitet. Der Prüfbogen besteht aus 9-Punkt hedonischen Skalen, JAR-Skalen (5-Punkt) und einem CATA-Prüfteil. Des Weiteren verfügt der Prüfbogen über persönliche Angaben, wie Alter, Geschlecht und die Häufigkeit des Konsumierens von Orangensaft. Die Proben sind mit einem dreistelligen Code angeordnet. Um die Aufgabe und Skalennutzung nachvollziehen zu können, befindet

sich für jede zu bearbeitende Kategorie eine Aufgabenstellung und Skalenerklärung. Der Prüfbogen befindet sich im Anhang 1.1-1.5.

3.4 Preference Mapping

Die multivariate statistische Methode „Preference Mapping“ wird häufig genutzt, um eine Darstellung des Zusammenhangs zwischen den sensorischen Produkteigenschaften und den hedonischen Akzeptanzdaten eines Verbrauchertestes zu ermöglichen (Carr und Meilgaard et al., 2015.). Das Preference Mapping ist ein wichtiges Werkzeug für die Marktforschung (Van kleef et al., 2006) und wird in drei verschiedene Varianten unterteilt: „Internal pre.map.“, „External pref.map.“ und „Extended pref. map.“. Entscheidend für alle drei Varianten ist, dass nicht weniger als 6 Proben in die Untersuchung herangezogen werden. Je mehr Proben untersucht werden und je höhere Unterscheidungskriterien vorhanden sind, desto höher ist die Aussagekraft der Ergebnisse (Bongartz und Mürset, 2011). Preference Mapping versucht die Präferenz bzw. Konsumentenbewertungen, mit den Produkteigenschaften in Beziehung zu bringen und festzustellen wie die Produkteigenschaften, die Konsumentenbeliebtheit beeinflusst (Van kleef et al., 2006). Das Verständnis der Methode beschränkt sich darauf, dass die Annahme getroffen wird, dass Konsumenten und geschulte Prüfer gleich bzw. ähnlich bewerten. Somit stellt sich hier die Frage, wie sehr die Konsumenten ein Produkt mögen. In dieser Methode wird nicht darauf eingegangen, wie das Produkt wahrgenommen wird bzw. findet keine Beschreibung statt. Für Solche Informationen können JAR- und CATA-Fragen zuständig sein (Ares et al., 2010). In dieser Arbeit wird das Extended pref.map. nicht in Betracht gezogen.

3.4.1 Internal Preference Mapping

Das Internal Preferece Mapping (MDPREF: Multidimensional Preference Analysis) nutzt die Beliebtheitsdaten der Konsumenten, die als interne Daten bezeichnet werden, um die Produkte dann anhand eines PCA-Plots (Principial Component Analysis) zu projizieren. Bei dem External pref.map. werden für die Aufgabe die deskriptiven Attribute aus externen Quellen herangezogen (Carr und Meilgaard et al., 2015). Durch die Benutzung des MDPREF kann keine Aussage getroffen werden, warum die Produkte gefallen oder missfallen. (Bongartz und Mürset, 2011, und Schröder, 2008). Somit werden die Präferenzen der Konsumenten in den Vordergrund gestellt (Lawless und Heymann, 2010). Mit Hilfe des internal pref. maps. ist es Betrieben möglich einen

klaren Vorteil in der Handlungsfähigkeit im Marketingbereich und neue Ideen in der Produktentwicklung zu schaffen (van Kleef et al., 2006). Um die Daten der Konsumenten graphisch darzustellen wird eine sogenannte Hauptkomponentenanalyse (HKA) durchgeführt. Das Ziel des HKA ist es eine möglichst geringe Anzahl an Dimensionen zu schaffen, die den größten Anteil der verschiedenen Konsumentenbewertungen überbringen. Dies geschieht durch eine Dimensionsreduzierung, indem das PCA die Komponenten mit dem größten und kleinsten Einfluss bestimmt (Ellert, 2006). Im ersten Schritt werden die Attribut-Daten durch die Konsumenten-Daten ersetzt. Dadurch ergibt sich ein Datensatz, der aus hedonischen Konsumentenbeurteilungen besteht, wo die Produkte (Objekte) in der Zeile und die Konsumentenbewertungen in die Spalte (Variable) eingetragen sind. Mit diesen Datensatz können nun die zwei Hauptkomponenten ermittelt werden, die mit Hilfe des zweidimensionalen Biplots dargestellt werden (Schröder, 2008). Die erste Hauptkomponente ist die Achse, durch die der größte Anteil der Daten erklärt wird. Auf die zweite Hauptkomponente, die senkrecht zur ersten Dimension der Hauptkomponente liegt, entfallen die restlichen Daten. Der zweidimensionale Biplot wird als eine Art Landkarte verstanden, in der die Lage der Konsumenten und der Produkte von besonderer Bedeutung sind. Eine Verwertung und Berechnung der Daten wird automatisch mittels Xlstat Version 2017.7 durchgeführt.

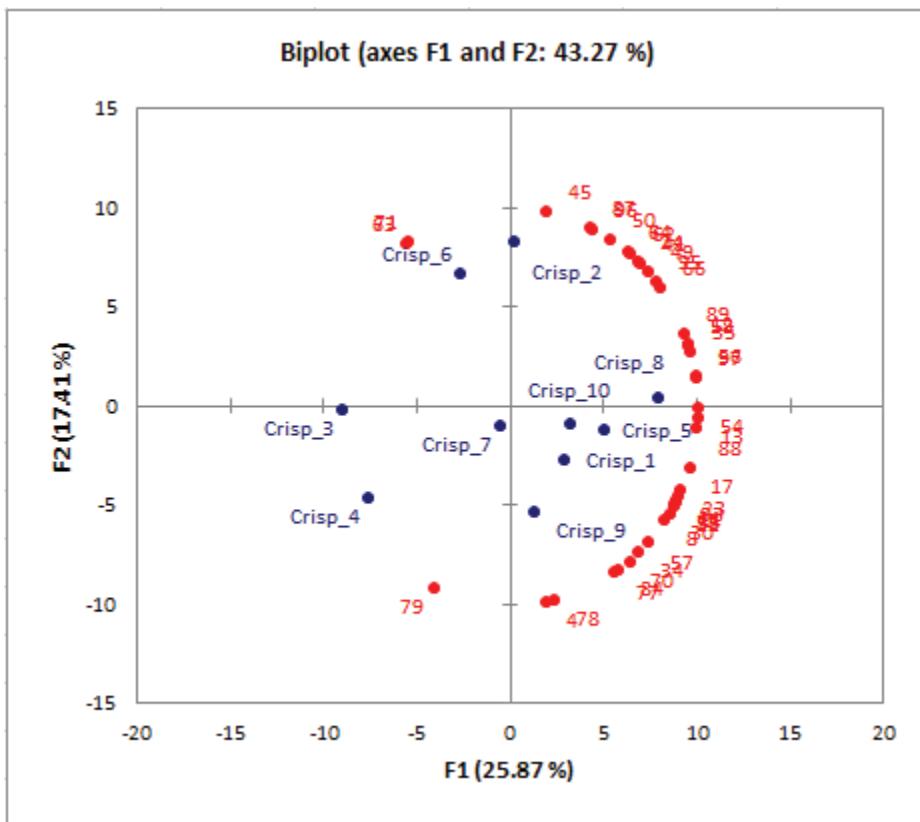


Abbildung 6: Beispiel zweidimensionaler Biplot eines internal pref.maps. (Xlstat, 2017)

In der Abbildung 6 ist ein Beispiel des zweidimensionalen Biplots zu sehen, bei dem die Schritte des Internal Preference Mappings durchgeführt wurden. Die roten Punkte stellen die Konsumenten dar, die für eine bessere Veranschaulichung nicht als Vektoren dargestellt werden. Des Weiteren sind die verschiedenen Produkte, als blaue Punkte, gekennzeichnet. Hierbei handelt es sich um ein Beispiel der Homepage Xlstat, wo verschiedene Kartoffelchipssorten geprüft wurden. Die Interpretation des internal pref maps. ist simple. Je mehr Konsumenten um eine bestimmtes Produkt liegen, desto beliebter ist das Produkt bei den Prüfern. Außerdem wird sichtbar, welche Probanden die Produkte bevorzugen. So ist zu sehen, dass „Crisp 3“, ähnlich wie „Crisp 4“, am weitesten von den Konsumenten entfernt ist und somit keine positive Präferenz erfährt. Die Sorten 1, 2, 5, 8, 9, 10 genießen eine hohe und gute Präferenz.

3.4.2 External Preference Mapping

Um das External Preference Mapping (PREFMAP) durchführen zu können, werden analytische Daten, wie z.B. aus einer Profilprüfung und Konsumentendaten, miteinander verbunden. Somit werden hier die deskriptiven - mit den hedonischen Daten vereint. Diese Methode eignet sich für die Identifizierung wichtiger Faktoren, die für eine Akzeptanz der Produkte von großer Bedeutung sind. Außerdem kann ermittelt werden, in welcher Abhängigkeit die Attribute aus der Profilierung zu der Beliebtheit des Konsumenten aus dem Verbrauchertest, stehen (Bongartz und Mürset, 2011). Hier wird die sensorische Wahrnehmung der Probanden mit in den Vordergrund gestellt (Lawless und Heymann, 2010). Mit Hilfe des external pref. maps. können mehr lebensmitteltechnologische Aufgaben bewältigt werden (van Kleef et al., 2006). Im ersten Schritt wird eine PCA-Analyse mit den externen Daten durchgeführt. Anschließend können die Konsumentendaten mit Hilfe der Regressionsanalyse in den Biplot der deskriptiven Daten eingefügt werden. Die Regressionsanalyse beinhaltet diverse ineinander verschachtelte Modelle: vektorielles-, zirkuläre-, elliptisches- und quadratisches Modell. Diese vier Regressionsmodelle lassen sich wiederum in zwei Arten einteilen: lineare (Vektormodell)- und quadratische (Idealpunkt-Modell, beinhaltet das Zirkuläre, Elliptische und Quadratische) Regression. Das Vektormodell (siehe Abbildung 6) wurde im Kapitel 3.4.1 dargestellt. Die Vektoren wurden dabei für die Übersichtlichkeit entfernt und stattdessen Punkte dargestellt. Die Vektoren können vom Ursprung, bis hin zum beliebigen Punkt, gedacht werden. Das Idealpunktmodell beschäftigt sich mit der Frage, ob es nicht auch Produkteigenschaften gibt, die in nur einer bestimmten Menge vom Konsumenten bevorzugt werden. Somit geht das Modell davon aus, dass es eine optimale Kombination sensorischer Attribute für ein bestimmtes Produkt gibt. Daraus ergibt sich

das Ziel, die Idealpunkte des Konsumenten zu bestimmen, die diese optimale Kombination sensorischer Attribute repräsentiert.

Der Biplot des PREFMAP's (Abbildung 7) ist ähnlich Aufgebaut wie der Plot, in der Abbildung 6. Es wird zusätzlich zu den beiliegenden Produktakzeptanzen die jeweiligen Attribute mittels eines Korrelationskreises der HKA eingefügt.

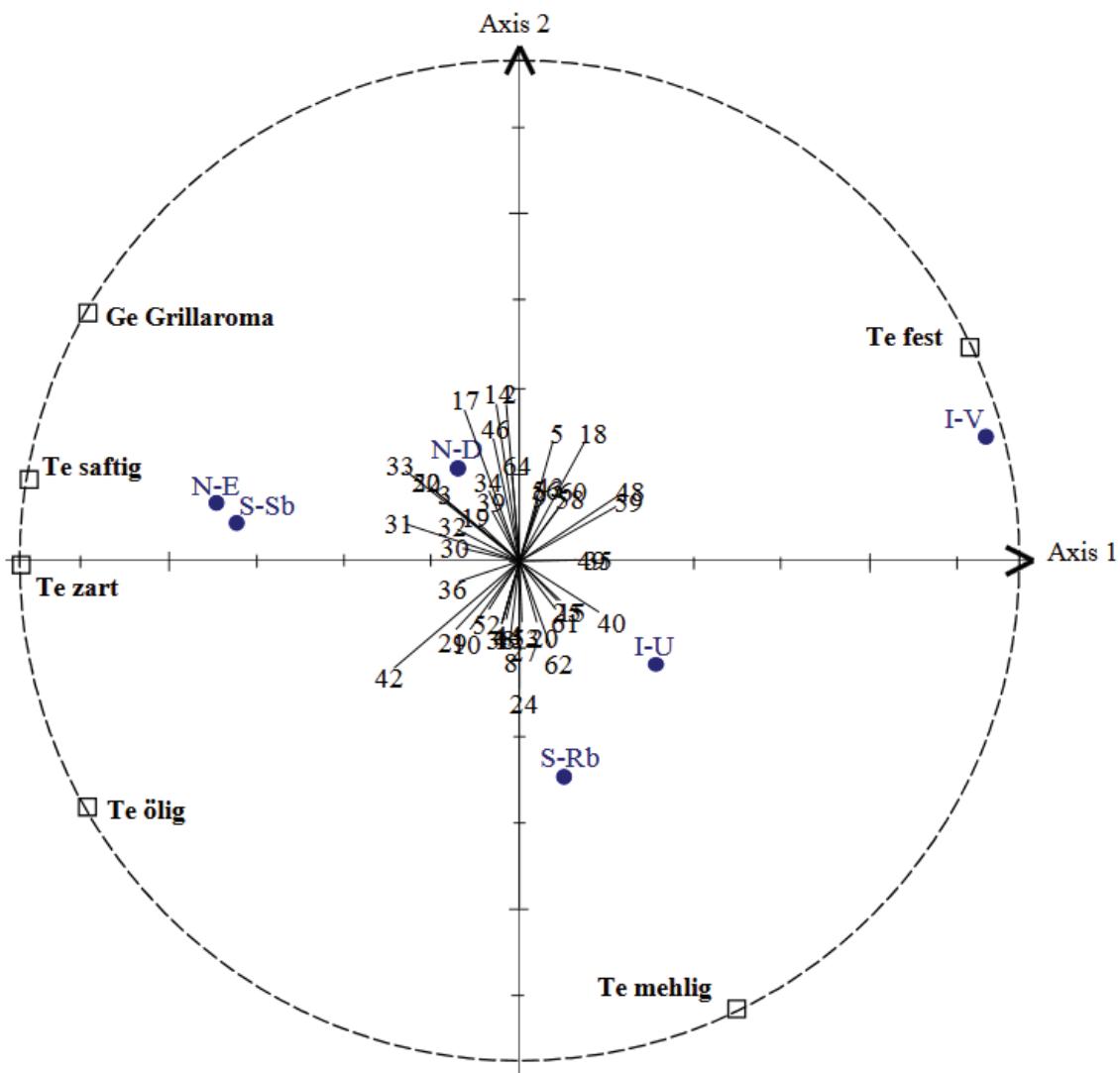


Abbildung 7: Beispiel external preference mapping Biplot (Ellert, 2006)

Die Lage dieser Attribute gibt Aufschluss über die Eigenschaften der Produkte, die in der Nähe der Attribute liegen. Anhand des Beispiels der Abbildung 7, wird die Aussage verdeutlicht. Es ist z.B. zu erkennen, dass das Produkt S-RB in die Richtung mehlig in der Textur tendiert und das Produkt I-V eine besonders feste Struktur aufweist. Ebenfalls wird in dieser Abbildung die Interpretation des Biplots als Landkarte besonders an den Attributen „Textur fest“ und „Textur zart“ ersichtlich, da diese gegenüberliegen (Ellert, 2006). Es kann auch beobachtet werden, dass die

Vektoren der Prüfer, im Gegensatz zu den Vektoren der Attribute oder Produkte, sehr nahe am Ursprung sind. Dies hat die Bedeutung, dass die Verbraucher nicht eindeutig in eine bestimmte Richtung, z.B. die Richtung „mehlig“ gehen. Würde der Verbraucher sagen, dass Produkt ist nur mehlig, dann wäre die Vektorlänge identisch der Attributlänge. Wäre eine Attribut als Vektor nahe am Ursprung, so würde das bedeuten, dass das Panel sich anhand dieses Attributs nicht einig in der Bewertung ist. Dieses könnte an einem fehlenden Verständnis des Attributes liegen oder sogar einer fehlenden Wahrnehmung.

Ein weiterer Schritt ist die Entwicklung eines sogenannten „Countour Plots“. Mit Hilfe dessen, kann dargestellt werden, wie viele verschiedene Konsumentengruppen die unterschiedliche Produktvorlieben aufweisen. Außerdem kann ermittelt werden wie viel Prozent der Konsumenten diese Bereiche bevorzugen.

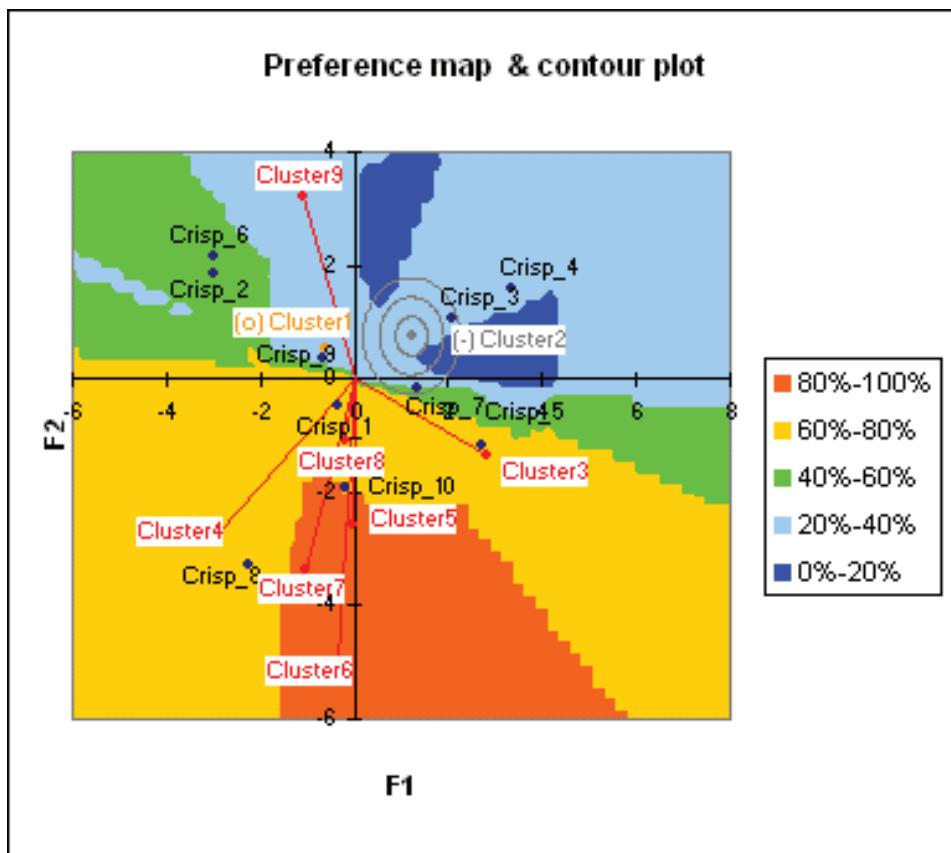


Abbildung 8: Beispiel eines Countour Plots (Xlstat, 2017)

Die Abbildung 8 präsentiert einen Countour Plot anhand des Beispiels der Kartoffelchipsprüfung. Das Beispiel wurde nach der External pref.map Methode und dem quadratischen Model berechnet. Anhand der farblichen Bereiche, werden die Teile der Konsumenten in Prozent dargestellt. Die verschiedenen Vektoren oder auch „Cluster“ genannt, stellen die verschiedenen Kon-

sumentengruppen dar. Es ist zu sehen, dass Probanden des „Cluster 6“, die Chipssorte 8 bevorzugen, jedoch gegensätzliche Ansichten aufweisen wie z.B. die Probanden des „Cluster 9“. Zudem kann behauptet werden, dass 80-100% der Prüfer ähnlich wie „Cluster 7“ beurteilen bzw., das Produkt 8 präferieren (Xlstat, 2017).

3.5 Untersuchungsmaterial

Die zu untersuchenden Proben sind sechs verschiedene Orangensafte (Tabelle 5), die sich in ihrem Fruchtgehalt, aber auch in der Zusammensetzung unterscheiden. Alle Orangensafte sind bekannte Markenprodukte, die im Einzelhandel erhältlich sind.

Tabelle 5: Produkte für die Untersuchung

| Marke | Bezeichnung | Hersteller | MhD | Charge |
|--|--|-------------------------------|----------|------------------------|
| Punica (40% Fruchtgehalt) | Orange plus Vitamin C | PepsiCo Deutschlad GmbH | 16.05.18 | 11:42 D |
| TiP (mind. 50% Fruchtgehalt) | Orangen Nektar aus Orangensaft -konzentrat | Für: real,- Handels GmbH | 01.06.18 | N CVK12 244 02:12 |
| Amecke (100% Fruchtgehalt, davon 47% Orangensaft aus Orangensaftkonz., | Amecke+Zink und Vitamin C | Amecke Fruchtsaft GmbH&Co. KG | 31.03.18 | 00:42 24083 SN28352 |
| Hohes C (100% Fruchtgehalt aus Orangensaftkonz.) | Hohes C Orange 100% Saft | Eckes-Granini | 16.03.18 | F2 03:40 230 |
| Amecke (100% Frucht) | Sanfte Säfte Mandarine-Orange | Amecke Fruchtsaft GmbH&CO KG | 31.07.18 | 10:40 16259 SN29675 |
| Hohes C (100% Fruchtsaft, davon 57% Orangensaft | Hohes C plus Magnesium & B-Vitaminen | Eckes-Granini | 02.06.18 | F2 06:08 248 |

Die Chargennummer und das Mindesthaltbarkeitsdatum verändern sich im Laufe der Untersuchungen nicht. Das Produkt „Amecke + Zink und Vitamin C“ ist ein Saft mit 100% Fruchtgehalt, wobei nur 47% Orangensaft aus Orangensaftkonzentrat vorhanden sind. Zudem besteht der Saft aus 30% Mandarinensaft, Apfelsaft aus Apfelsaftkonzentrat und 5% Grapefruitsaftkonzentrat. Der „Amecke Sanfte Säfte Mandarine-Orange“-Saft weist ebenfalls einen Fruchtgehalt von 100% auf, der aus 60% Orangensaft aus Orangensaftkonzentrat und 40% Mandarinendirektsaft zusammengesetzt ist. Ein Saft mit 100% Fruchtgehalt aus Fruchtsaftkonzentrat ist auch der „Hohes C plus Magnesium & B-Vitaminen“. Davon sind 57% Orangensaft, 28,5% Apfelsaft, 7,5% Traube, 5% Limette, 1,5% Acerola und Bestandteile von Zitronensaft. Somit ist dies kein reiner Orangensaft, sondern ein Mehrfruchtsaft. Lediglich der „Hohes C Orange“ ist ein reiner Orangensaft, der 100% aus der Frucht stammt. Der „TiP-Orangen Nektar“ ist mit seinem 50% Fruchtgehalt ein Nektar.

Zum Neutralisieren der Produkte wird natriumarmes, natürliches Mineralwasser ohne Kohlensäure der Marke „Christinen“, mit der Bezeichnung „Carat Naturelle“ verwendet. Das MhD ist bis 24.08.19 und die Chargennummer mit L3474 17:28 angegeben. Für den Verbrauchertest wird zusätzlich ein Knäckebrot der Marke „Burger“ mit der Bezeichnung „Das Landknäcke: Urtyp herzlich köstlich“ beigelegt. Das MhD ist bis zum 04.18 und die Chargennummer lautet L1609925. In der Profilprüfung wird das Knäckebrot bei Bedarf bzw. auf Wunsch des Probanden ausgegeben.

3.6 Vorbereitung

3.6.1 Probenvorbereitung

Die Vorbereitung der Proben ist für beide Versuchsreihen (Profil- und Konsumentenprüfung) identisch. Jede der sechs verschiedenen Orangensafte wird in einen transparenten, geruchs- und geschmacksneutralen 0,2-Liter-Plastikbecher gefüllt. Dabei beträgt die Füllmenge ca. 50 ml. Des Weiteren sind alle Proben mit einem dreistelligen Zufallscode, der von der Software „FIZZ“ festgelegt wurde, randomisiert.

Das Darreichen der Proben erfolgt nach dem sogenannten „Lateinischen Quadrat“. Somit erhält jeder Proband eine unterschiedliche Reihenfolge der Proben. Für die Profilprüfung erhalten die Proben für jede der drei Sitzungen eine neue Produktcodierung.

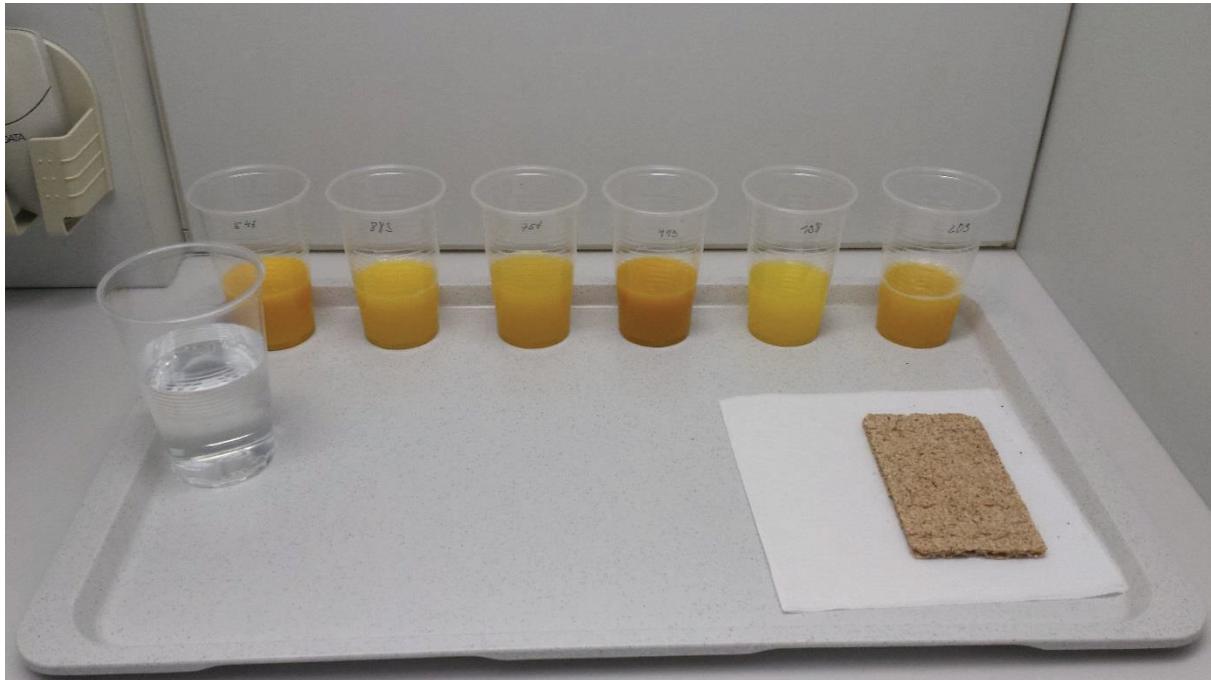


Abbildung 9: Probenpräsentation

In der Abbildung 9 ist die Probenpräsentation für die Profilprüfung und den Verbrauchertest zu sehen. Zusätzlich erhält der Proband zur Neutralisation ca. 0,3 Liter natürliches Mineralwasser, ohne Kohlensäure, in einem transparenten, geruchs- und geschmacksneutralen 0,4-Liter-Plastikbecher. Ebenfalls wird ein Knäckebrot auf einer Serviette zur Verfügung gestellt, wenn das Wasser nicht zum Neutralisieren ausreicht. Bei Bedarf wird mehr Wasser, Knäckebrot oder Probenmaterial ausgehändigt.

3.7 Beschreibung der Probanden

3.7.1 Geschultes Panel

Für die Bearbeitung der Fragestellung wird zum einen eine Profilierung der sechs Orangensafte durchgeführt. Um die Profilprüfung nach DIN 10967 durchführen zu können, wird ein geschultes Panel benötigt. Hierfür werden zwei Panels mit einer Größe von 13 (Gruppe 2) und 14 Prü-

fern (Gruppe 1) gebildet. Die Probanden bestehen aus Studenten der Lebensmitteltechnologie: 3. Semester, der Hochschule Neubrandenburg. Die Probanden wurden in zwei Sitzungen auf das Produkt geschult, indem diese das Produkt untersuchten und Attribute sammelten. Anschließend wurden die Attribute in einer Gruppendiskussion auf Verständnis und Richtigkeit überprüft. Im weiteren Verlauf fand ein Testdurchlauf mit der aktuellen Attributliste statt, mit dem Hintergrund, dass jeder Proband ein Gefühl und Verständnis für die Nutzung der Attribute erhält. Nach der Schulung fand die Profilierung, die innerhalb 3 Sitzungen durchgeführt wurde, in einem Prüfraum nach ISO8589, statt. Hierfür bestand die Aufgabe, die Produkteigenschaften, die in der Gruppe festgelegt wurden, anhand ihrer Intensität, zu beurteilen. Die Intensitäten der einzelnen Attribute wurden mit Hilfe einer Intensitätsskala an einem Computer festgehalten.

3.7.2 Konsumentenpanel

Das Konsumentenpanel besteht im Gegensatz zu der Profilierung, aus nicht produktgeschulten Verbrauchern der Stadt Neubrandenburg. Die Versuchsdurchführung wurde am 27.11.18 vor der Mensa, an der Hochschule Neubrandenburg durchgeführt. Die Rekrutierung erfolgte nach dem Zufallsprinzip, sodass Personen willkürlich angesprochen wurden. Eine Voraussetzung für die Teilnahme am Verbrauchertest wurde damit erfüllt, dass der Proband keine Abneigung gegen Orangensaft hat. Nach der DIN 10974 sollten nicht weniger als 60 Personen an dem Verbrauchertest teilnehmen, um nicht nur eine qualitative, sondern auch quantitative Aussage treffen zu können. Zudem wird damit die Messgenauigkeit erhöht. Es wird empfohlen, eine Panelgröße von 80-120 Personen zu haben. An dem Verbrauchertest haben 98 Personen teilgenommen. Die Aufgabe der Prüfer bestand darin, die persönliche Präferenz bzw. Beliebtheit und die Intensität der Attribute „Farbe im Aussehen“, „orangenartig im Geschmack“ und „bitter im Geschmack“ anzugeben. Des Weiteren sollten die Produkte anhand vorgegebener Attribute überprüft werden (siehe Anhang 1.5).

4. Ergebnisse

4.1 Verbrauchertest

4.1.1 Angabe zur Person

Bei dem Verbrauchertest vor der Mensa der Hochschule Neubrandenburg haben insgesamt 98 Personen teilgenommen. Von den 98 Teilnehmern waren 58 weibliche und 40 männliche Probanden zu verzeichnen. Die Altersspanne lag mit 70 Probanden im Alter zwischen 18 und 25 Jahren. Wie die Abbildung 10 verdeutlicht, haben 24 Personen, die zwischen 26 und 35 Jahren, eine Person zwischen 36 und 45 Jahren und 3 Personen über 45 Jahren alt, an dem Verbrauchertest teilgenommen.

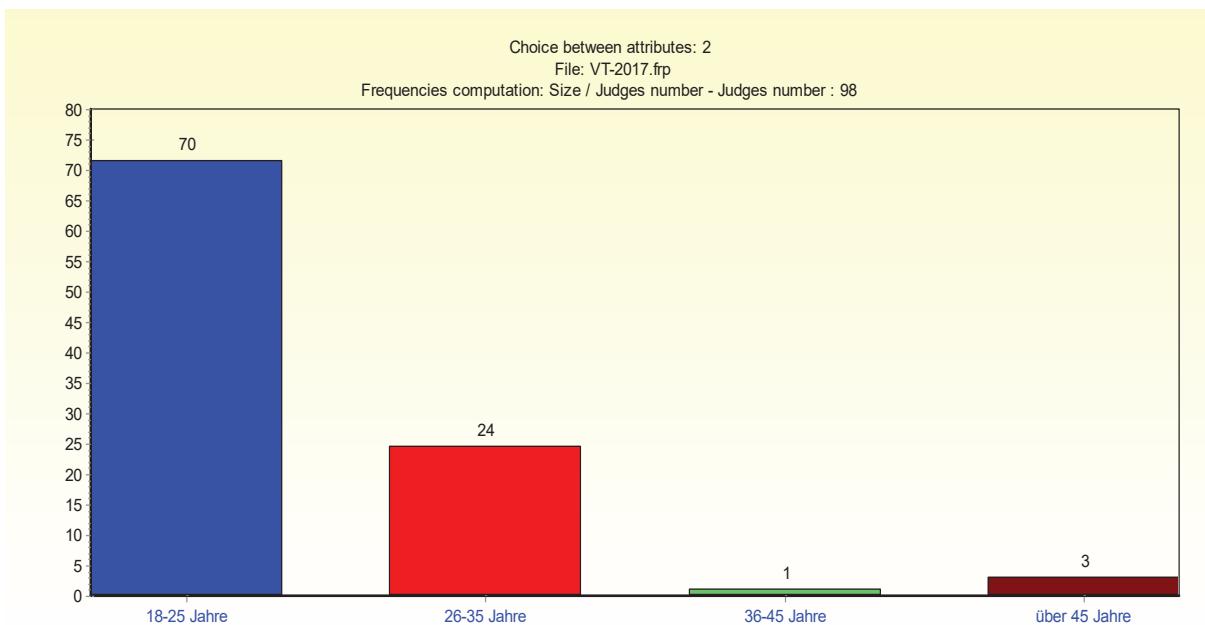


Abbildung 10: Altersaufteilung der Probanden

Unter anderem wurde in dem Verbrauchertest auch der Konsumabstand von Orangensaften abgefragt. Hierbei wurde in der Abbildung 11 ersichtlich, dass der größte Teil der Probanden, 36 Personen, den Orangensaft seltener als 2-3x im Monat konsumieren. Bevor ein Proband den Verbrauchertest durchführen konnte wurde dieser befragt, ob eine Allergie oder Abneigung gegenüber Orangensaften besteht.

In diesem Fall haben die 36 Personen keine Abneigung, jedoch werden oftmals andere Produkte (andere Säfte, Cola, etc.) gekauft.

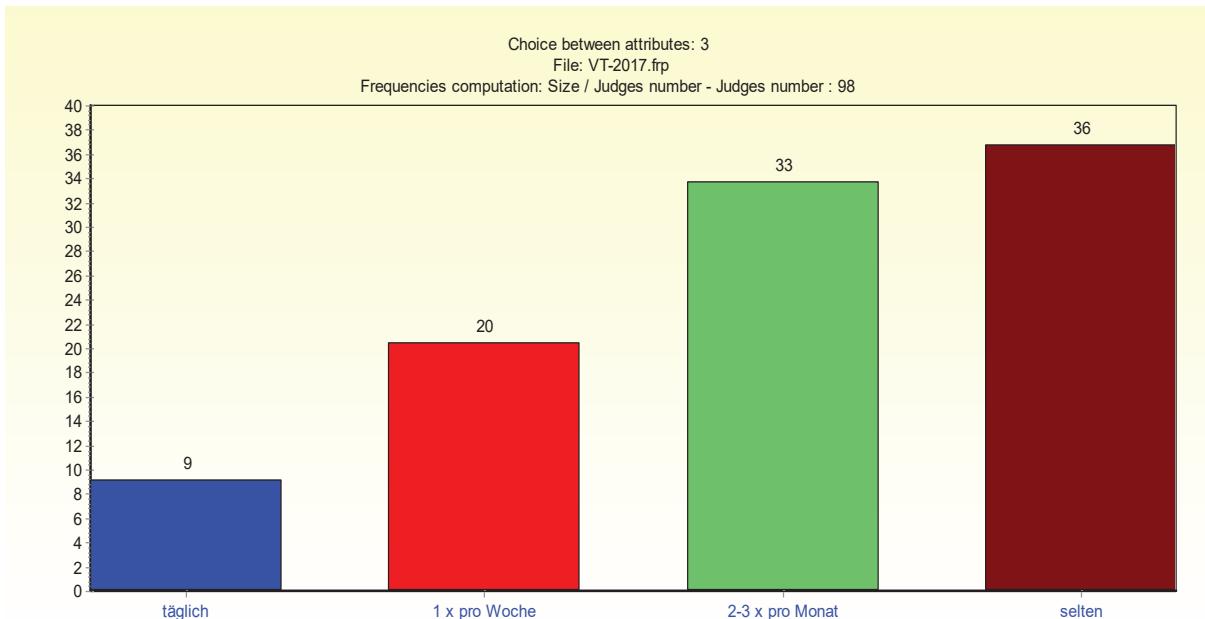


Abbildung 11: Häufigkeit des Orangensaftverzehrs

Die zweitgrößte Gruppe mit 33 Personen, konsumieren den Saft mehrmals die Woche, 20 von 98 sogar Wöchentlich und 9 Personen täglich.

4.1.2 Beliebtheitsprüfung

Die Aufgabe der Probanden in der Beliebtheitsprüfung bestand, die Produkte nach persönlicher Vorliebe eine Note zu vergeben. Die Noten wurden auf einer 9-Punkt hedonischen Skala notiert (siehe Kapitel 3.2.1).

Tabelle 6: Durchschnittsbewertung mit Standardabweichung

| Level | Name | Size | Mean | Std. Dev. |
|-------|-----------------|------|------|-----------|
| 1 | PUNICA | 98 | 3,86 | 2,07 |
| 2 | TIP-O-NEKTAR | 98 | 4,91 | 2,26 |
| 3 | AM-ZINK-VIT C | 98 | 5,16 | 2,16 |
| 4 | HOHES C- ORANGE | 98 | 6,08 | 2,04 |
| 5 | AM-MANDAR-ORANG | 98 | 5,9 | 2,08 |
| 6 | HOHES C- O-T-L | 98 | 4,62 | 2,22 |

In der Tabelle 6 sind die getesteten Fruchtsäfte und deren Durchschnittsbewertung mit der Standardabweichung angegeben. Der reine Fruchtsaft „Hohes C-Orange“ ist mit 6,08 Punkten im Durchschnitt am besten bewertet worden. Zusätzlich hat diese Probe mit 2,04 Punkten die geringste Standardabweichung. Somit liegt der Saft im Bereich „gefällt etwas“. Am schlechtesten wurde der Saft der Marke „Punica“ bewertet. Dieser erhielt im Durchschnitt eine Benotung von 3,86 Punkten. Hier kann die Schlussfolgerung getroffen werden, dass der Saft im „Missfallen“-Bereich liegt. Die anderen Säfte liegen im neutralen Bereich bis hin zum „gefällt etwas“.

Tabelle 7: ANOVA Beliebtheitsprüfung

| Sources of variation | D.F. | S.S. | M.S. | Comp. F | Proba. |
|----------------------|------|---------|------|---------|----------------|
| Product | 5 | 334,48 | 66,9 | 15,98 | <0,0001 *** |
| Judge | 97 | 636,07 | 6,56 | 1,57 | 0,0012 ** |
| Residuals | 485 | 2030,85 | 4,19 | | |
| Total | 587 | 3001,4 | | | |

* significant at 5 % ** significant at 1 % *** significant at 0,1 %

Die Anova-Tabelle gibt hier die Auskunft, dass die Produkte höchst signifikant sind. Dies bedeutet, dass die Probanden die Produkte als unterschiedlich wahrgenommen haben. Eine hohe Signifikanz bei Prüfern hat zur Folge, dass die Prüfer mit einer Wahrscheinlichkeit von 1% unterschiedlich bewertet haben bzw. verschiedene Geschmäcker und Vorlieben besitzen. Mit dem Post hoc Test nach Fischer LSD bei einem L.S.D. (least significant difference) von 5% wird dieses noch einmal verdeutlicht

Tabelle 8: Post hoc Test, Faktor „Produkt“ mit L.S.D.: 0,57

| Level | Size | Mean | Groups |
|-----------------|------|------|--------|
| HOHES C- ORANGE | 98 | 6,08 | A |
| AM-MANDAR-ORANG | 98 | 5,9 | A |
| AM-ZINK-VIT C | 98 | 5,16 | B |
| TIP-O-NEKTAR | 98 | 4,91 | B |
| HOHES C- O-T-L | 98 | 4,62 | B |
| PUNICA | 98 | 3,86 | C |

In der Tabelle 8 sind die Orangensaft-Säfte mit ihren Durchschnittsbewertungen und der dadurch entstandenen Gruppeneinteilungen zu sehen. Die Spalte mit den Gruppen gibt Auskunft, welche Säfte ziemlich ähnlich und welche unterschiedlich in der Beliebtheitsprüfung abgeschnitten haben. Der L.S.D. gibt die Grenzen der Gruppen an, ab wann das Produkt nicht mehr „A“, sondern

„B“ ist. Auch hier wird deutlich, dass „Hohes C-Orange“ (Gruppe A) deutlich anders bewertet wurde, als das Produkt „Punica“ (Gruppe C).

Bei dem Verbrauchertest macht es keinen Sinn, sich einen Post hoc Test mit dem Faktor „Prüfer“ genauer unter Betracht zu ziehen, da die Beliebtheitsprüfung rein subjektiv ist und aus diesem Grund verschiedene Bewertungen mit ca. 25 oder mehr Gruppeneinteilungen vorhanden sind.

Die Beliebtheit der Produkte ist in dem Verbrauchertest höchst Signifikant unterschiedlich.

4.1.3 JAR-Fragen

Die JAR-Fragen bzw. die Attribute: Farbintensität, orangenartig GS und bitter GS, werden im ersten Schritt ebenfalls mit Hilfe der Varianzanalyse ausgewertet. Für ein leichteres Verständnis wird die ANOVA-Tabelle zusammengefasst.

Tabelle 9: ANOVA der 3 JAR-Attribute

| Analysis of variance | | Farbintensität | orangenartig GS | bitter GS |
|-----------------------------|--|-----------------------|------------------------|------------------|
| Sources of variation | | Proba. | Proba. | Proba. |
| Product | | <0,0001 *** | <0,0001 *** | <0,0001 *** |
| Judge | | <0,0001 *** | 0,0003 *** | <0,0001 *** |

* significant at 5 % ** significant at 1 % *** significant at 0,1 %

In der Tabelle 9 werden die Signifikanzen der drei JAR-Attribute, bezogen auf die Faktoren Produkt und Prüfer, dargestellt. Alle drei Attribute zeigen eine Höchstsingifikanz in Produkt und Prüfer an. Das bedeutet, dass die Produkte anhand dieser Attribute unterscheidbar sind und dass die Prüfer die Produkte anhand der Attribute unterschiedlich bewertet haben. Somit gibt es keine einheitliche Bewertung, bzw. die Probanden sind sich bei den JAR-Fragen nicht einig. Durch den Einblick in die Rohdaten können Beispiele hervorgebracht werden. So beurteilt Proband 1 den „Punica“-Saft in der Kategorie „Farbintensität“ mit „viel zu intensiv“ und der Proband 30 beurteilt das gleiche Produkt und in der gleichen Kategorie mit „viel zu schwach“. Gründe können verschiedene Wahrnehmungen, Missverständnis der Skala, aber auch mögliche Ablenkungsfaktoren, wie Lärm, Gerüche etc., die eine Konzentrationsstörung zur Folge haben, sein.

In der Tabelle 10,11 und 12 wird die Gruppeneinteilung der Säfte nach den Attributen dargestellt.

Tabelle 10: Post hoc-Test Farbintensität mit LSD: 0,24

| Level | Size | Mean | Groups Farbin-tensität LSD: 0,24 |
|-----------------|-------------|-------------|---|
| AM-MANDAR-ORANG | 98 | 3,67 | A |
| AM-ZINK-VIT C | 98 | 3,61 | A |
| HOHES C- ORANGE | 98 | 2,88 | B |
| HOHES C- O-T-L | 98 | 2,81 | B |
| TIP-O-NEKTAR | 98 | 2,21 | C |
| PUNICA | 98 | 1,97 | D |

Tabelle 11: Post hoc-Test orangenartiger GS mit LSD: 0,27

| Level | Size | Mean | Groups oran-genartig GS LSD: 0,27 |
|-----------------|-------------|-------------|--|
| AM-MANDAR-ORANG | 98 | 2,72 | A |
| HOHES C- ORANGE | 98 | 2,67 | A |
| AM-ZINK-VIT C | 98 | 2,47 | AB |
| HOHES C- O-T-L | 98 | 2,4 | B |
| TIP-O-NEKTAR | 98 | 1,97 | C |
| PUNICA | 98 | 1,88 | C |

Tabelle 12: Post hoc-Test bitter GS mit LSD: 0,28

| Level | Size | Mean | Groups bitter GS LSD: 0,28 |
|-----------------|-------------|-------------|-----------------------------------|
| HOHES C- O-T-L | 98 | 3,37 | A |
| AM-ZINK-VIT C | 98 | 3,35 | AB |
| HOHES C- ORANGE | 98 | 3,07 | BC |
| AM-MANDAR-ORANG | 98 | 2,91 | C |
| PUNICA | 98 | 2,61 | D |
| TIP-O-NEKTAR | 98 | 2,33 | E |

Die Produkte können mit einem LSD von 0,27 in vier Gruppen eingeteilt werden, mit einem LSD von 0,27 können die Produkte im orangenartigen Geschmack in drei Gruppen unterteilt werden. Durch das Attribut bitter im Geschmack entstehen sogar fünf Gruppen. Die „Amecke“-

und „Hohes C“- Produkte sind in der Farbintensität ähnlich, bzw. in der gleichen Gruppe. Es kann beobachtet werden, dass bei dem Attribut „bitter im Geschmack“ (Tab. 12) die meisten Gruppen vorhanden sind. Dies hat zur Folge, dass die Produkte alle unterschiedlich in der bitteren Note wahrgenommen wurden. Ebenfalls ist eine Überschneidung der Gruppen in der Tabelle 11 und 12 vorhanden. Das Produkt „Amecke Zink plus Vitamin C“ ist in beiden Kategorien (orangenartig GS und bitter GS) in der Gruppe A und der Gruppe B. Dies bedeutet, dass das Produkt in diesen Kategorien sowohl den Produkten in der Gruppe A, aber auch in der Gruppe B ähnlich ist.

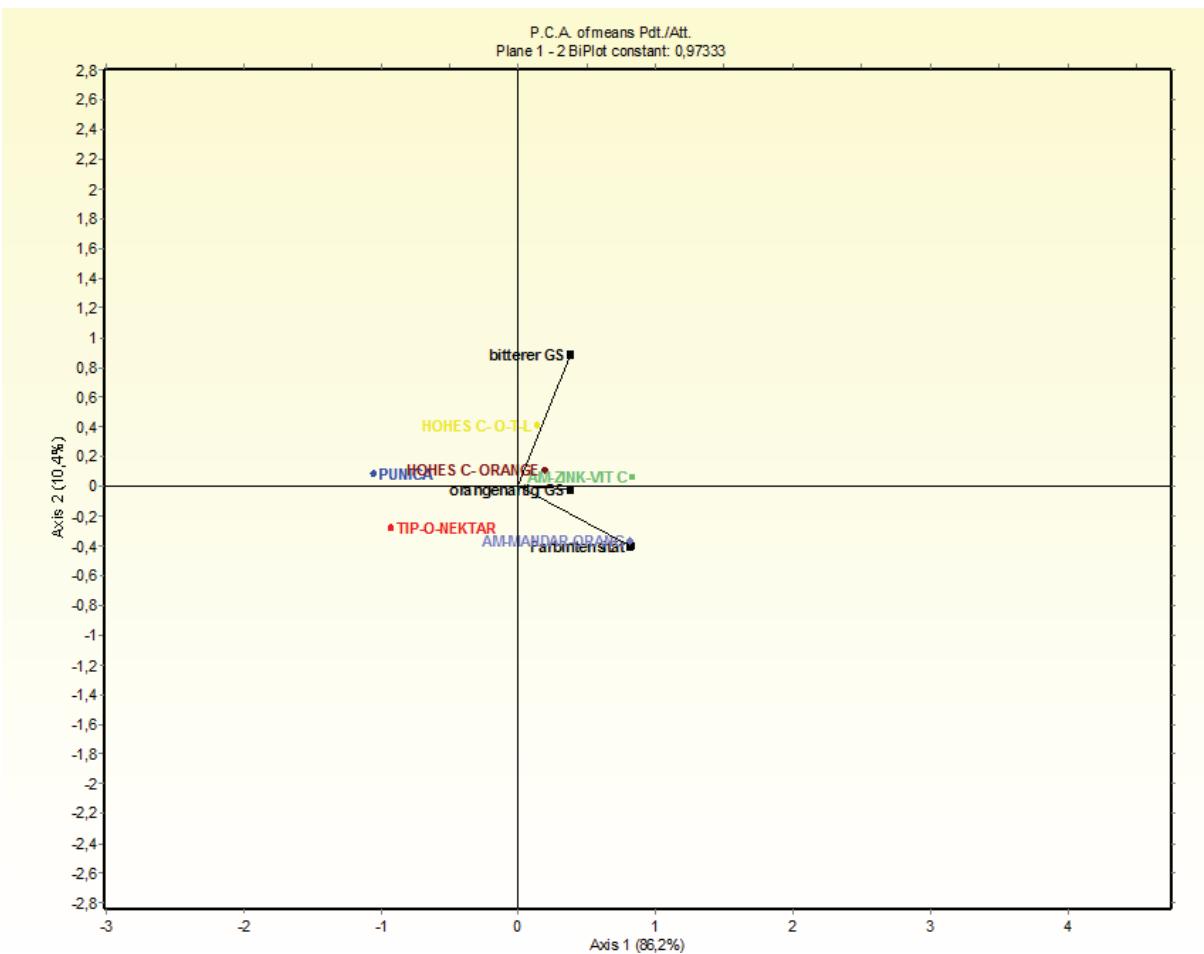


Abbildung 12: PCA-Plot der JAR-Attribute (Achse 1 + Achse 2 = 96,6%)

In der Abbildung 12 wird eine HKA anhand eines PCA-Biplots dargestellt. Die erste Hauptkomponente (X-Achse) wird mit 86,2% und die zweite HK (Y-Achse) wird mit 10,4% dargestellt. Im ersten Moment wird ersichtlich, dass die Produkte nah beieinanderliegen, sodass die Behauptung getroffen werden kann, dass diese Produkte ähnliche Bewertungen erhalten haben, bzw. die Produkte konnten untereinander nur schwer unterschieden werden. Das Produkt „Hohes-C O-T-L“

ist jedoch weiter entfernt, als der Saft von „Amecke Mandarine“ und der Nektar von „Tip“, so dass zwischen „Hohes-C O-T-L“ und den anderen zwei Produkten größere Unterschiede getroffen wurden. Werden die Attribute mit einbezogen, so ist ersichtlich, dass das „Amecke Mandarine“ eine besonders hohe Farbintensität und das „Hohes-C O-T-L“ eine besonders bittere Note aufweist. „Hohes C- Orange“, aber auch „Amecke Zink plus Vitamin C“ besitzen einen besonders hohen orangenartigen Geschmack.

4.1.4 Penalty-Analyse

Um eine Auskunft zu erhalten, wie sehr die Produkte an Beliebtheit verlieren, wenn diese nicht JAR sind und wenn die Erwartung (JAR) nicht erfüllt wurde, werden die sogenannte „Mean Drop“ und die „Penalties“ berechnet. Die Berechnung erfolgt nach der Beschreibung im Kapitel 3.2.3.2.

4.1.4.1 Punica

Im ersten Schritt werden von jeder Probe die fünf JAR-Kategorien zu drei Kategorien (schwach, JAR, intensiv) zusammengefasst.

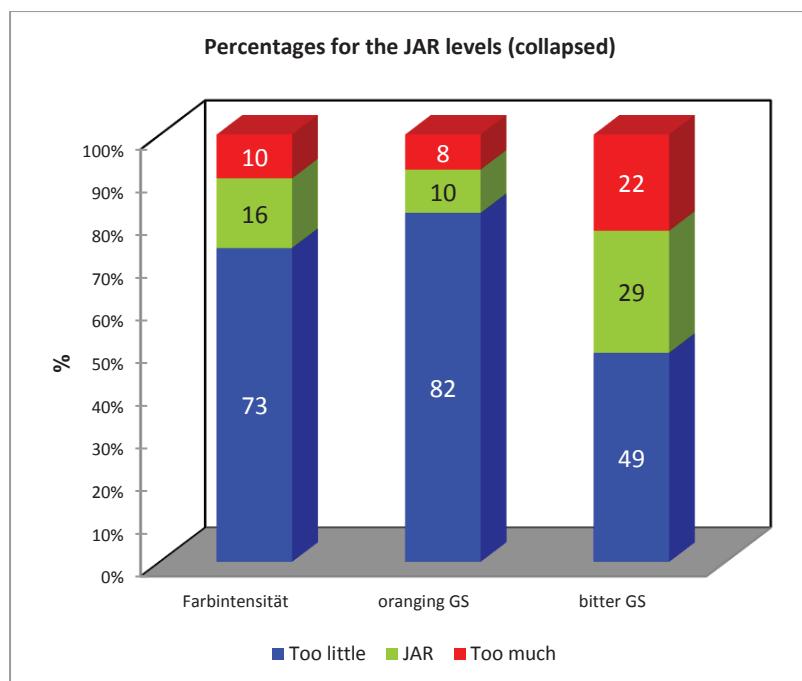


Abbildung 13: JAR-Bewertungen Punica

In der Abbildung 13 sind die Intensitätsbewertungen des Produktes „Punica“, die anhand der Verbraucher durchgeführt wurden, zu sehen. Es ist ersichtlich, dass die Probanden den Saft in allen drei möglichen Attributen als zu schwach empfanden. Der orangige Geschmack wurde sogar von 82% der Konsumenten als nicht intensiv genug gewählt. Fast die Hälfte der Teilnehmer empfand die bittere Note zu gering und nur 29% empfanden diese als genau richtig.

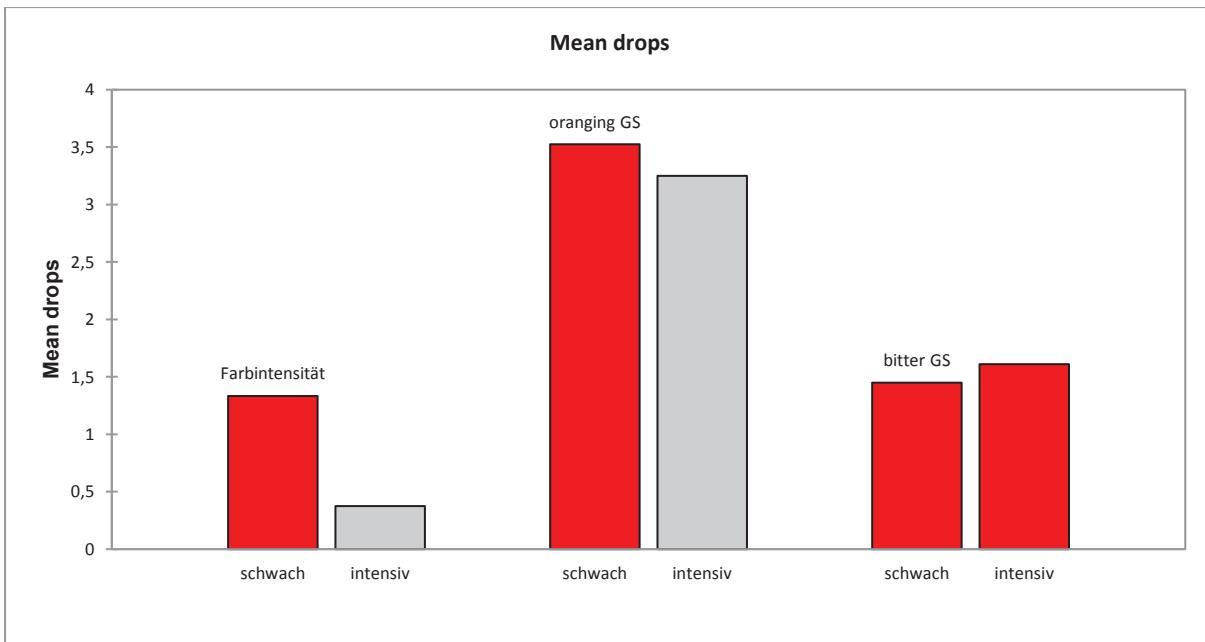


Abbildung 14: Mean Drops des Punica Saftes

Anhand der Mean Drops in der Abbildung 14 kann die Aussage getroffen werden, dass der Saft mit einer Signifikanz von 5% die meisten Punkte (3,5) in der Beliebtheit verliert, wenn der orangige Geschmack zu schwach ist. Für die JAR-Kategorie orangiger Geschmack „zu intensiv“ und auch Farbintensität „zu intensiv (0.375)“ konnte keine Signifikanz ermittelt werden, da der %-Satz der Fälle unter der 20%-Schwelle liegt. In der Farbintensität „zu schwach“ liegt ein Mean Drop von 1,33 vor und ist ebenfalls signifikant. Der bittere Geschmack hat ähnlich hohe Mean Drops, wobei der „zu intensiv“ höher ist. Hieraus kann interpretiert werden, dass der Saft in Richtung mehr orangiger Geschmack und einer höheren Farbintensität überarbeitet werden könnte.

Die Strafpunkte in der Abbildung 15 zeigen an, wie viele Punkte an Beliebtheit verloren gehen, wenn die Erwartung (JAR) des Konsumenten, bzw. die „genau richtig“-Kategorie, nicht erfüllt wurde.

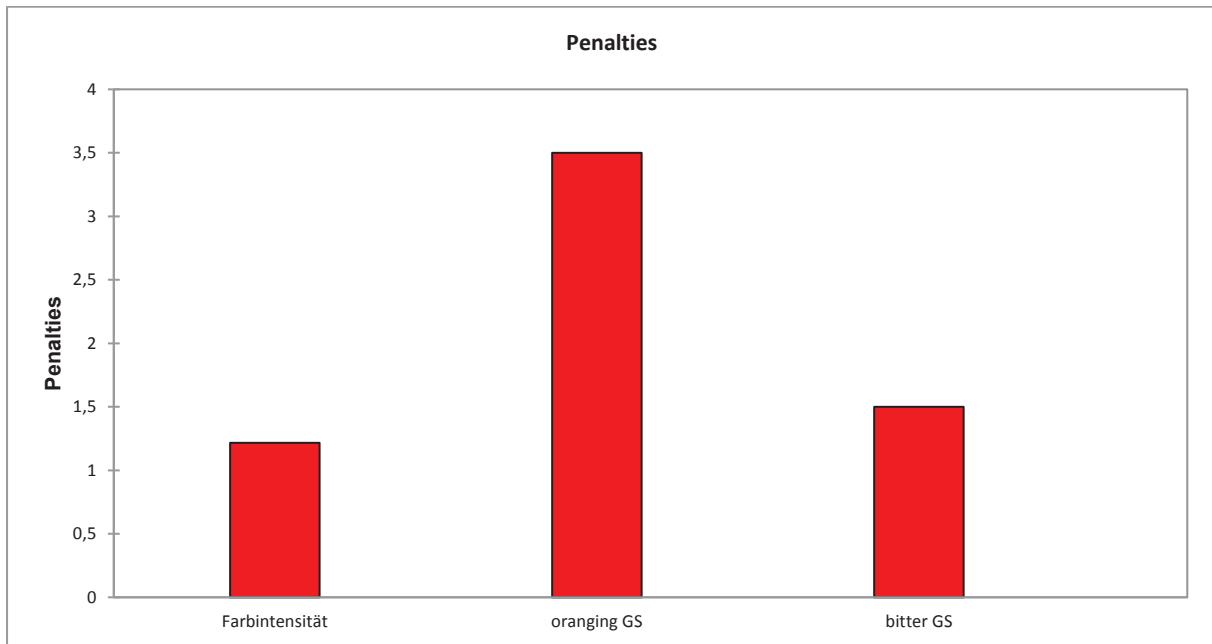


Abbildung 15: Penalties des Punica Saftes

Auch hier ist ein ähnliches Ergebnis zu betrachten, sodass der orangige Geschmack bei „Punica“ dem Kunden sehr wichtig ist und sich hiermit die höchsten Strafpunkte ergeben. Alle drei Attribute sind signifikant, wobei in der Farbintensität die wenigsten Strafpunkte vorhanden sind.

4.1.4.2 Tip Nektar

Der „TIP Orangennektar“ zeigt eine ähnliche Aufteilung der JAR-Kategorien in der Abbildung 16 wie der „Punica“-Saft. Auch hier überwiegt in allen drei Attributen die Kategorie „zu schwach“. Des Weiteren sind die Bewertungen in der Kategorie „zu intensiv“ in allen drei Produkteigenschaften sehr gering.

Der bittere Geschmack ist mit 43% in der „genau richtig“-Kategorie am höchsten vertreten.

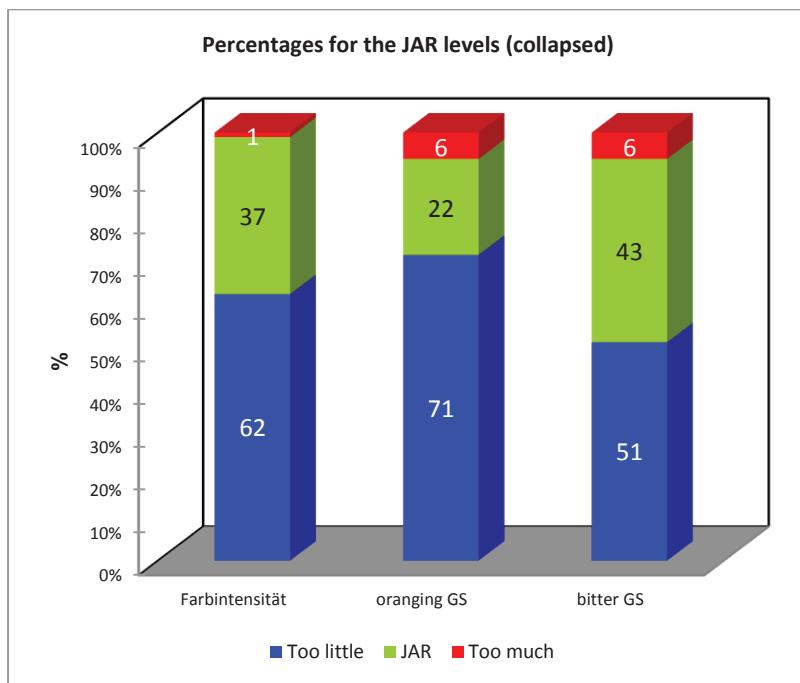


Abbildung 16: JAR-Bewertung TIP Nektar

Der Mean Drop in der Farbintensität „intensiv“ (Abbildung 17) ist am höchsten zu betrachten. Jedoch kann diese nicht mit der Signifikanz untermauert werden. In allen drei Produktmerkmalen liegt nur in den Bereichen „zu schwach“ eine Signifikanz vor. Der orangige Geschmack „zu schwach“ ist ähnlich wie bei dem „Punica“-Saft mit einem hohen Mean Drop belegt.

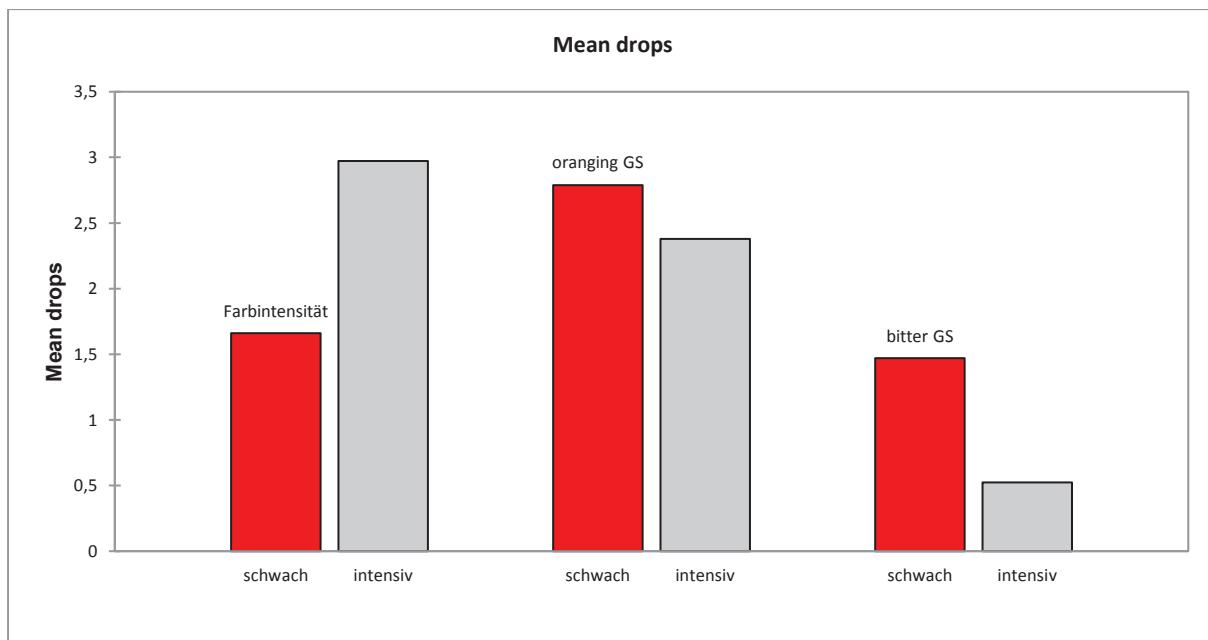


Abbildung 17: Mean Drops des TIP-Nektars

In der Abbildung 18 können die höchsten Strafpunkte im orangigen Geschmack festgemacht werden. Somit werden bei dem Nichterfüllen der „genau richtig“ Kategorie 2,75 Punkte in der Beliebtheit abgezogen. Also ist dem Verbraucher ein orangiger Geschmack auch bei diesem Produkt sehr wichtig und der Verbraucher wünscht sich diesen intensiver.

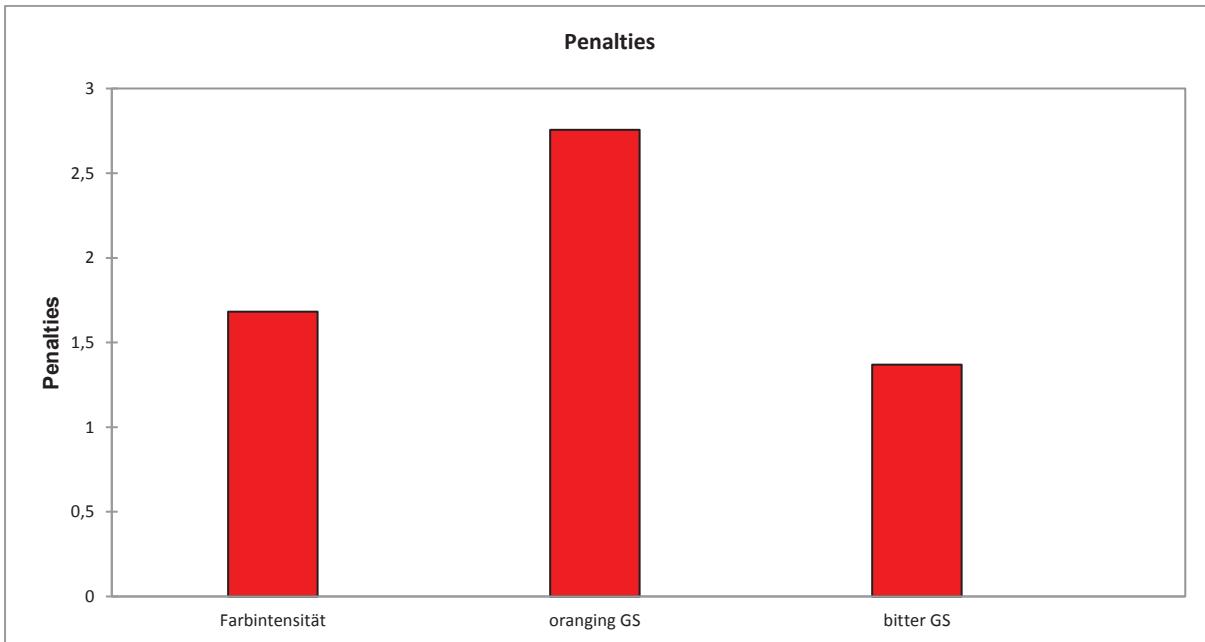


Abbildung 18: Penalties des TIP-Nektars

Auch im bitteren Geschmack und der Farbintensität kann minimal in die Richtung „mehr“ entwickelt werden.

4.1.4.3 Amecke plus Zink und Vitamin C

Die Bewertungen der JAR-Kategorie (Abbildung 19) bei „Amecke plus Zink und Vitamin C“ sind durchwachsen. So ist die Farbintensität mit über 50% als zu viel und mit 39% als genau richtig bewertet worden. Der orangige Geschmack wurde von 55% der Verbraucher als zu wenig eingestuft, wobei auch 32% der Meinung sind, dass dieser genau richtig ist.

Der bittere Geschmack ist nicht eindeutig erklärt. 42% der Verbraucher empfanden ihn als zu intensiv, 39% als genau richtig und nur 19% als zu schwach.

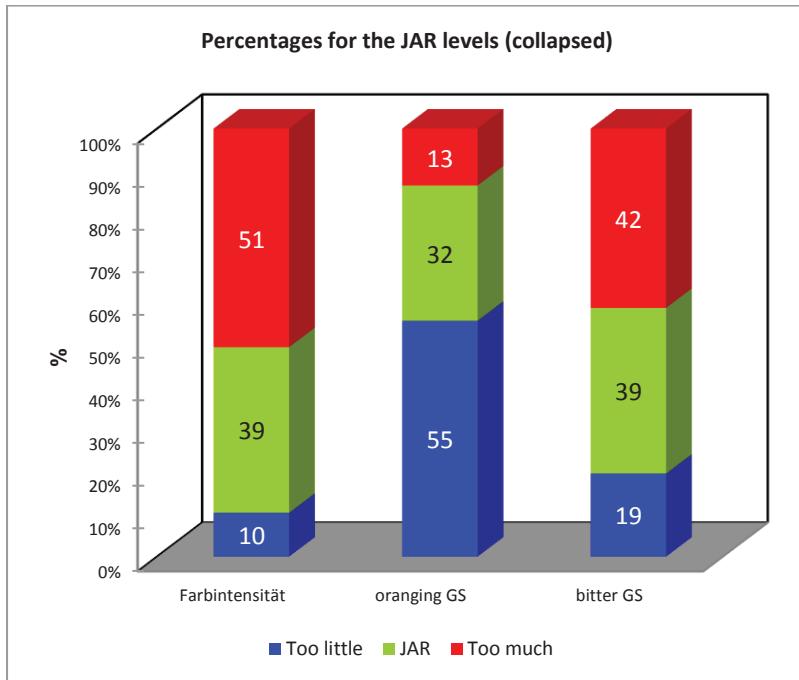


Abbildung 19: JAR-Bewertung Ameckle plus Zink und Vitamin C

Bei diesem Fruchtsaft kann der höchste Mean Drop (Abbildung 20) von ~ 2 in der Kategorie orangiger Geschmack „zu schwach“ verzeichnet werden. Dicht gefolgt von dem zu intensiven Geschmack, der aber nicht mit einer Signifikanz belegt werden kann. In der Farbintensität ist minimal der Bereich des zu intensiven größer, jedoch liegt hier bei einer 5%-Wahrscheinlichkeit keine Signifikanz vor

Der bittere Geschmack ist in der Kategorie „zu intensiv“ mit einer Signifikanz und einem Mean Drop von 1,55 abgebildet. Der Bereich „zu schwach“ konnte nicht ermittelt werden, da nur 19% der Verbraucher für diese Kategorie gestimmt haben, jedoch 20% die zu erreichende Schwelle ist. Somit erhält das Produkt die höchsten Abzüge in der Beliebtheit, wenn der orangige Geschmack zu schwach ist. Aber auch Abzüge, wenn die bittere Note dominiert.

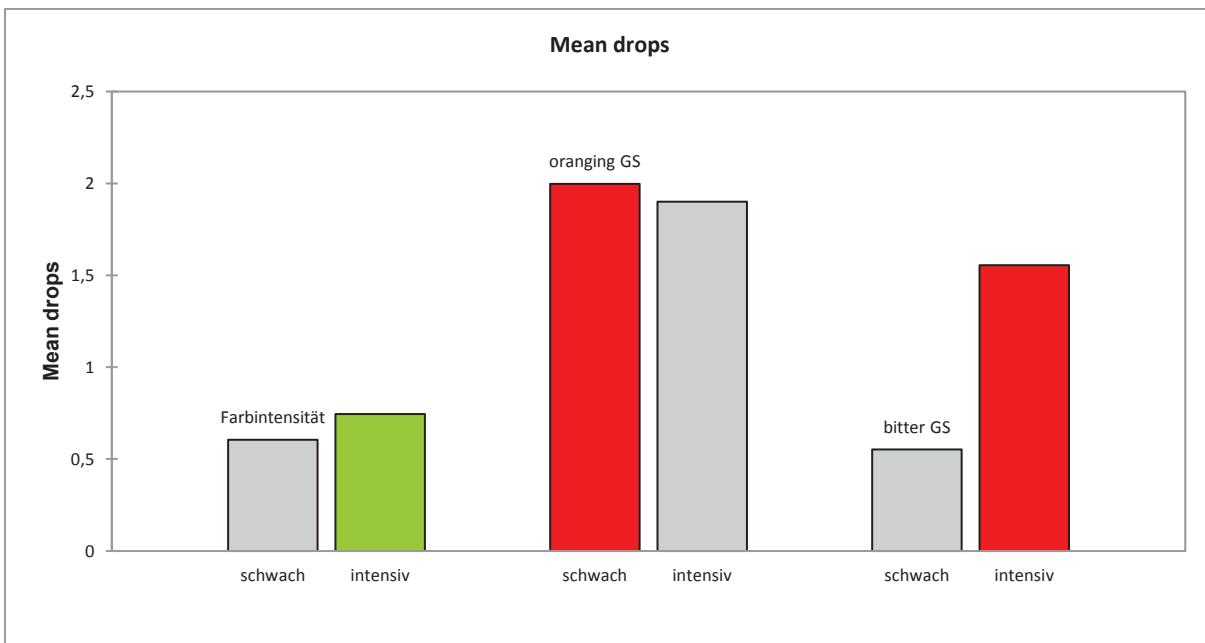


Abbildung 20: Mean Drops des Amecke plus Zink und Vitamin C Saftes

Die Penalties in der Abbildung 21 untermauern die Aussage zusätzlich. Mit einem Penalty von 1,98 im orangigen Geschmack, verzeichnet der Saft seine höchsten Strafpunkte bei der Nichterfüllung des JAR-Bereiches. Der bittere Geschmack wird mit 1,24 Strafpunkten verzeichnet. Diese beiden Kategorien verfügen über eine Signifikanz. Das Attribut Farbintensität kann bei der Wahrscheinlichkeit von 5% keine Signifikanz vorzeigen.

Um den Saft zu optimieren, müsste der Saft in die Richtung mehr Geschmack und weniger bitter überarbeitet werden.

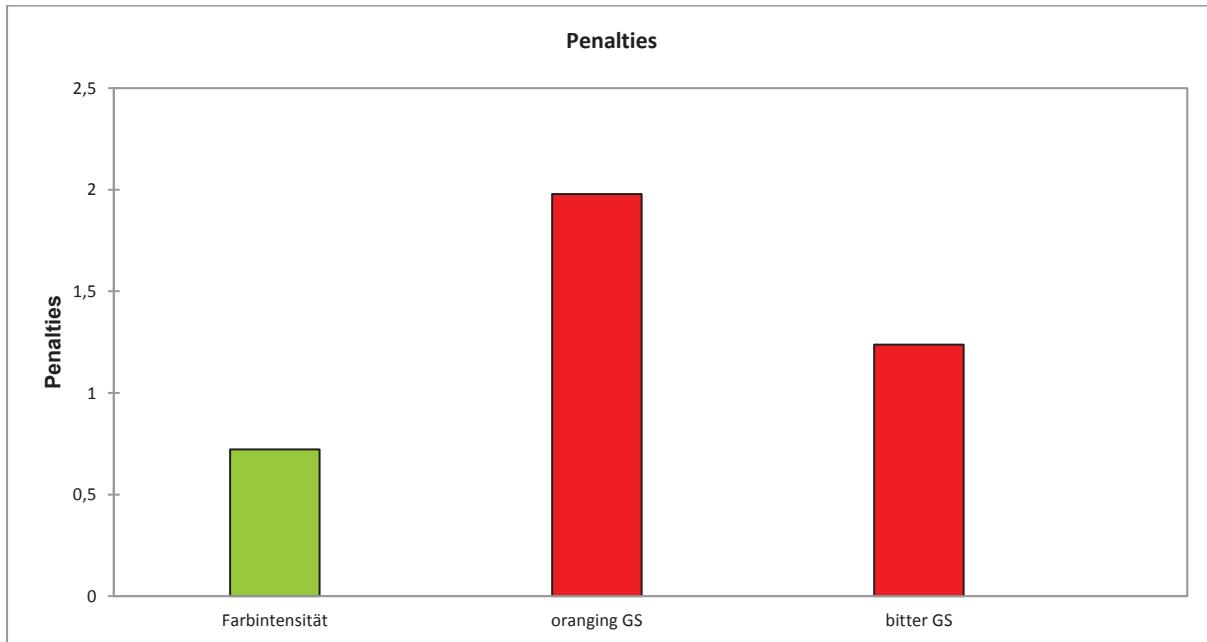


Abbildung 21: Penalties des Amecke plus Zink und Vitamin C Saftes

In der Farbintensität kann keine klare Aussage geschaffen werden, ob und in welche Richtung der Saft verändert werden sollte, da keine Signifikanz im Bereich „zu schwach“ berechnet werden konnte und im Bereich „zu intensiv“ das Signifikanzniveau von 5% nicht ausreichend ist.

4.1.4.4 Hohes C 100% Orangensaft

Der Saft „Hohes C 100% Orangensaft“ besitzt die meisten Nennungen in der JAR-Kategorie (Abbildung 22). Vor allem die Farbintensität wurde mit 66% am meisten als genau richtig empfunden. Auch der orangige Geschmack wurde von fast der Hälfte der Konsumenten als genau richtig empfunden, jedoch auch von 39% als zu schwach.

Der bittere Geschmack ist mit 51% als genau richtig empfunden worden. 29% fanden den Saft zu intensiv und 20% zu schwach bitter.

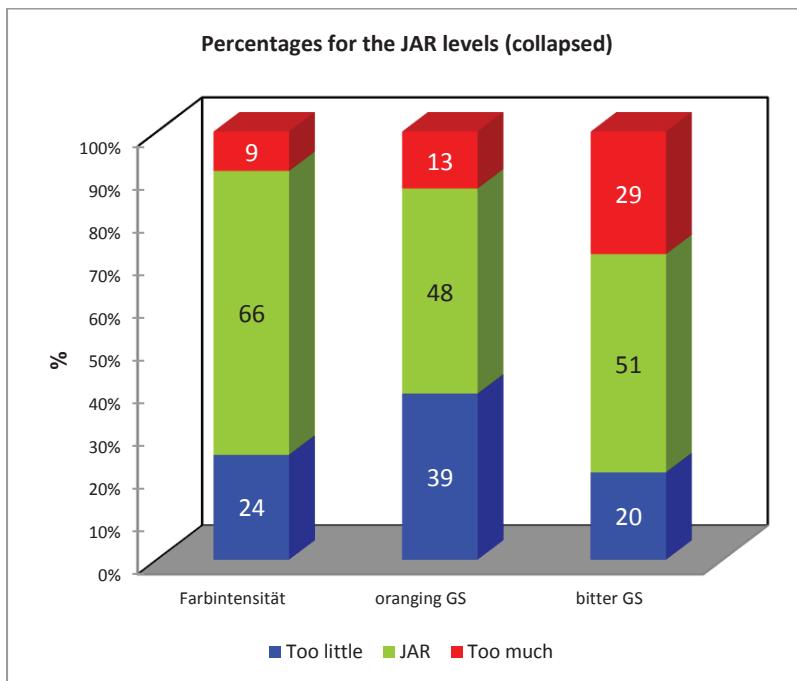


Abbildung 22: JAR-Bewertung Hohes C 100% Orangensaft

Der Fruchtsaft besitzt ähnlich hohe Mean Drops (Abbildung 23), wobei der orangige Geschmack mit 2,18 Punkten in der Kategorie „zu schwach“ mit einer Signifikanz den höchsten Mean Drop verzeichnet. Die Mean Drops des Attributes bitter, zeigen in beiden Kategorien eine Signifikanz in einem ähnlich hohen Bereich („zu schwach“: 1,65 und „zu intensiv“: 1,69) auf. Somit kann die Aussage getroffen werden, dass der Verbraucher gleichviele Punktabzüge in der Beliebtheit geben würde, wenn die bittere Note zu schwach oder zu stark ist. Die Farbintensität im Bereich „zu schwach“ überliegt eindeutig dem Bereich „zu intensiv“ und ist signifikant.

Somit würde der Verbraucher die meisten Abzüge in der Beliebtheit bei einem zu schwachen Geschmack geben.

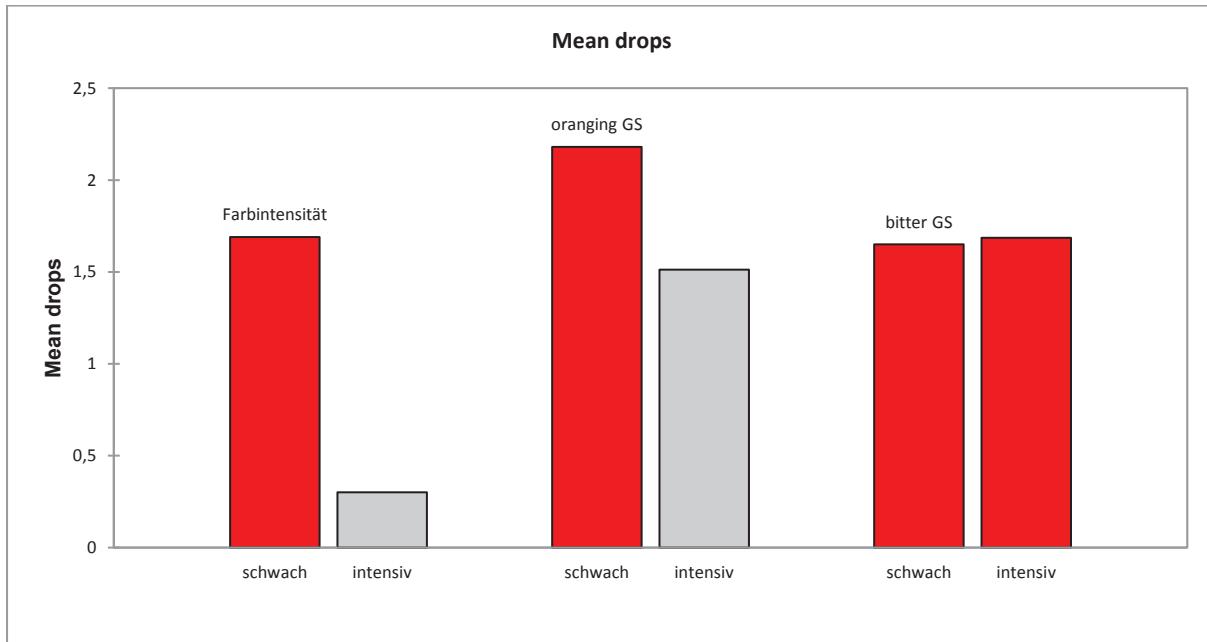


Abbildung 23: Mean Drops des Hohen C 100% Orangensaftes

Die Abbildung 24 belegt die Aussage, dass die meisten Strafpunkte (2,0) bei dem orangigen Geschmack vergeben werden, wenn dieser die Erwartungen nicht erfüllt. Der bittere Geschmack ist mit 1,67 Strafpunkten im Nicht-JAR-Bereich ebenfalls hoch, dicht gefolgt von der Farbintensität mit 1,31 Strafpunkten.

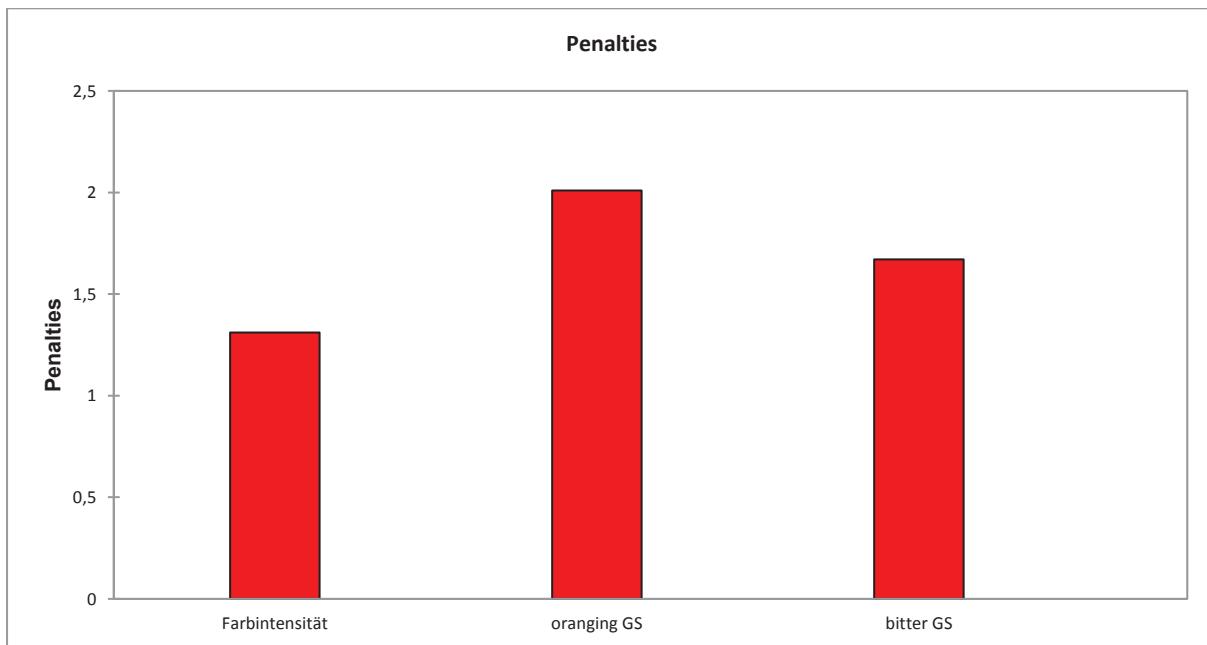


Abbildung 24: Penalties des Hohen C 100% Orangensaftes

Das Produkt müsste hinsichtlich des intensiveren Geschmackes verändert werden, jedoch kann es dazu kommen, dass die 48% der Verbraucher, die das Produkt als genau richtig empfinden, den Geschmack dann als zu intensiv betrachten. Die Farbintensität erlangt 66% der Konsumentenstimmen im JAR-Bereich und liegt nahe bei der 70%-Akzeptanzgrenze. Somit kann vorsichtig gesagt werden, dass die Farbintensität größtenteils von den Konsumenten akzeptiert wurde. Der bittere Geschmack ist schwierig zu verbessern, da 50% der Verbraucher im JAR-Bereich liegen, dieses aber für eine Akzeptanz nicht ausreicht. Mit 29% im „zu intensiv“ und 20% „zu schwach“ ist es schwierig, eine direkte Verbesserungsrichtung vorzugeben. Da der Bereich „zu intensiv“ jedoch minimal höher ist, wäre eine Verbesserung, indem die bittere Note reduziert wird, sinnvoll.

4.1.4.5 Amecke Mandarine und Orange

Der „Amecke Mandarine und Orange“ Fruchtsaft verfügt ebenfalls über eine hohe JAR-Bewertung (Abbildung 25). In den Attributen orangiger und bitterer Geschmack ist mit 50% der „genau richtig“-Bereich am höchsten belegt. Anschließend folgen die „zu schwach“-Nennungen. Die Farbintensität ist mit 50% zu intensiv, jedoch mit 44% genau richtig. Die Konsumenten sind sich mit 6% jedoch einig, dass die Farbintensität nicht zu gering ist.

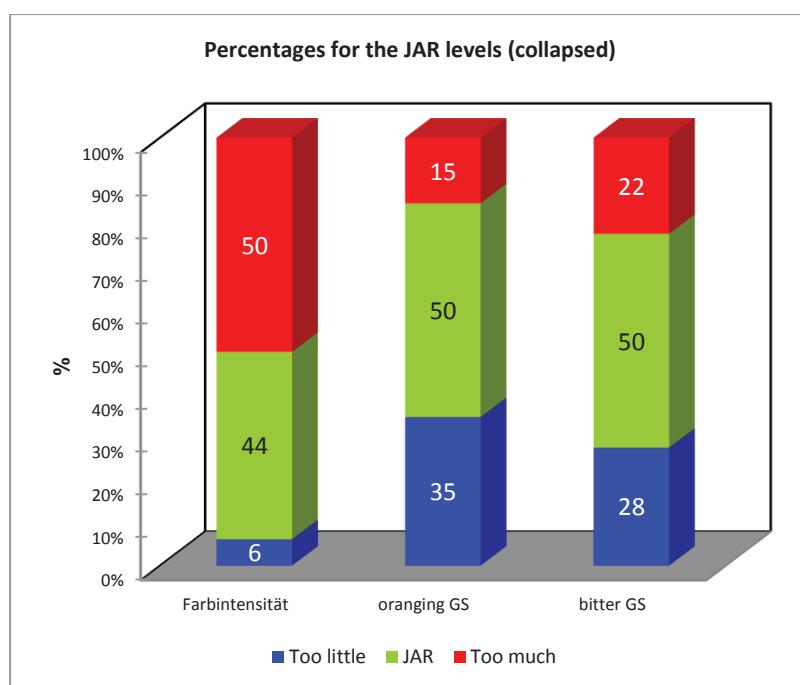


Abbildung 25: JAR-Bewertung Amecke Mandarine und Orange

Die Mean Drops (Abbildung 26) verzeichnen bei diesem Saft diesmal in einer anderen Kategorie den Maximalwert. Der bittere Geschmack ist im Bereich „zu intensiv“ signifikant mit dem höchsten Mean Drop von 2,58 belegt. In dem Bereich „zu schwach“ würden 1,38 Beliebtheitspunkte abgezogen werden

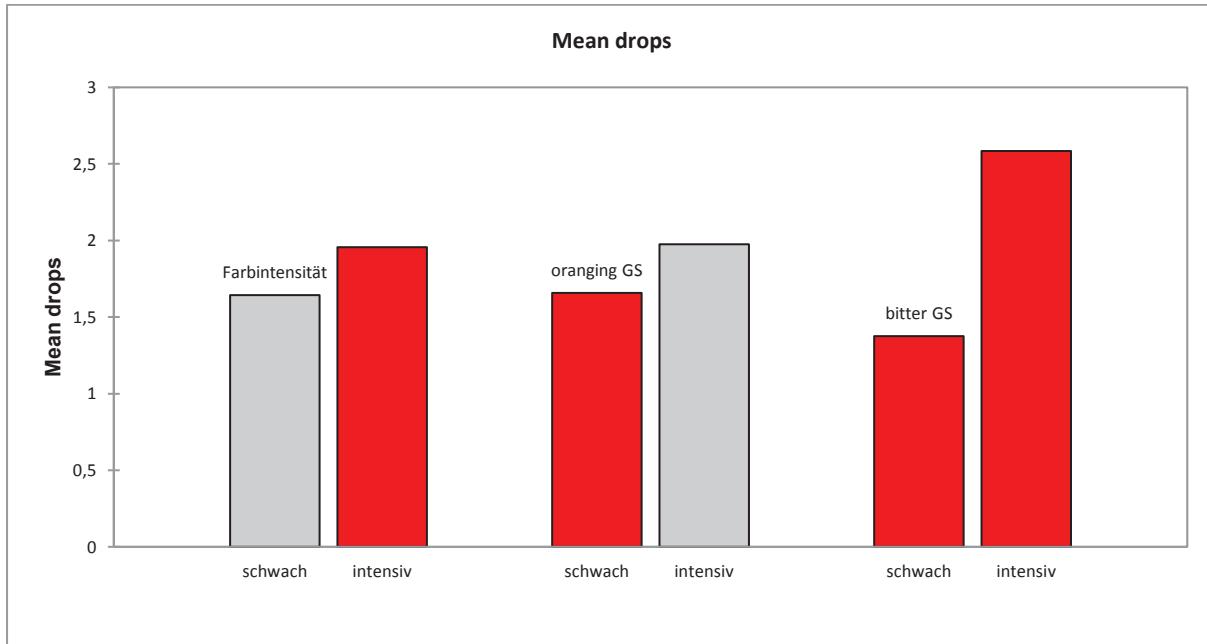


Abbildung 26: Mean Drops des Amecke Mandarine und Orange Saftes

Die Strafpunkte der Farbintensität und des bitteren Geschmackes sind mit 1,922 und 1,918 nahezu gleich hoch (Abbildung 27), sodass bei dem Nichterreichen der Erwartungen in beiden Bereichen die gleichen Abzüge in der Beliebtheit gemacht werden.

Der orangige Geschmack ist jedoch mit 1,76 Strafpunkten ähnlich hoch. Es könnte die Aussage getroffen werden, dass dem Verbraucher bei diesem Saft die Veränderungen in den drei Attributen gleich wichtig ist.

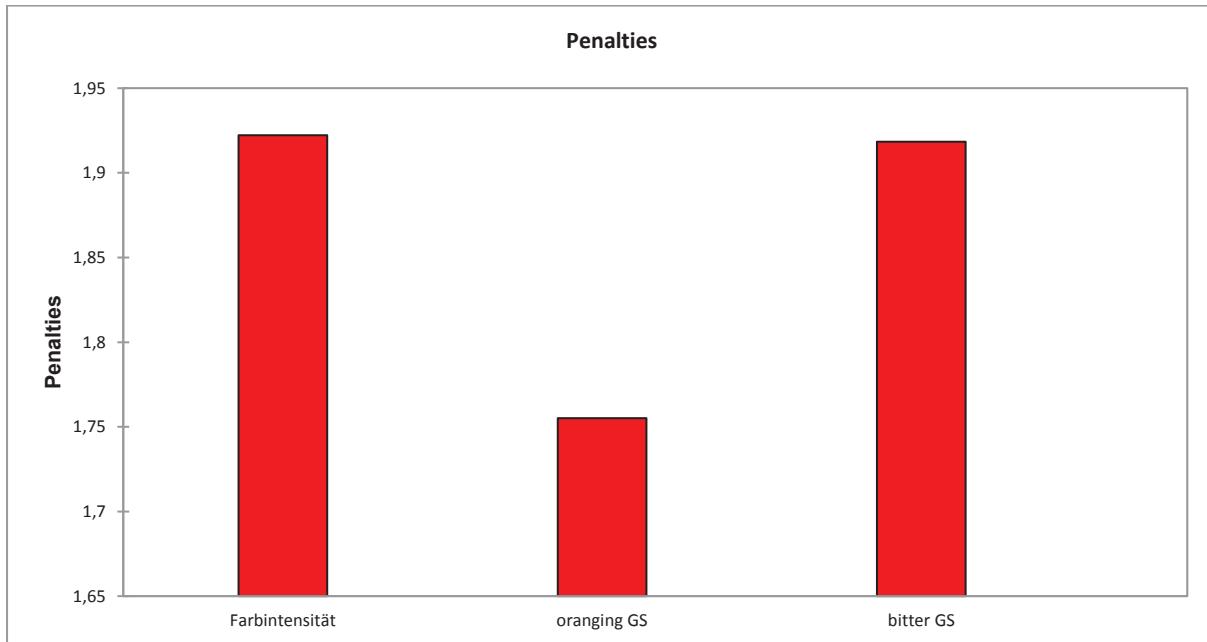


Abbildung 27: Penalties des Amecke Mandarine und Orange Saftes

Um den Bereich „genau richtig“ zu verbessern, kann das Produkt im orangigen und bitteren Geschmack in die Richtung „mehr Geschmack“ verbessert werden. Die Farbintensität kann verringert werden. Dadurch würde sich die Produktakzeptanz verbessern.

4.1.4.6 Hohes C O-T-L

Die JAR-Leves des Markenfruchtsafts „Hohes C Orange-Traube-Limette“ zeigt in jedem der drei Attribute eine andere Präferenz (Abbildung 28). Somit sind 56% der Konsumenten sich einig, dass die Farbintensität „genau richtig“ sei. Dagegen sprechen die 31% für „zu schwach“ und 13% für „zu viel“.

Der orangige Geschmack wurde hingegen mit 59% als „zu schwach empfunden. Die Kategorien „genau richtig“ und „zu stark“ beinhalten jeweils 20% der Stimmen. Zu viel bitteren Geschmack

wurde von fast 50% der Verbraucher bemängelt. Gut ein Drittel war mit der bitteren Note zufrieden und 22% würden sich über eine stärkere Ausprägung freuen.

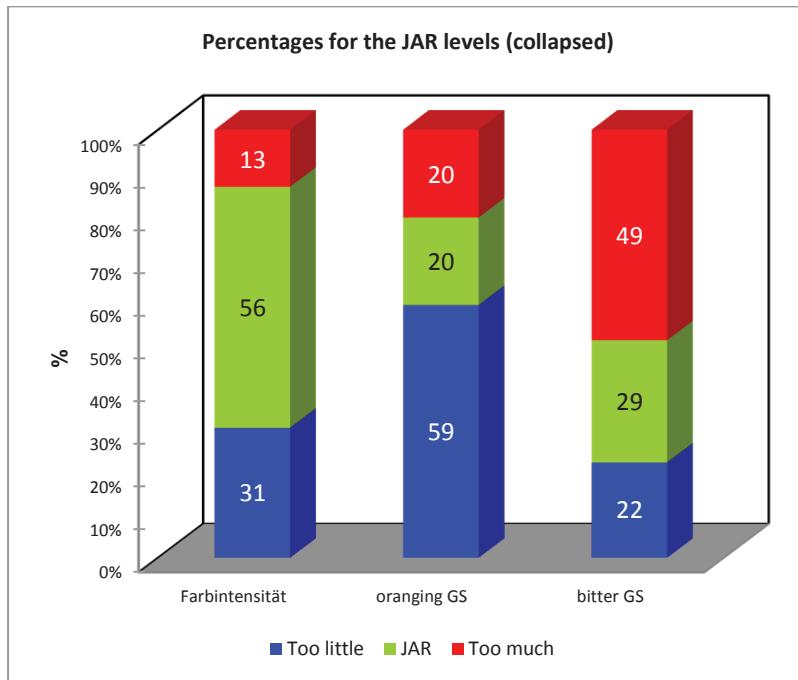


Abbildung 28: JAR-Bewertung Hohes C O-T-L

In diesem Fall (Abbildung 29) weisen die Mean Drops eine gewisse Auffälligkeit auf. Der Mean Drop der Farbintensität ist bei „zu schwach“ sehr niedrig (unter 0,5) und besitzt auch keine Signifikanz auf dem 5% Niveau. Jedoch ist festzustellen, dass bei „zu intensiv“ keine Signifikanz vorhanden ist und der Mean Drop sogar negativ (-0,04) ist. Auch hier zeigt sich wieder, dass der orangige Geschmack in „zu schwach“ den höchsten Abzug (2,85) in der Beliebtheit aufweisen würde. Der orangige Geschmack „intensiv“, aber auch die beiden Kategorien des bitteren Geschmacks, sind nahezu gleich zu verzeichnen.

Sowohl der orangige aber auch bittere Geschmack verfügen über eine Signifikanz.

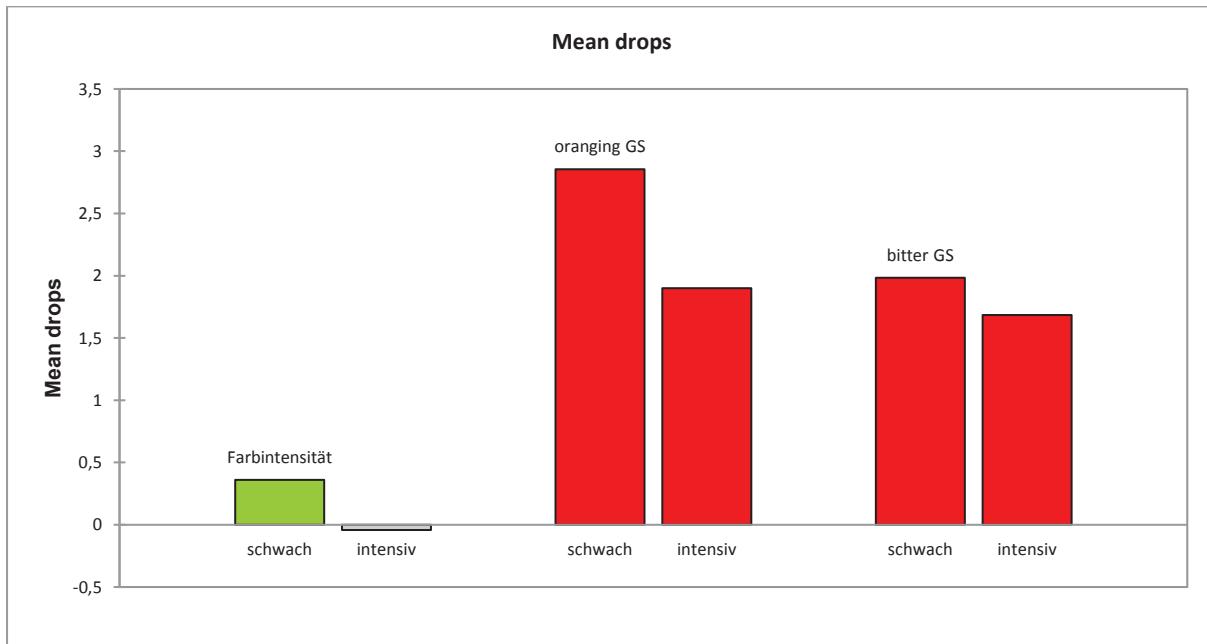


Abbildung 29: Mean Drops des Hohen C O-T-L Saftes

Die Strafpunkte untermauern die Mean Drops, indem die Farbintensität mit gerade mal 0,24 Punkten und keiner Signifikanz, der orangige Geschmack mit den höchsten Penalties von 2,61 und der bittere Geschmack mit 1,78 Strafpunkten verzeichnet ist (Abbildung 30).

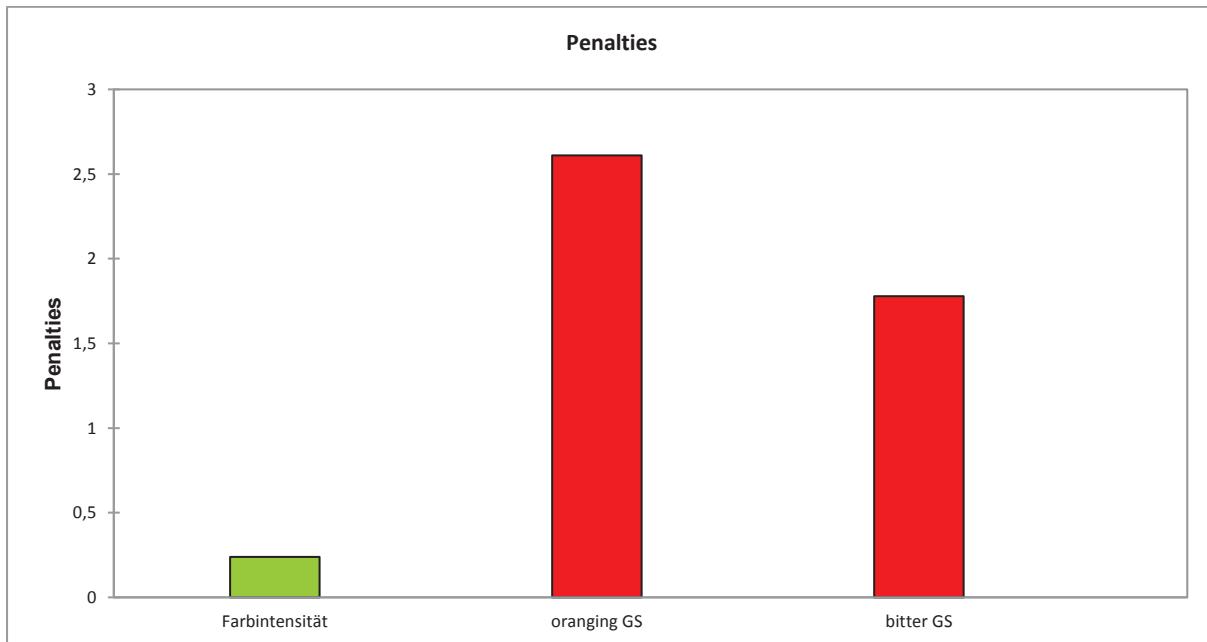


Abbildung 30: Penalties des Hohen C O-T-L Saftes

Aus diesen Ergebnissen kann nun die Schlussfolgerung gezogen werden, dass für die Verbesserung der Beliebtheit des Hohen C O-T-L Saftes eine Verstärkung des orangigen Geschmacks aber auch eine Reduzierung der bitteren Note stattfinden muss. Für die Verbesserungsvorschläge der Farbintensität kann keine genaue Aussage getroffen werden, da in beiden Fällen keine Signifikanz vorliegt.

Im Großen und Ganzen wird deutlich, dass fünf von sechs Säften („Amecke Mandarine und Orange“ hat den höchsten Mean Drop und Penalty im bitteren Geschmack „zu wenig“) einen besonders hohen Mean Drop und auch hohe Strafpunkte in dem orangigen Geschmack, vor allem in der Kategorie „zu wenig“, verzeichnen. Dadurch kann die Behauptung aufgestellt werden, dass eine hohe Konsumentenerwartung an den orangigen Geschmack vorliegt. Wird der orangige Geschmack näher zur „genau richtig“ Kategorie verbessert, erhöht sich auch die Beliebtheit der Säfte, es kann aber bei dem Nichterfüllen zu einer starken Senkung führen. Lediglich der „Hohes C 100% Orangensaft“ hat in der Kategorie Farbintensität „genau richtig“ eine annähernde Akzeptanz mit 66% erreicht.

4.1.5 CATA

Für die Auswertung des CATA-Fragebogens des Verbrauchertestes wird im ersten Schritt ein Überblick geschaffen, wie viele Attribute durch die Konsumenten benutzt wurden. In der Abbildung 31 wird ein Balkendiagramm dargestellt. Der grüne Teil des Balkendiagramms hat die Aussage, wie viele Konsumenten in Prozent das Attribut gewählt haben. Die roten Balken treffen die Aussage, wie viel Prozent der Konsumenten das Produkt nicht gewählt haben. Dadurch wird deutlich, dass das Attribut „süß“ mit 50% im grünen Bereich am häufigsten durch die Konsumenten benutzt wurde, um ein Attribut zu beschreiben.

Ebenfalls häufig benutzt wurden „gelblich“, „orangenartig“, „orangenfarbend“ und künstlich. Das Attribut „bräunlich“ wurde mit 17% am seltensten benutzt.

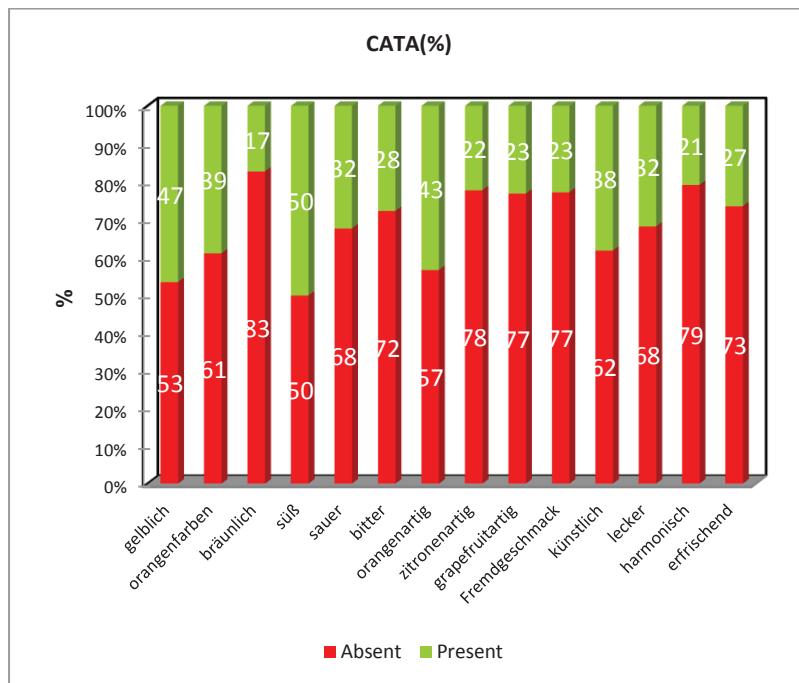


Abbildung 31: Nutzung der CATA-Attribute in Prozent

Des Weiteren wurde ein sogenannter „Cochran's Q-Test“ durchgeführt. Bei diesem Test wird deutlich, welche Produkte sich anhand der Attribute deutlich voneinander unterscheiden, aber auch welche Produkte z.B. besonders süß und welche am wenigsten süß sind.

Tabelle 13: Cochran's Q-Test

| Attributes | p-values | Punica | Am. Mand. | Hohes C | | Tip O-Nektar | Hohes C O-T-L |
|-----------------|----------|------------|-------------|------------|-------------|--------------|------------------|
| | | | | 100% Or. | Am. Zink | | |
| gelblich | 0,000 | 0.872 (d) | 0.074 (a) | 0.606 (bc) | 0.053 (a) | 0.723 (cd) | 0.468 (b) |
| orangefarben | 0,000 | 0.032 (a) | 0.681 (d) | 0.351 (bc) | 0.649 (d) | 0.191 (b) | 0.436 (c) |
| bräunlich | 0,000 | 0.021 (a) | 0.457 (b) | 0.011 (a) | 0.426 (b) | 0.043 (a) | 0.085 (a) |
| süß | 0,000 | 0.606 (b) | 0.553 (ab) | 0.351 (a) | 0.372 (a) | 0.734 (b) | 0.394 (a) |
| sauer | 0,000 | 0.181 (a) | 0.223 (ab) | 0.564 (c) | 0.383 (bc) | 0.181 (a) | 0.415 (bc) |
| bitter | 0,000 | 0.266 (b) | 0.266 (b) | 0.309 (b) | 0.426 (b) | 0.074 (a) | 0.330 (b) |
| orangenartig | 0,000 | 0.277 (a) | 0.574 (b) | 0.532 (b) | 0.404 (ab) | 0.543 (b) | 0.277 (a) |
| zitronenartig | 0,000 | 0.330 (b) | 0.106 (a) | 0.309 (b) | 0.106 (a) | 0.170 (ab) | 0.319 (b) |
| grapefruitartig | 0,000 | 0.160 (ab) | 0.223 (ab) | 0.202 (ab) | 0.447 (c) | 0.064 (a) | 0.298 (bc) |
| Fremdgeschmack | 0,000 | 0.351 (cd) | 0.160 (abc) | 0.064 (a) | 0.255 (bcd) | 0.149 (ab) | 0.394 (d) |
| künstlich | 0,000 | 0.670 (b) | 0.170 (a) | 0.202 (a) | 0.245 (a) | 0.511 (b) | 0.500 (b) |
| lecker | 0,000 | 0.170 (a) | 0.415 (b) | 0.404 (b) | 0.383 (ab) | 0.319 (ab) | 0.223 (ab) |
| harmonisch | 0,000 | 0.096 (a) | 0.394 (b) | 0.202 (ab) | 0.234 (ab) | 0.181 (a) | 0.149 (a) |
| erfrischend | 0,231 | 0.245 (a) | 0.245 (a) | 0.309 (a) | 0.245 (a) | 0.340 (a) | 0.213 (a) |

Die Interpretation der Tabelle 13 erfolgt, indem die von Xlstat berechneten Zahlenwerte und die Buchstaben genauer betrachtet werden. Ein besonders hoher Wert bedeutet, dass hier die höchste Intensität gespürt wurde und ein niedriger Wert hat zur Folge, dass bei dem Produkt die geringste Intensität vorhanden war. So ist z.B. „Punica“ besonders gelb, aber diese Probe ist nicht signifikant gelber als „Tip O-Nektar“, denn beide besitzen den Buchstaben „d“ und sind somit in einer Gruppe. Der Saft „Am. Zink plus Vitamin C“ besitzt die geringste gelbe Note, ist aber nicht signifikant „gelber“ als „Am. Mandarine-Orange“. Die p-values geben Auskunft über die signifikanten Unterschiede und haben somit direkt mit der Gruppeneinteilung in Buchstaben zu tun. Bei einem p-value von 0,05 weisen alle Attribute einen Wert von „0“ auf. Dies bedeutet, dass die Produkte anhand der Attribute unterschiedlich sind. Lediglich das Attribut „erfrischend“ hat keine Signifikanz. Dies hat zur Folge, dass alle Produkte der gleichen Gruppe „a“ angehören. Die Werte der Produkte sind auch ähnlich hoch. Somit können die Produkte anhand des Attributes „erfrischend“ nicht unterschieden werden.

Um herauszufinden welche Attribute für eine besondere Beliebtheit sorgen, wird ein PCA-Plot angefertigt.

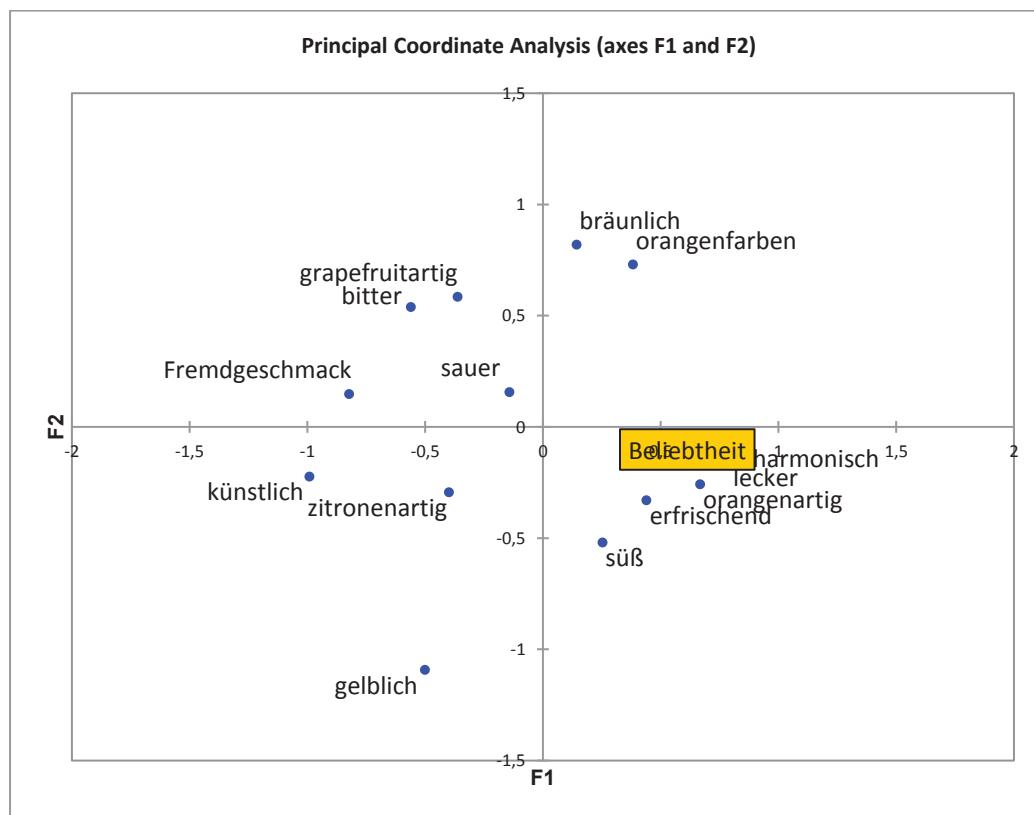


Abbildung 32: Beliebtheitsüberblick anhand eines PCA-Plots

In der Abbildung 32 ist ein PCA-Plot mit den verwendeten Attributen zu sehen.

Außerdem ist ein gelbes Feld zu sehen, welches die Beliebtheit darstellt. Je näher die Attribute an der „Beliebtheit“ liegen, desto wichtiger sind diese. Somit wird deutlich, dass ein harmonischer und leckerer Eindruck, die Orangenartigkeit, ein süßer Geschmack und ein erfrischendes Gefühl für eine hohe Beliebtheit sorgen. Dahingegen ist ein Fremdgeschmack oder eine hohe künstliche Eigenschaft kontraproduktiv. Ebenfalls ersichtlich ist, dass Prüfer die ein süßes und orangenartiges Produkt bevorzugen, eine Abneigung gegenüber bitteren und grapefruitartigen Produkten aufweisen und umgekehrt genauso. Des Weiteren korreliert das Attribut „bitter“ mit „grapefruitartig“. Somit interpretiert der Verbraucher, wenn er „grapefruitartig“ herausschmeckt, dies direkt mit einer bitteren Note. „Orangenfarbend“ und „gelblich“ werden durch das Konsumentenpanel eindeutig unterschieden, wohingegen „orangenfarbend“ und „bräunlich“ stark korrelieren. Es kann die Aussage getroffen werden, dass „orangenfarbend“ ähnlich wie „bräunlich“ interpretiert wird und eher zur Verwirrung führt, als das es hilft.

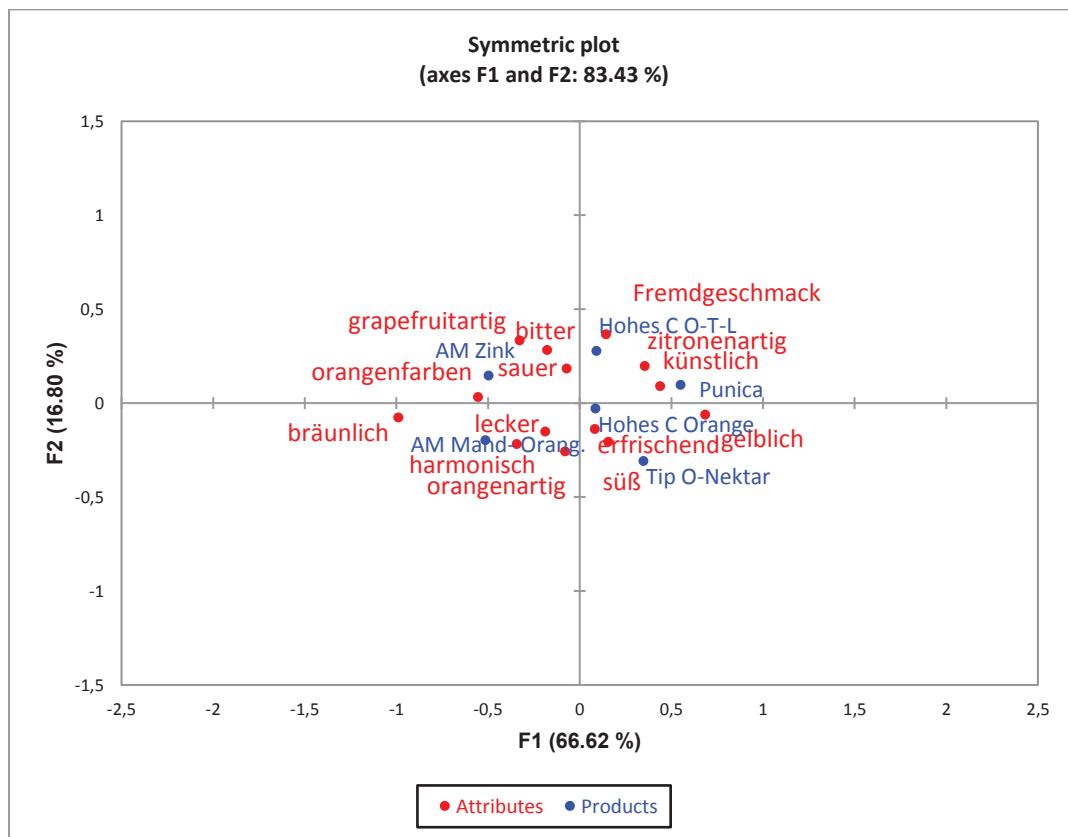


Abbildung 33: PCA Plot mit Produkten und Attributen

Anhand der Abbildung 33 kann nun gesehen werden, welche Produkteigenschaften die getesteten Produkte aufweisen. Die Produkte „TIP O-Nektar“ und „Hohes C 100% Orange“ sind besonders süß, erfrischend und gelblich. Außerdem ist der Saft „Amecke Mandarine-Orange“ be-

sonders lecker, harmonisch und orangenartig. Somit weisen diese Produkte genau die Merkmaleigenschaften auf, die in der Abbildung 32 für eine besonders hohe Beliebtheit verantwortlich sind. Das Produkt „Punica“ wird als besonders künstlich und zitronenartig wahrgenommen. In der ähnlichen Position befindet sich das Produkt „Hohes C O-T-L“, welches zusätzlich über einen Fremdgeschmack verfügt und sauer ist. Diese zwei Produkte beinhalten die Attribute, die für eine schwache Beliebtheit verantwortlich sind.

Um herauszufinden welche Attribute die „must haves“ sind, also welche Produkteigenschaften vorhanden sein sollten, um eine hohe Beliebtheit des Produktes zu gewährleisten, wird ein sogenannter „Mean Impact“ erstellt, auf einem Signifikanzniveau von 5%.

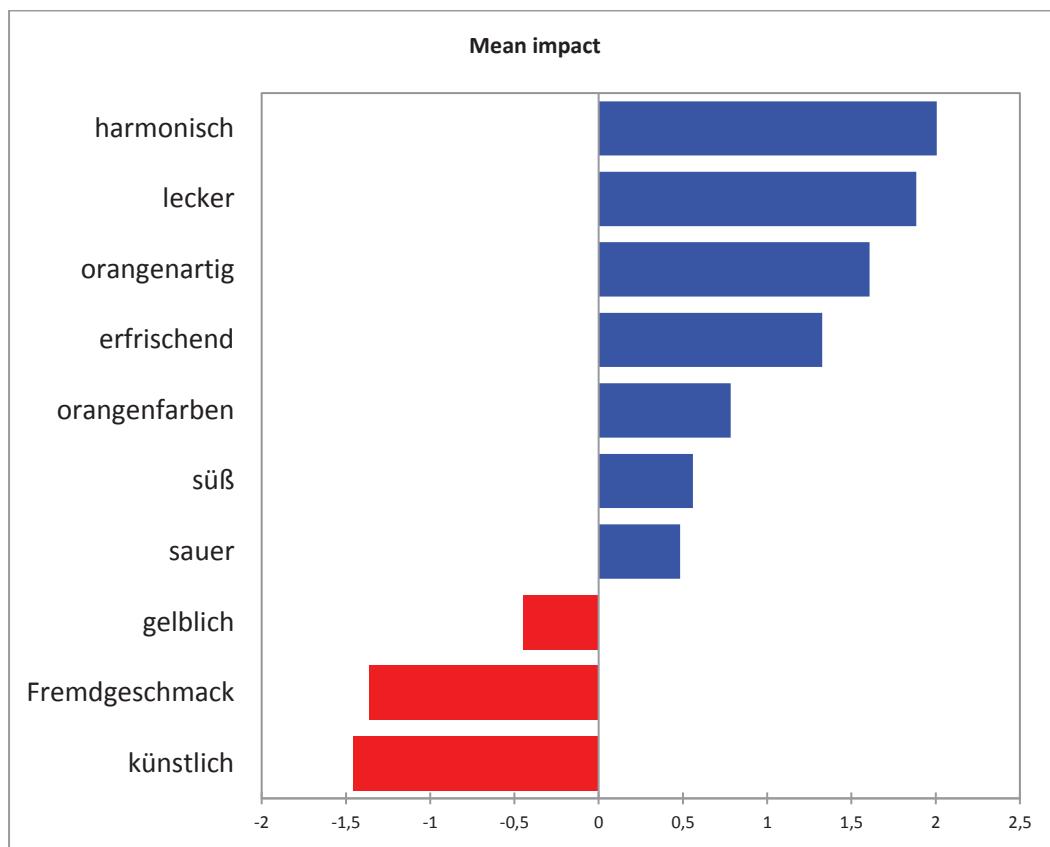


Abbildung 34: Mean Impact

Der „Mean Impact“ in der Abbildung 34 zeigt, dass ein Produkt besonders harmonisch, lecker, orangenartig erfrischend, orangefarbig, süß und zugleich sauer sein muss, um eine hohe Beliebtheit zu erhalten. Eine hohe künstliche und gelbliche Eigenschaft, sowie ein Fremdgeschmack sollten in einem Fruchtsaft nicht vorhanden sein. Für die Attribute „bräunlich“, „bitter“, „zitronenartig“ und „grapefruitartig“ konnte kein Mean Impact bestimmt werden, da diese Attribute keine Signifikanz aufweisen. Somit kann anhand dieser vier Attribute kein Zusammenhang mit der Beliebtheit hergestellt werden.

4.2 Profilprüfung

Die Profilprüfung wurde in einem Sensorikraum nach ISO 8589 durchgeführt. Der Sensorikraum war bei allen drei Sitzungen gleichmäßig auf 20,2°C +/- 0,1°C temperiert. In den drei Durchgängen betrug die relative Luftfeuchtigkeit 59% +/- 1%.

4.2.1 Gruppe 1

Um einen ersten Überblick über die Bewertung des Panels zu erhalten, wird ein Anova-Graph projiziert (Abbildung 35), in dem die Attribute mit den Signifikanzen angezeigt werden. Die Gruppe 1 hat für ihre Untersuchung 14 Attribute verwendet, davon sind 12 höchstsignifikant, ein Attribut hochsignifikant und ein Attribut nichtsignifikant. Die Abbildung 35 sagt aus, mit welcher Signifikanz die Produkte anhand dieser Attribute unterschieden werden konnten. So wird deutlich, dass das Panel anhand des Attributes „süßlich Geruch“ die Produkte nicht eindeutig unterscheiden konnte.

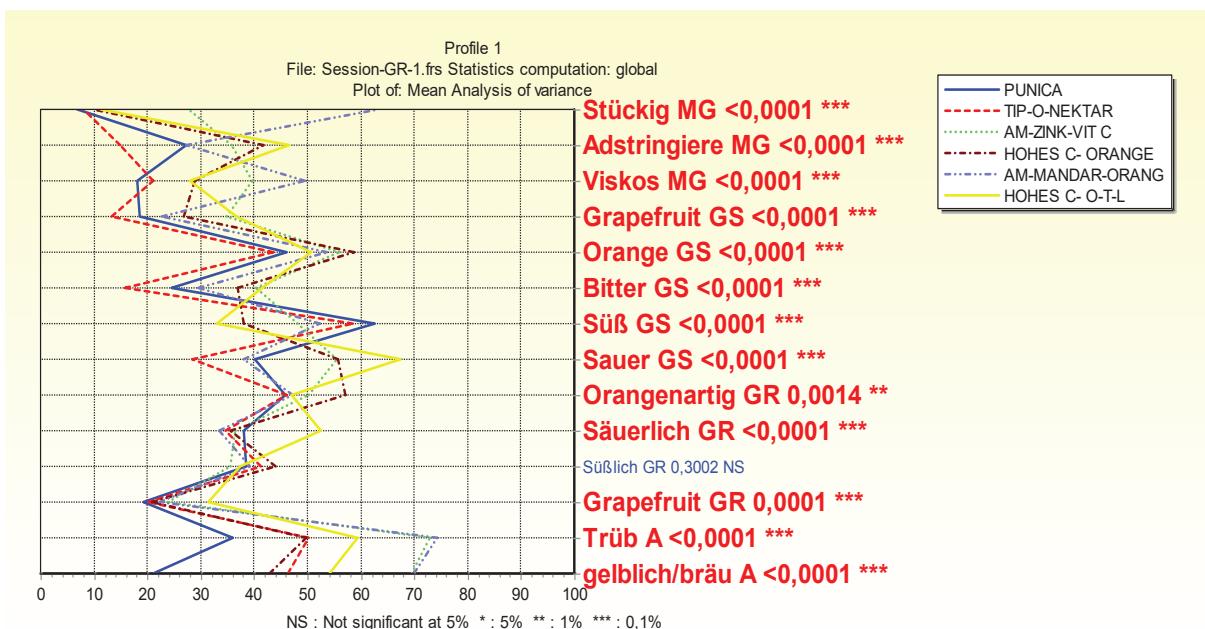


Abbildung 35: Anova-Graph der Gruppe 1

Betrachtet man für das Attribut „süßlich Geruch“ die durchschnittlichen Bewertungen der Probanden, so reichen die durchschnittlichen Skalenwerte von 0 – 57,39. Die Anova-Tabelle zeigt auch eindeutig, weshalb es zu einer Nichtsignifikanz kommt. Es ist zu sehen, dass der Faktor „Produkt“ über keine Signifikanz verfügt, das heißt, die Produkte konnten anhand des Attributes

nicht unterschieden werden. Dafür liefert die Tabelle 13 eine Höchstsignifikanz in dem Faktor Prüfer. Das hat die Aussage, dass die Prüfer alle unterschiedlich mit diesem Attribut umgehen. Dies erklärt auch die breite Skalennutzung von 0- 57,39. Bei der Interaktion zwischen Produkt und Prüfer liegt keine Signifikanz vor, somit sind Wechselwirkungen zwischen den zwei Faktoren ausgeschlossen.

Tabelle 14: Anova „süßlich Geruch“ Gruppe 1

| Sources of variation | D.F. | S.S. | M.S. | Comp. F | Proba. |
|------------------------|------|----------|---------|---------|----------------|
| Product | 5 | 1827,21 | 365,44 | 1,22 | 0,3002 |
| Judge | 13 | 66416,75 | 5108,98 | 17,1 | <0,0001 *** |
| Interaction 1-2 | 65 | 26886,18 | 413,63 | 1,38 | 0,0507 |

Um die Durchschnittsbewertungen der einzelnen Säfte visuell zu verdeutlichen, wird die Abbildung 36 genauer betrachtet. Der dicke schwarze Strich, der von links nach rechts nahezu waagerecht durch die Abbildung geht, ist der Gruppendurchschnitt. Die unterschiedlichen Bewertungen des Panels sind hier deutlich. Es ist zu erkennen, dass der Prüfer „Al“ bei dem Saft „Hohes C O-T-L“ die höchste säuerliche Geruchsnote wahrnimmt. Wohingegen der Proband „Han“ bei dem gleichen Saft keine Geruchsnote wahrnimmt.

Der Proband „Gro“ bewertet die Säfte alle im niedrigen Skalenniveau unter 20, außer minimal über 20 bei dem „Hohes C O-T-L“. Wohingegen Prüfer „Hel“ und „Erm“ im oben Skalenbereich die Bewertungen abgeben und „Beh“ die Skalenbereiche weiter ausnutzt.

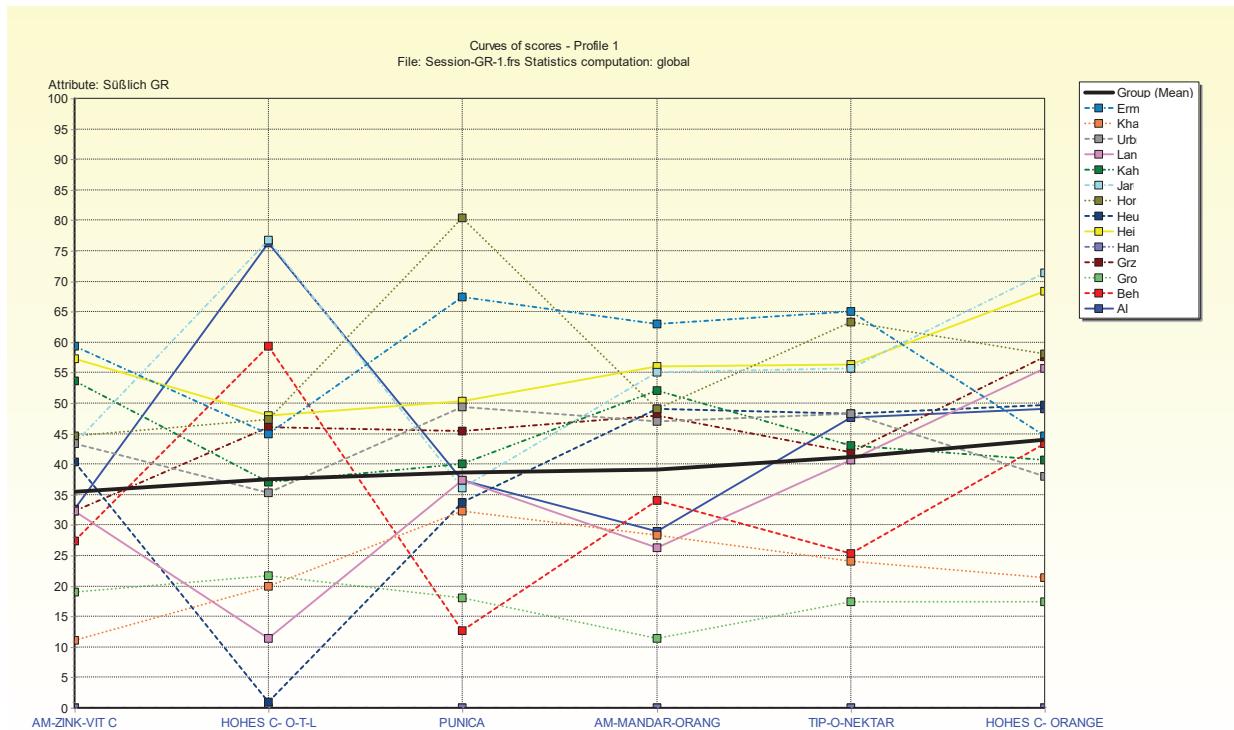


Abbildung 36: Durchschnittsbewertungen des Attributes „süßlich Geruch“

Die Abbildung 36 verdeutlicht, dass das Panel sich nicht einig ist, was dieses Attribut betrifft. Als Vergleichsbeispiel dient die Abbildung 37, die die Durchschnittsbewertungen des Attributes „gelblich/bräunlich Aussehen“ zeigt. Hier sind die meisten Bewertungen ähnlich und nahe dem Gruppendurchschnitt. Somit bewerten die Probanden nicht nahezu gleich, da die Skalennutzung unterschiedlich ist, jedoch besitzt jeder Graph die gleiche Aussage. In der Abbildung 35 ist auch zu sehen, dass dieses Attribut höchstsignifikant ist.

In der Abbildung 37 wird der durchschnittliche Intensitätsverlauf der Produkte anhand der Konsenssteigerung (dicker schwarzer Strich), dargestellt. Die Produkte liegen dabei in Rangfolge von weniger intensiv bis hin zu stark intensiv. Das Panel ist sich im Großen und Ganzen über die Steigerung einig. Es existieren aber auch einige Ausreißer, wie z.B. Proband „Grz“ bei dem Produkt „Tip O-Nektar“. Hier steigt die Intensität, jedoch fällt diese bei diesem Prüfer. Ähnlich bei „Erm“ und dem Produkt „Amecke Zink plus Vitamin C“.

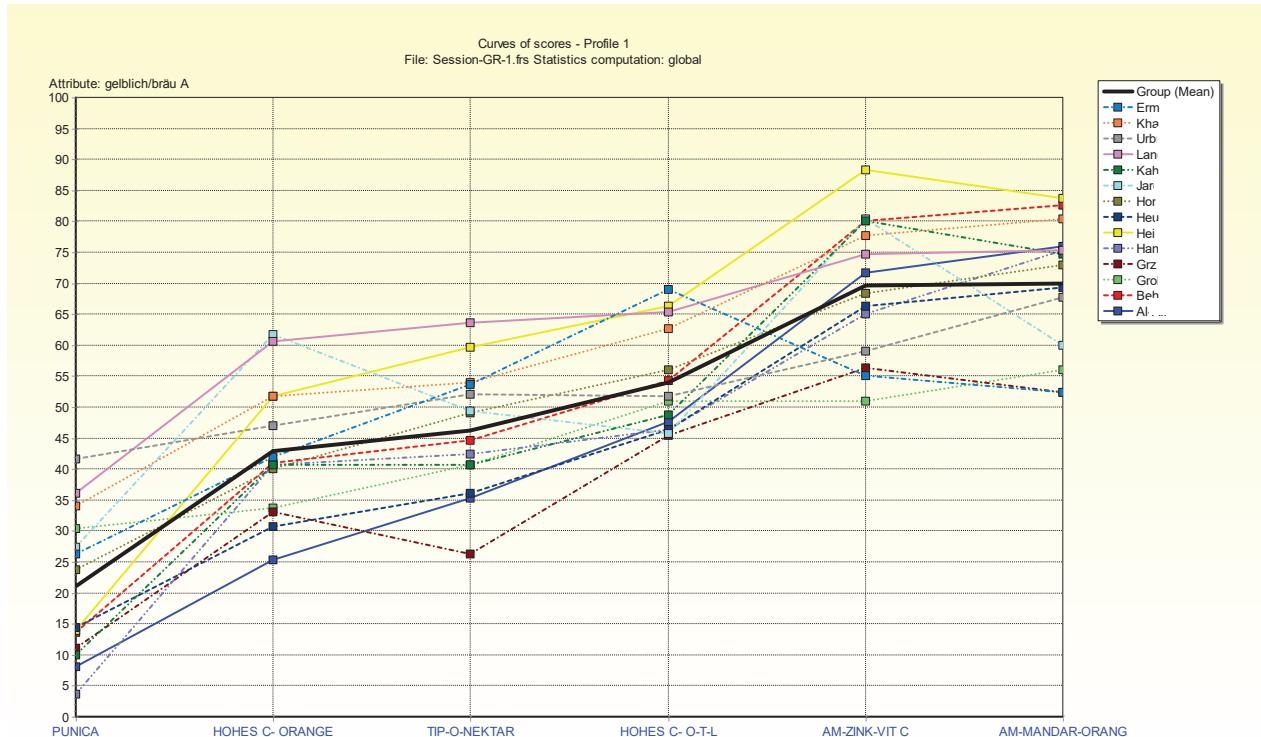


Abbildung 37: Durchschnittsbewertungen anhand des Attributes „gelblich/bräunlich Aussehen“

Werden die Rohwerte betrachtet, so wird besonders in der Kategorie „Geruch“ deutlich, dass ein Proband mit seinen Werten heraussticht. Der Prüfer „Han“ hat in allen Attributen der Kategorie „Geruch“ eine durchschnittliche Bewertung kleiner 1. Hier sollte der Proband aus dem Panel herausgezogen werden und ein klarendes Gespräch und gegebenenfalls Schulungen im Bereich „Geruch“ durchgeführt werden.

Der PCA-Plot (Abbildung 38) zeigt 90,1% der Daten auf zwei Achsen. Die erste Achse (x-Achse), kann auch als die Farbintensitätsachse interpretiert werden, da die farbgebenden Attribute den geringsten Winkel zur Achse aufweisen. Die zweite Achse (Y-Achse), kann als die Geschmacks- und Geruchsachse, vor allem als „süß-sauer“-Achse betrachtet werden. Auch hier besitzen die Attribute „sauer“ und „süß“ einen sehr niedrigen Achsenradius. Des Weiteren ist

„süß“ ca. 180° von „sauer“ entfernt und somit ein Gegensatz. Eine weitere Bemerkung wird anhand der Positionen der Produkte gefällt.

Je näher ein Produkt am Ursprung liegt, desto weniger unterscheidet es sich von den anderen Produkten. In diesem Fall trifft es am ehesten auf das Produkt „Amecke Zink plus Vitamin C“ zu.

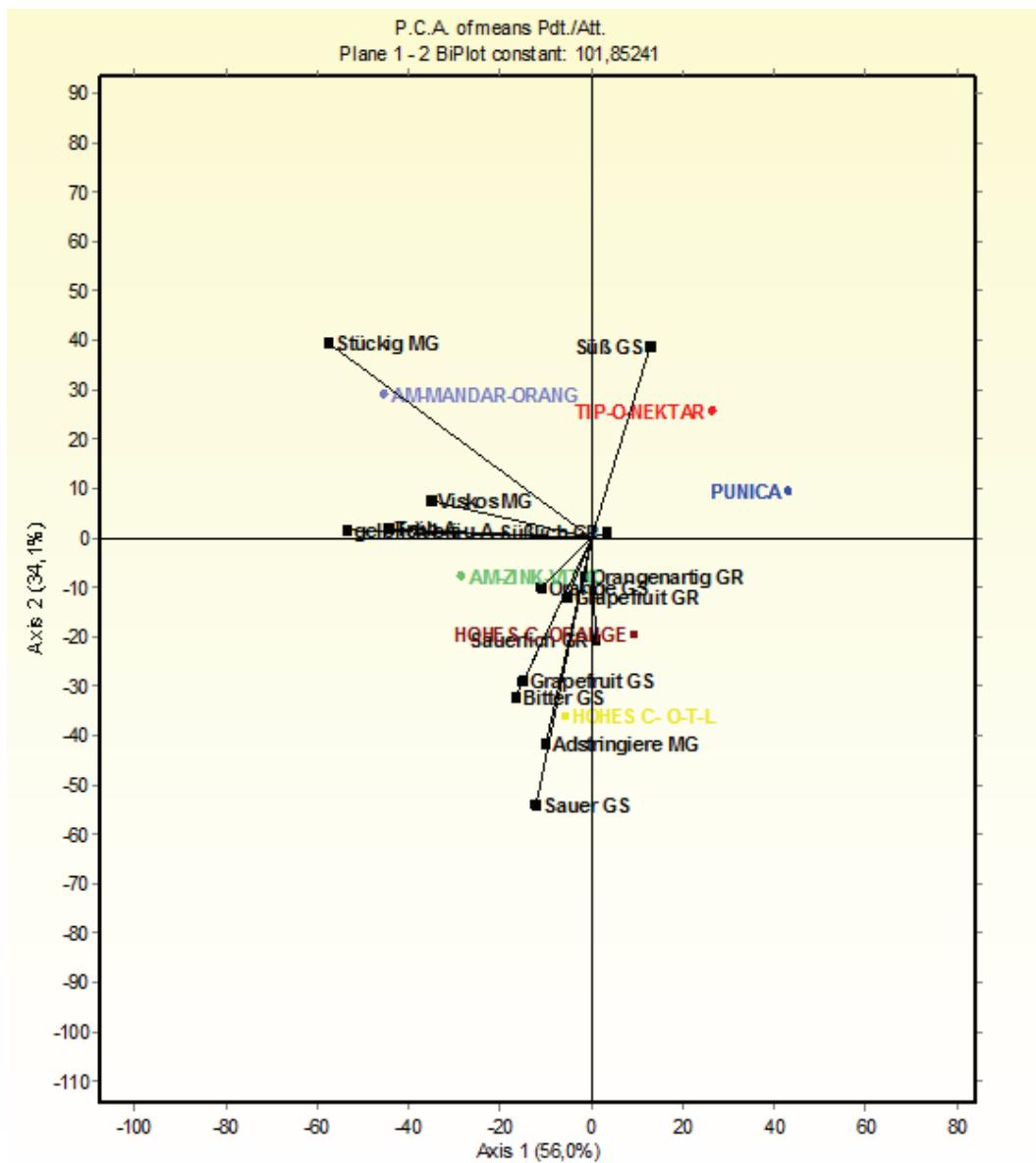


Abbildung 38: PCA-Plot Gruppe 1

Es werden die einzelnen Produkte dargestellt und deren Produkteigenschaften, die durch das Panel bestimmt wurden. Es ist zu erkennen, dass das Produkt „Amecke Mandarine-Orange“ stückig und viskos im Mundgefühl, aber auch süß ist. Besondere Merkmale des Produktes „Tip

Orangennektar“ ist die Süße, genau wie bei „Punica“. Außerdem ist „Punica“ gelblich und am klarsten. Diese Aussage kann getroffen werden, da die Attribute „bräunlich-orange“ und „trüb“ genau auf der anderen Seite der Y-Achse liegen. Das Produkt „Amecke Zink plus Vitamin C“ ist besonders gelb, besitzt aber auch einen grapefruit- und orangenartigen Geruch. Des Weiteren wird es mit einem hohen orangigen Geschmack in Bezug gebracht. Das „Hohes-C O-T-L“ ist recht bitter und adstringierend, aber auch deutlich saurer als z.B. das „TIP Orangennektar“. Auch zu erkennen ist, dass eine bittere Note direkt mit Grapefruit in Verbindung gebracht wird, da diese zwei Produkteigenschaften sehr nah beieinanderliegen. Das Attribut „süßlich Geruch“ ist sehr nahe am Schnittpunkt der Achsen, was bedeutet, dass die Probanden ziemliche Schwierigkeiten haben eine eindeutige Eingliederung bzw. Aussage über dieses Attribut zu treffen. Somit besitzt dieses Attribut eine sehr geringe Unterscheidungskraft. Die Produkte „Amecke Mandarine-Orange“ und „Hohes- C O-T-L“ sind am weitesten voneinander entfernt, sodass gesagt werden kann, dass die beiden Produkte vom der Gruppe 1 als am unterschiedlichsten empfunden werden.

Nach der ISO DIN EN 11132 kann die Aussage getroffen werden, dass das Panel von 14 Merkmalseigenschaften 13 signifikant unterschieden hat. Dadurch besitzt das Panel eine hohe Leistungsfähigkeit, es muss jedoch anhand des Attributes „süßlich Geruch“ geschult werden, bzw. wird Klärungsbedarf benötigt, ob dieses Attribut sinnvoll ist. Die Homogenität des Panels ist gering, da von 14 Attributen 10 Attribute eine signifikante Interaktion zwischen Prüfer und Probe aufwiesen. Ein geschultes Panel darf keine Wechselwirkungen aufweisen.

Mit Hilfe der Tabelle 15 kann nun die Leistungsprüfung der Einzelpersonen betrachtet werden. Es folgte eine Berechnung der Tabellen nach DIN EN ISO 11132 für das Attribut „orangiger Geschmack“.

Tabelle 15: Leistungsprüfung der Einzelpersonen Gruppe 1

| | Individuelle syst. Abweichung | S.D. der Syst. Abweichungsterme von einzelnen Prüfern | Mittelwert der Standardabweichung |
|------------|--------------------------------------|--|--|
| Al | -7,13 | 16,09 | 29,96 |
| Beh | -10,18 | 11,15 | 35,6 |
| Gro | -21,85 | 10,19 | 9,79 |
| Grz | 9,43 | 13,1 | 34,44 |
| Han | -7,85 | 24,4 | 33,08 |
| Hei | 22,15 | 6,75 | 13,65 |
| Heu | -13,24 | 22,04 | 25,16 |
| Ho | 1,15 | 14,6 | 22,13 |
| Jar | 12,43 | 8,31 | 34,64 |
| Kah | -13,79 | 14,34 | 38,31 |
| Lan | -6,24 | 4,14 | 32,15 |
| Urb | 33,93 | 15,2 | 27,1 |
| Kha | 5,49 | 19,86 | 18,77 |
| Erm | -4,29 | 13,4 | 7,38 |

Die Interpretation erfolgt so: Eine besonders hohe individuelle systematische Abweichung im negativen Bereich bedeutet, dass der Prüfer deutlich kleinere Zahlenwerte als Bewertung benutzt hat, als die übrigen Mitglieder. In diesem Fall hat der Prüfer „Gro“ mit -21,85 die niedrigste individuelle syst. Abweichung. Somit benutzt dieser Proband niedrigere Zahlenwerte zur Bewertung als der Rest des Panels. Der Proband „Urb“ liegt mit einer individuellen syst. Abweichung von 33,93 am höchsten. Somit benutzt er besonders große Zahlen zur Bewertung des Attributes. Die Standardabweichung der syst. Abweichungsterme von einzelnen Prüfpersonen deutet auf die Konsistenz der Prüfer hin. Der Prüfer „Heu“ besitzt den höchsten Wert mit 22,04. Werden die Rohdaten von Prüfer „Heu“ in Betracht gezogen bzw. die systematischen Abweichungen im Einzelnen, so wird deutlich, dass seine niedrigste Abweichung vom Panel -51,3 und seine höchste Abweichung 12,31, beträgt. Dies hat zur Folge, dass Proband „Heu“ die geringste Konsistenz des ganzen Panels besitzt. Dahingegen besitzt Proband „Lan“ mit 4,14 die geringste Standardabweichung der syst. Abweichungsterme. Dies zeigt sich auch anhand der syst. Abweichung in Excel. Maximale Abweichung vom Panelmittelwert: -12,3 und niedrigste Abweichung des Panelmittelwertes: -1,36. Bei diesem Probanden ist auch die individuelle systematische Abweichung am höchsten.

chung im Vergleich zum Restpanel gering. Somit besitzt dieser Proband eine hohe Leistungsfähigkeit, was das Arbeiten mit dem Attribut „orangiger Geschmack“ angeht.

Die Wiederholpräzision wird anhand des Mittelwertes der Standardabweichung festgehalten. Dabei gilt je niedriger der Wert, desto ähnlicher war die Bewertung in den drei Sitzungen. Die Prüfer „Erm“ (7,38) und „Gro“ (9,79) besitzen den niedrigsten Mittelwert der Standardabweichung. Somit bewerten diese Prüfer in den 3 Durchgängen immer ähnlich. Das bedeutet, dass diese Prüfer eine sehr gute Wiederholbarkeitsrate aufweisen. Dahingegen besitzen die Prüfer „Kah“ (38,31), „Jar“ (34,64), „Grz“ (34,44), „Han“ (33,08), „Beh“ (35,1) und „Lan“ (32,15) einen Mittelwert von über 30. Somit besitzen diese Prüfer eine hohe Spannweite der Bewertung, in den drei Sitzungen. „Kah“ besitzt mit 38,31 den höchsten Mittelwert der Standardabweichung und somit die schlechteste Wiederholbarkeitsrate. Werden die Rohdaten betrachtet, so wird deutlich, dass „Han“ in fünf Attributen die schlechteste Wiederholbarkeit des gesamten Panels aufweist. Proband „Kah“ besitzt starke Schwierigkeiten mit dem Attribut „stückig MG“. „Beh“ sollte in den Attributen „grapefruit im Geruch“, „sauer im Geschmack“ und „viskos MG“ geschult werden.

4.2.2 Gruppe 2

Wie bei Gruppe 1 wird für den ersten Überblick der Anova-Graph projiziert. Die Abbildung 39 zeigt, dass die Gruppe 2 insgesamt mit 17 Attributen gearbeitet hat. Von diesen 17 Attributen sind alle, außer das Attribut „süßlich Geruch“, höchstsignifikant. Das bedeutet, dass die Produkte anhand dieser 16 Attribute signifikant zu unterscheiden sind. Wie bereits Gruppe 1, hat das Panel Schwierigkeiten mit dem Attribut „süßlich Geruch“.

Für das nicht signifikante Attribut wird die Anova-Tabelle (Tabelle 16) zur Betrachtung herangezogen. Somit kann herausgefunden werden, an welchem Faktor die Nichtsignifikanz zu verantworten ist.

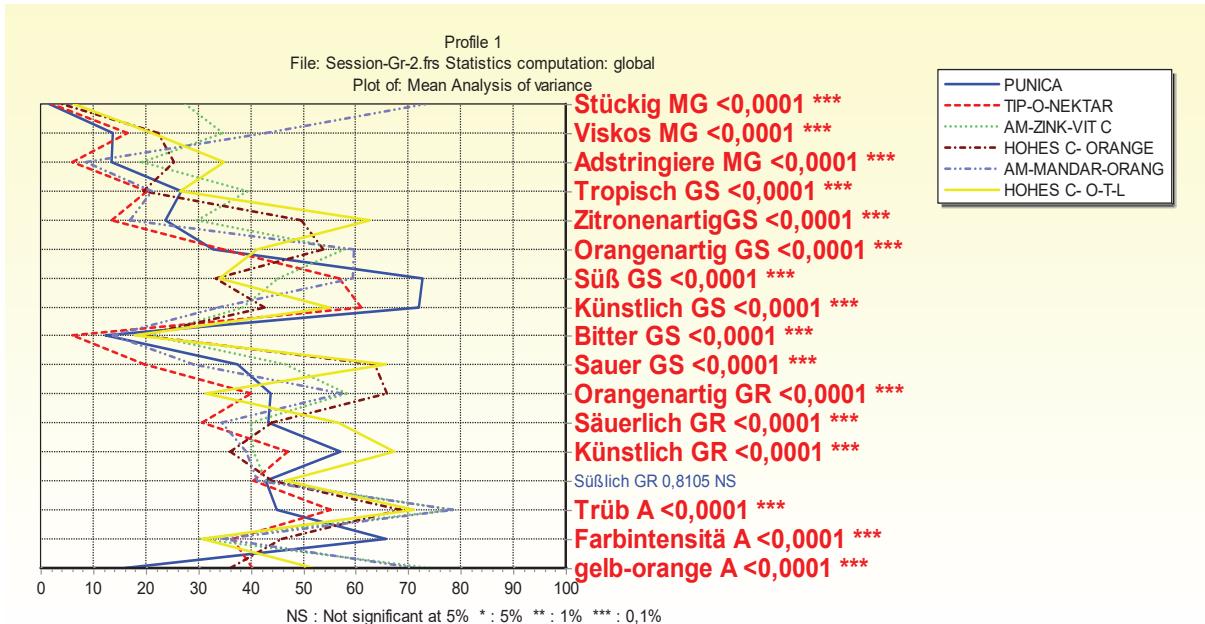


Abbildung 39: Anova-Graph Gruppe 2

Es ist zu erkennen, dass der Faktor „Produkt“ nicht signifikant ist. Aus diesem Grund kann die Annahme getroffen werden, dass die Produkte anhand des Attributes „süßlich Geruch“ nicht unterschieden werden können. Des Weiteren besteht eine Höchstsignifikanz bei dem Faktor Prüfer, was bei einem geschulten Panel ein negatives Anzeichen ist. Die Prüfer waren sich mit ihrer Bewertung nicht einig und so wie bei Gruppe 1 ist hier Gesprächsbedarf notwendig. Außerdem besteht zusätzlich eine Hochsignifikanz in den Wechselwirkungen zwischen Produkt und Prüfer.

Tabelle 16: Anova-Tabelle „süßlicher Geruch“ Gruppe 2

| Sources of variation | D.F. | S.S. | M.S. | Comp. F | Proba. |
|------------------------|------|----------|--------|---------|-------------|
| Product | 5 | 883,87 | 176,77 | 0,45 | 0,8105 |
| Judge | 12 | 67745,98 | 5645,5 | 14,47 | <0,0001 *** |
| Interaction 1-2 | 60 | 39974,63 | 666,24 | 1,71 | 0,0046 ** |

Um sich diese Aussagen visuell zu verdeutlichen, wird die Gesamtbewertung der einzelnen Säfte, des Attributes „süßlich im Geruch“ durch die Abbildung 40 dargestellt. Es ist klar zu sehen, dass das gesamte Panel durcheinander bewertet. So nutzen Prüfer „Pau“ und „Spl“ die oberen Skalenbereiche und z.B. „Tob“ die unteren Skalenbereiche. Es ist kein eindeutiges Fließen mit dem Gruppendurchschnitt zu erkennen.

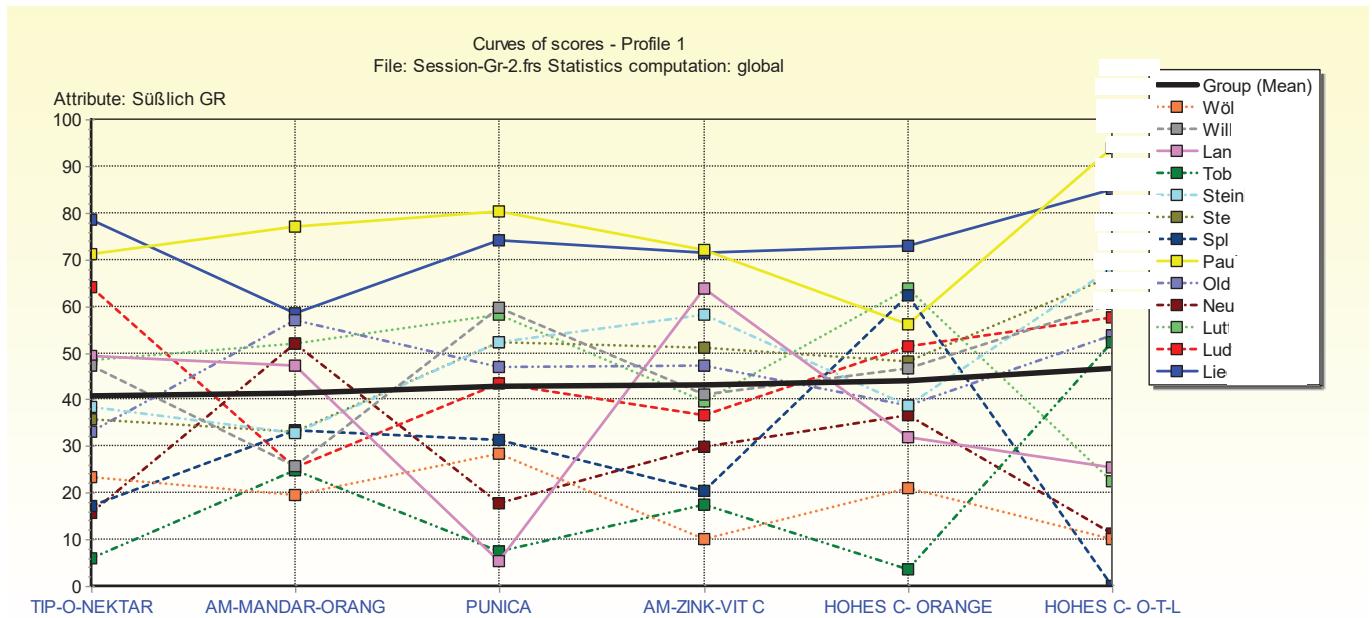


Abbildung 40: Durchschnittsbewertung des Attributes „süßlich im Geruch“

Als Vergleich wird, so wie bei Gruppe 1, der visuelle farbliche Aspekt genommen. In der Abbildung 41 ist eine nahezu ideale Bewertung zu sehen, die mit dem Gruppendurchschnitt einen Fluss ergibt. Lediglich geringe Abweichungen, vor allem bei dem Saft „Hohes-C O-T-L“ sind zu verzeichnen.

Im Vergleich zur Gruppe 1 schneidet die Gruppe 2 in dem Attribut „gelb-orange“ eindeutig besser ab (optischer Vergleich Abbildung 37 und 41).

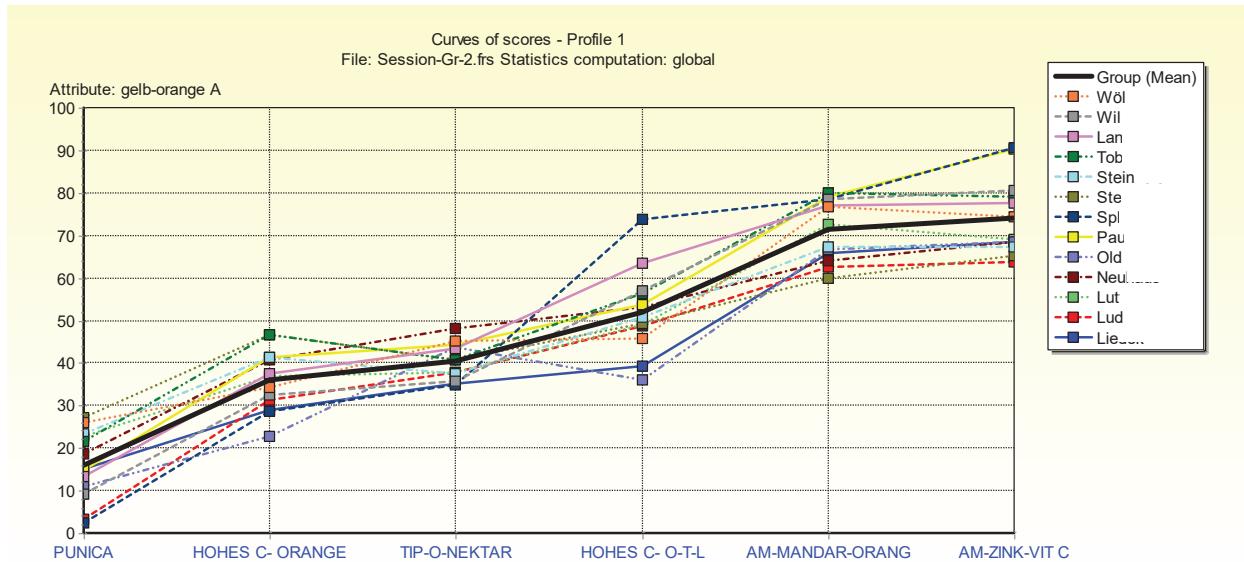


Abbildung 41: Durchschnittsbewertungen des Attributes „gelb-orange Aussehen“

In der Abbildung 42 ist ein PCA-Plot mit 82,8% der Werte dargestellt, woran deutlich zu erkennen ist, dass auch hier die Y-Achse als süß-sauer-Vektor interpretiert werden kann. Die X-Achse kann dahingegen als „Mundgefühl“-Achse bzw. stück – homogen bzw. viskos – flüssig interpretiert werden. Auch hier wird festgestellt, dass sich das Produkt „Amecke Zink plus Vitamin C“ am wenigsten von den anderen Produkten unterscheidet. Der „Amecke Mandarine-Orange“-Saft ist so wie in der Gruppe 1, besonders stückig im Mundgefühl, aber auch viskos. Ebenfalls, wie bei Gruppe 1, hat die Gruppe 2 den „Tip O-Nektar“ und „Punica“ als besonders süß empfunden. Der Saft „Punica“ ist zusätzlich besonders hell in der Farbintensität und besitzt einen künstlichen Geschmack. Das Produkt „Hohes-C O-T-L“ ist besonders zitronenartig und sauer im Geschmack, zusätzlich wurde dieser Saft als adstringierend, säuerlich im Geruch, aber auch bitter empfunden. Ebenfalls bitter und säuerlich im Geruch, sowie trüb, ist der „Hohes-C Orange“. Der „Amecke Zink plus Vitamin C“-Saft ist intensiv gelb-orange und besitzt einen ausgeprägten orangenartigen Geschmack. So wie bei Gruppe 1 ist auch bei Gruppe 2 erkennbar, dass das Attribut „süßlich Geruch“ sehr nahe am Schnittpunkt der zwei Achsen liegt, sodass auch hier die Aussage getroffen werden kann, dass das Panel Schwierigkeiten hat, die Produkte eindeutig anhand des Attributes zu differenzieren.

Die Gruppe 2 hat die Säfte „Hohes-C O-T-L“ und „Amecke Mandarine-Orange“, ähnlich wie die Gruppe 1, unterschiedlich bewertet. Dies ist daran zu erkennen, dass unter anderem die zwei Produkte am weitesten voneinander entfernt sind. Ebenfalls sind die zwei genannten Säfte unterschiedlich zu dem Saft „Punica“.

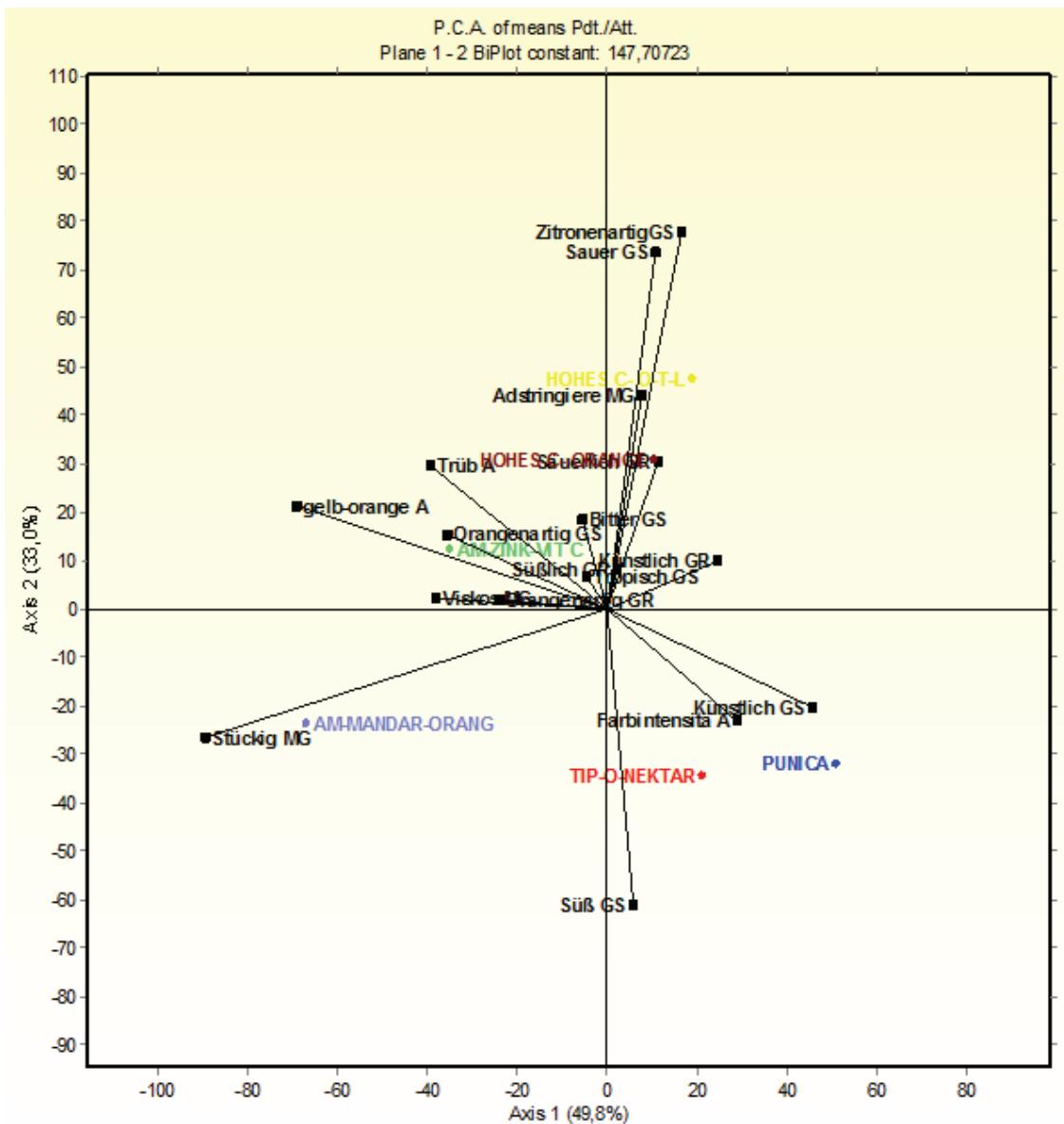


Abbildung 42: PCA-Plot Gruppe 2

Nach der DIN EN ISO 11132 kann die Aussage getroffen werden, dass das Panel die Produkte anhand von 17 Attributen signifikant unterscheiden musste. Das Panel hat die Produkte mit 16 Attributen höchstsignifikant unterschieden. Aus diesem Grund hat auch dieses Panel eine sehr gute Leistungsfähigkeit. Vergleicht man die beiden Panels so kann gesagt werden, dass Gruppe 2 94,1% der Merkmalseigenschaften höchstsignifikant unterschieden. Gruppe 1 hat weniger Attri-

bute gehabt und lediglich 12 mit einer Höchstsignifikanz, also 85,7% unterschieden. Die Gruppe 2 verfügt 14 Interaktionen zwischen Panel und Prüfer, somit sind nur drei Attribute (viskos Mundgefühl, trüb Aussehen, gelb-orange Aussehen) nicht signifikant in der Interaktion. Auch hier muss das Panel weitere Schulungen erhalten, damit die Wechselwirkung keine Signifikanz erreicht.

Tabelle 17: Leistungsprüfung der Einzelpersonen Gruppe 2

| | individuelle syst. Abweichung | S.D. der Syst. Abweichungsterme von einzelnen Prüfern | Mittelwert der Standardabweichung |
|--------------|--------------------------------------|--|--|
| Lie | 19,98 | 7,28 | 16,90 |
| Lud | -16,29 | 15,08 | 24,03 |
| Lutt | -24,46 | 8,76 | 21,18 |
| Neu | -2,35 | 10,53 | 23,08 |
| Old | -7,24 | 7,49 | 19,69 |
| Pau | 17,76 | 4,19 | 21,43 |
| Spl | -9,59 | 20,9 | 29,01 |
| Ste | 8,87 | 15,43 | 18,40 |
| Stein | -3,96 | 14,35 | 24,59 |
| Tob | -0,63 | 11,97 | 23,62 |
| Lank | 16,48 | 326 | 24,56 |
| Will | -4,35 | 14,16 | 22,64 |
| Wöl | 5,76 | 7,44 | 17,27 |

Nach der DIN EN ISO 11132 wurde die Tabelle 17 berechnet und zeigt die individuelle systematische Abweichung und die Standardabweichung der systematischen Abweichungsterme der Einzelprüfer. Die Prüfer „Lutt“ und „Lud“ haben eine sehr niedrige individuelle systematische Abweichung, was bedeutet, dass diese zwei Prüfer niedrige Zahlenwerte benutzt haben, um die Produkte anhand des orangigen Geschmackes zu bewerten. Dahingegen haben „Lie“, „Pau“ und „Lank“ einen höheren Wert, sodass davon auszugehen ist, dass die Prüfer höhere Werte zur Bewertung benutzt haben. Die Konsistenz der Probanden der Gruppe 2 wird anhand der Standardabweichung der systematischen Abweichungsterme dargestellt. Der Prüfer „Ste“ hat mit 15,43 die höchste Abweichung. Anhand der errechneten Daten beträgt seine niedrigste systematische Abweichung vom Panelmittelwert -22,41 und die höchste 27,44. Die Prüfer „Lud“, „Stein“ und „Will“ haben ebenfalls eine höhere Standardabweichung. Im Vergleich zur Gruppe 1 besitzt die Gruppe 2 eine niedrigere maximale Standardabweichung. Somit besitzt Prüfer „Ste“ die niedrigste Konsistenz im Panel der Gruppe 2, jedoch eine höhere Konsistenz als der Prüfer

„Heu“ aus der Gruppe 1. Mit einer Standardabweichung von 3,26 besitzt der Prüfer „Lank“ die beste Konsistenz und benutzt höhere Werte zur Bewertung der Proben (individuelle syst. Abweichung von 16,48), als der Panelmittelwert. Die höchste Abweichung vom Panelmittelwert beträgt beim Prüfer „Lank“ 21,59 und die niedrigste Abweichung 11,38. Dadurch ergibt sich eine niedrige Standardabweichung. Dies bedeutet, dass dieser Prüfer die beste Konsistenz des Panels und somit eine gute Leistungsfähigkeit besitzt, um die Produkte anhand des Attributes „oranger Geschmack“ zu beurteilen. Vergleicht man den Prüfer „Lank“ mit dem Prüfer mit der besten Konsistenz aus der Gruppe 1, so stellt sich heraus, dass der Prüfer „Lank“ insgesamt die beste Konsistenz besitzt.

Die beste Wiederholbarkeit besitzen die Prüfer „Lie“ (16,90) und „Wöl“ (17,27). Im Gegensatz zu der Gruppe 1 ist diese „beste“ Wiederholbarkeit schlechter als die von „Erm“ (7,38) und „Gro“ (9,79). Jedoch ist hier der höchste Mittelwert der Standardabweichung von dem Prüfer „Spl“ mit 29,01 zu verzeichnen. Somit besitzt dieser Prüfer die schlechteste Wiederholbarkeit, jedoch eine bessere als die Prüfer „Kah“ (38,31), „Jar“ (34,64), „Grz“ (34,44), „Han“ (33,08) und „Lan“ (32,15) der Gruppe 1.

Durch das Betrachten der Rohdaten wird deutlich, dass Prüfer „Spl“ sehr starke Schwierigkeiten aufweist. Dieser Prüfer weißt die höchsten Werte in neun Attributen auf. Besonders gravierend ist da Arbeiten mit dem Attribut „zitronenartig im Geschmack“ (40,73) und „süß im Geschmack“ (38,05). Dieser Prüfer muss eine intensive Schulung durchlaufen oder aus dem Panel herausgenommen werden, um eine bessere Panelkonsistenz zu erlangen. Außerdem ist zu erkennen, dass das gesamte Panel im „stückig Mundgefühl“ eine hohe Standardabweichung zeigt (niedrigste Standardabweichung: 19,54, höchste Standardabweichung: 39,9). Prüfer „Lie“ besitzt fünf niedrigste Werte, „Wöl“ folgt mit vier niedrigsten Werten.

Im Vergleich kann die Aussage getroffen werden, dass die Gruppe 1 fünf Prüfer mit einem Wert über 30 und zwei Prüfer mit einem Wert unter 10 besitzt. Die Gruppe 2 hat keinen einzigen Prüfer der einen Mittelwert höher als 30 aufweist, aber auch keinen Prüfer der unter 10 liegt. Nichtsdestotrotz, ist im Vergleich der Werte zu erkennen, dass die Gruppe 2 mit einem Mittelwert von 22,03 eine bessere Wiederholbarkeit im Ganzen aufweist, als Gruppe 1 mit 25,87.

4.3 Preference Mapping

4.3.1 Internal Preference Mapping

Das Internal pref.map. wird mit Hilfe der hedonischen Daten des Konsumentenpanels durchgeführt. Durch diese Auswertungsmethode werden die Produkte als blaue Punkte und die Probanden als rote Punkte dargestellt. Je mehr Probanden um das Produkt liegen, desto beliebter ist der Saft.

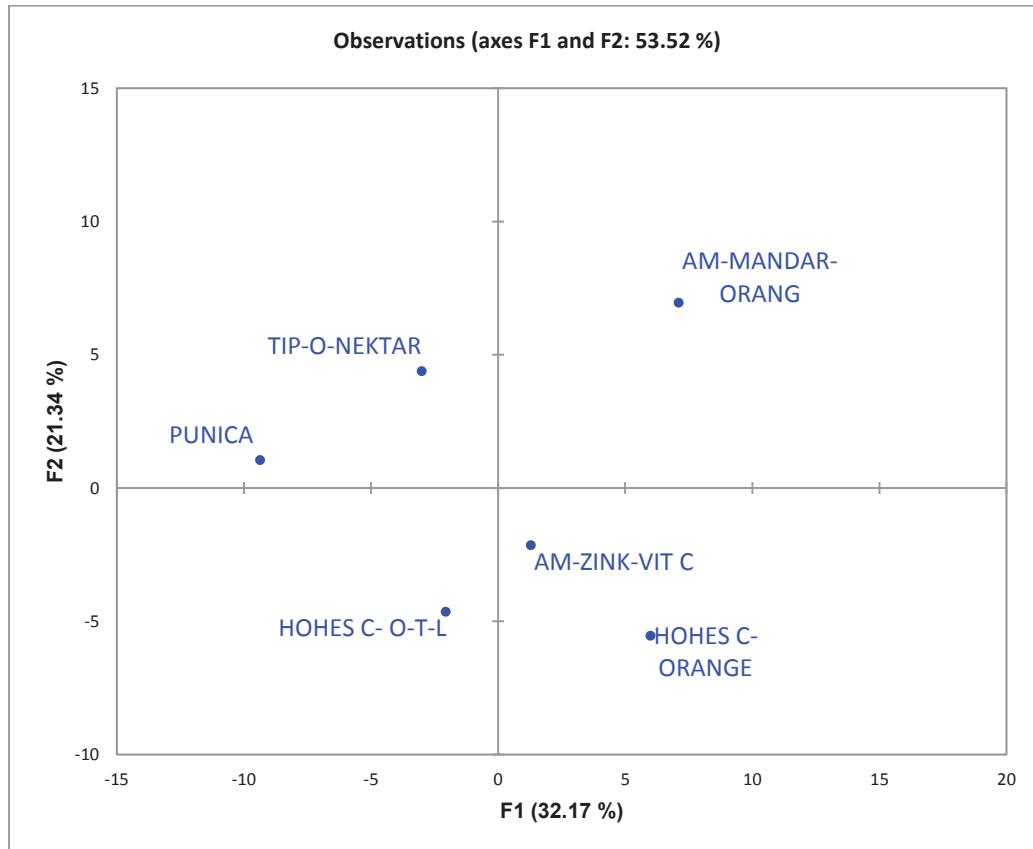


Abbildung 43: Platzierung der Säfte durch das Konsumentenpanel

In der Abbildung 43 sind die einzelnen Produkte dargestellt, die durch die Probanden des Konsumentenpanels platziert wurden. Hier ist zu erkennen, welche Produkte ähnlich bewertet wurden („Punica“ und „Tip-O-Nektar“) und welche Produkte z.B. stark unterschiedlich sind („Hohes C O-T-L“ und „AM-Mandarine-Orange“).

Im nächsten Schritt werden die Prüfer miteingefügt. Dieses stellt die Abbildung 44 dar.

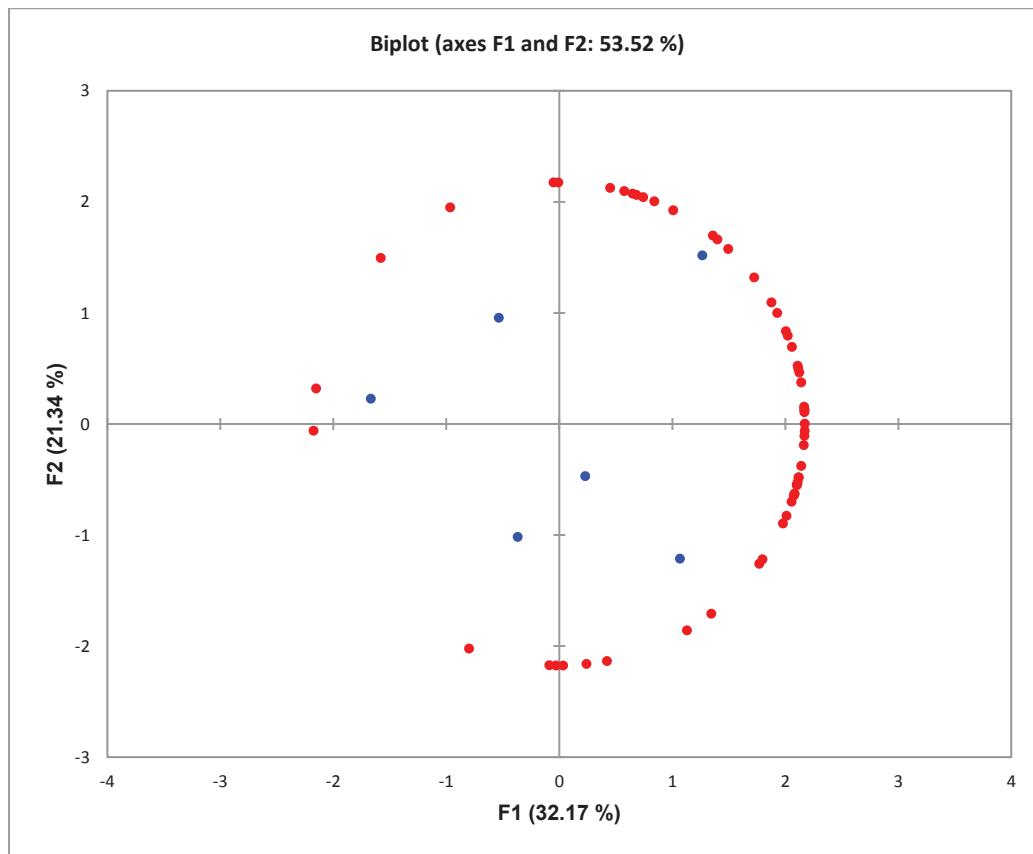


Abbildung 44: Produktpräferenz durch das Konsumentenpanel

Die Abbildung 44 kann interpretiert werden, indem die blauen Punkte die Produkte aus der Abbildung 43 darstellen und die roten Punkte jeden der 98 Prüfer zeigen. Je mehr Probanden um ein Produkt liegen, desto beliebter ist das Produkt. So wird deutlich, dass die Produkte „Punica“ und „Tip-O-Nektar“ die wenigste Beliebtheit erlangen. Das Produkt „Punica“ besitzt 2 von 98 Prüfern, die das Produkt gegenüber der anderen Produkte präferieren. „AM Mandarine-Orange“ und „Hohes C Orange“ besitzen eine sehr hohe Akzeptanz. Zudem ist auffällig, dass 32,17% der Daten in der ersten Dimension dargestellt werden können. Heymann und Lawless (2010) zitierten die Arbeit (How reliable are consumers? Comparison of sensory profiles from consumers and experts.) von Worch et al. (2009), dass ein untrainiertes Panel oft eine 15-24% Datenanalyse auf der ersten Dimension bietet, sodass mehrere Dimensionen dargestellt werden müssen. In diesem Fall besitzt das ungeschulte Panel eine deutlich höhere Darstellung. Dies hängt von der gleichmäßigen Punktevergabe des Panels ab. Je stimmiger das Panel ihre Analyse betreibt, desto besser

ist die Datendarstellung (Heymann and Lawless, 2010). Aus der Abbildung 44 geht hervor, dass viele Konsumenten an der X-Achse im Bereich +2 platziert sind, aber dass hier keiner der Säfte liegt. Dies zeigt, dass viele Prüfer eine bestimmte Präferenz aufweisen, die keiner der getesteten Säfte erfüllt. Hier ist also noch Bedarf vorhanden. Die Aufgabe der Produktentwicklung wäre nun, so ein Produkt zu entwickeln.

4.3.2 External Preference Mapping

Für die Auswertung des Preference Mapping wurde im ersten Schritt mit Hilfe der Daten des geschulten Panels eine „Principal Component Analysis“ durchgeführt. Das Ergebnis wird also anhand von PCA-Plots dargestellt.

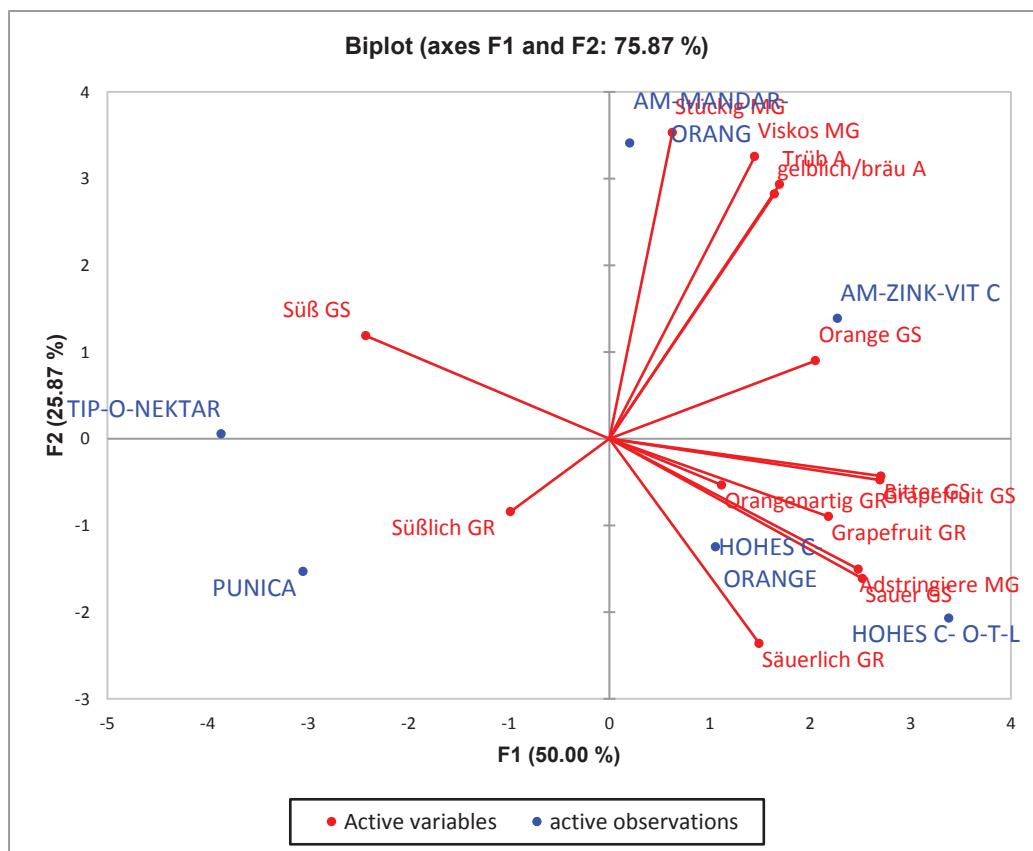


Abbildung 45: PCA-Bi-Plot des geschulten Panels Gruppe 1 anhand des Internal Preference Mapping

In der Abbildung 45 ist ein Biplot dargestellt, wo die Produkte und die Produkteigenschaften angezeigt sind. Es ist zu erkennen, dass „TIP-O-Nektar“ und „Punica“ als besonders süß im Geschmack empfunden werden. Der Vektor des Attributes „süßlich Geruch“ ist sehr kurz. Dies ist

die Folge daraus, dass dieses Attribut nicht gleichmäßig vom Panel zur Bewertung benutzt wurde (siehe Kapitel 5.2). Die Gruppe 1 konnte somit die Produkte deutlich voneinander trennen.

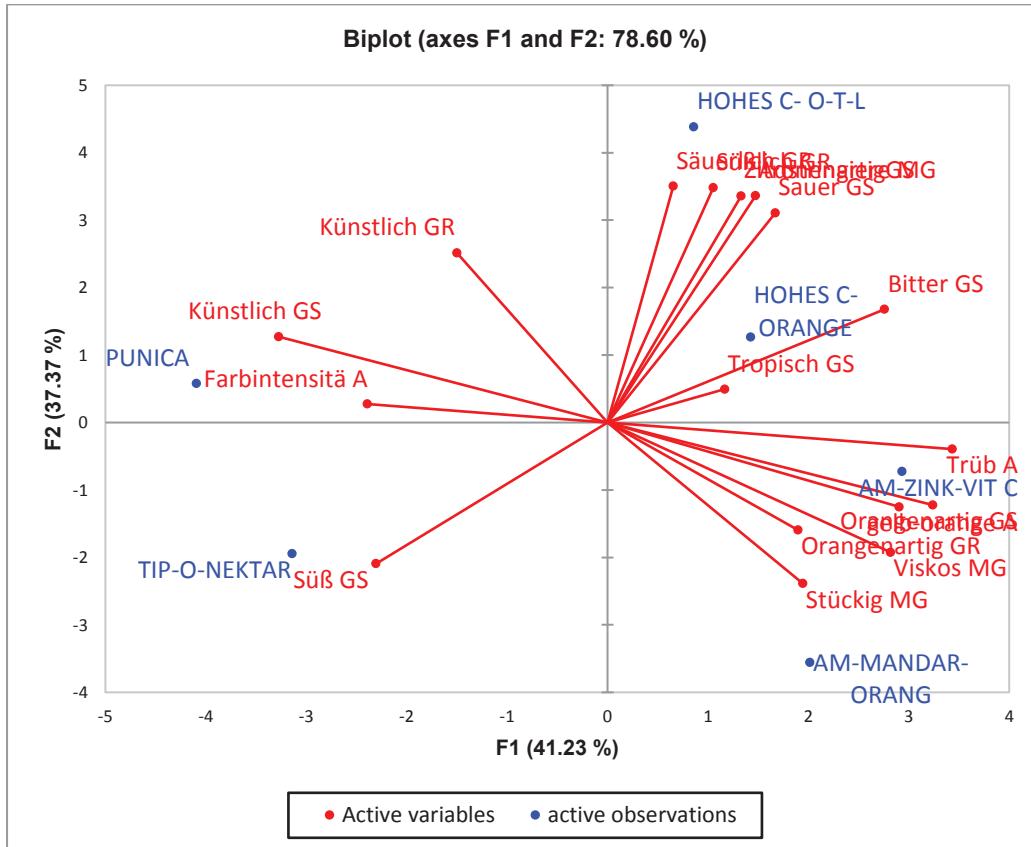


Abbildung 46: PCA-Bi-Plot des geschulten Panels Gruppe 2 anhand des Internal Preference Mapping

In der Abbildung 46, wird das Ergebnis der PCA der Gruppe 2 dargestellt. Die Gruppe 2 hat ebenfalls wie die Gruppe 1 die Produkte „TIP-O-Nektar“ und „Punica“ als süß im Geschmack empfunden. Zusätzlich wurde „Punica“ als künstlich im Geschmack und Geruch und sehr intensiv in der Farbintensität wahrgenommen. Beide Gruppen haben die Produkte „Punica“ und „AM-Zink-Vit C“ am unterschiedlichsten voneinander wahrgenommen. Dies wird durch die Entfernung der zwei Produkte auf dem Biplot deutlich. Die Gruppen sind sich ebenso einig, dass die zwei „Amecke“-Produkte besonders stückig, viskos und trüb sind, wobei das „AM-Zink-Vit C“ besonders orangig im Geschmack ist. Die „Hohes C“-Produkte sind bitter, sauer im Geschmack und besitzen einen säuerlichen Geruch. Beide Panels sind sehr ähnlich in ihrer Bewertung.

Im weiteren Schritt wurden die Beliebtheits-Daten des ungeschulten Panels mit für die Bewertung herangezogen und das PREFMAP durchgeführt.

In der Abbildung 47 ist die Preference Map zu sehen, die aus den Daten der Konsumenten und des geschulten Panels der Gruppe 1 entstanden ist.

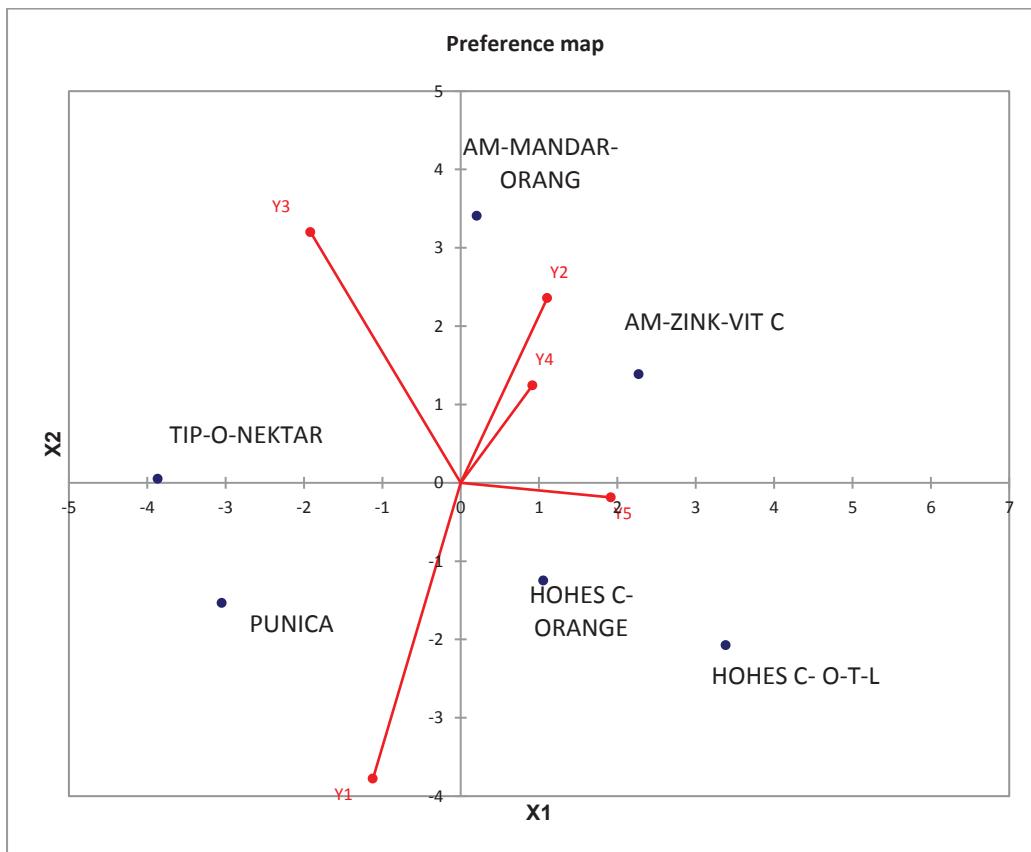


Abbildung 47: Darstellung der Konsumentengruppen anhand einer Preference Map

Die Y-Vektoren stellen die Personen-Gruppen dar, die ein Produkt bevorzugen. Somit stellt der Vektor „Y1“ die Personengruppe dar, die Punica bevorzugen. Um herauszufinden, wie viel Prozent der Konsumenten zufrieden mit den Produkten sind, wird ein sogenannter „Countar-Plot“ projiziert. Die Preference Map kann nun mit dem Counter Plot in Verbindung gebracht werden. Somit wird deutlich, wie viele Anteile der Konsumenten die Säfte bevorzugen und zusätzlich kann anhand einer Präferenz-Tabelle dargestellt werden, in welche Rangfolge die weiteren Säfte durch die Konsumentengruppen gebracht werden.

Die Abbildung 48 zeigt, dass 20-40% der Probanden das Produkt „Punica“ bevorzugen.

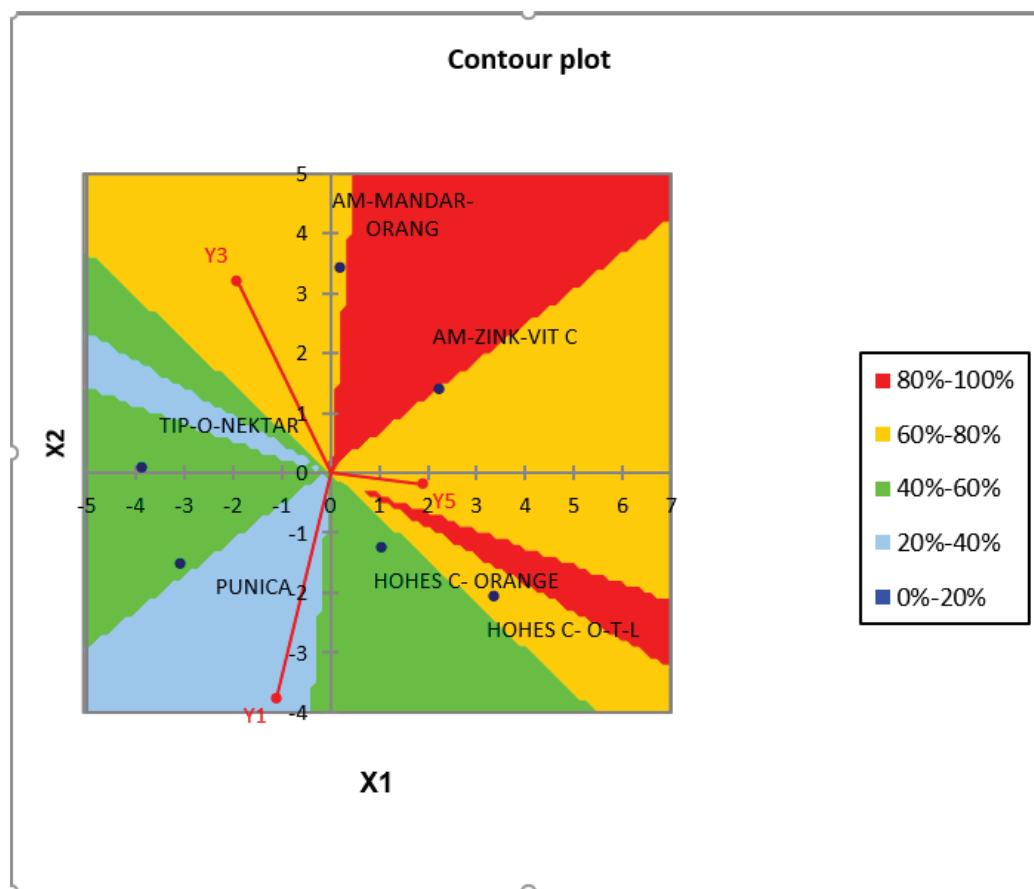


Abbildung 48: Countour Plot

„Hohes C-Orange“ wird von 40-60% als gut empfunden und „AM-Zink-Vit C“ liegt im Grenzbereich von 60-80% und 80-100%. Die Gruppen Y2 und Y4 sind im Counter-Plot nicht zu sehen, da diese im roten Bereich, also 80-100%-Bereich liegen. Zur weiteren Verdeutlichung der Graphik, wird die Tabelle 18 herangezogen.

Tabelle 18: Rangfolge der Beliebtheit der Gruppe 1

| Y1 | Y2 | Y3 | Y4 | Y5 |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| AM-MANDAR-ORANG | PUNICA | HOHES C- O-T-L | PUNICA | TIP-O-NEKTAR |
| AM-ZINK-VIT C | TIP-O-NEKTAR | HOHES C- ORANGE | TIP-O-NEKTAR | PUNICA |
| HOHES C- ORANGE | HOHES C- ORANGE | AM-ZINK-VIT C | HOHES C- ORANGE | AM-MANDAR-ORANG |
| HOHES C- O-T-L | HOHES C- O-T-L | PUNICA | HOHES C- O-T-L | HOHES C- ORANGE |
| TIP-O-NEKTAR | AM-ZINK-VIT C | TIP-O-NEKTAR | AM-ZINK-VIT C | AM-ZINK-VIT C |
| PUNICA | AM-MANDAR-ORANG | AM-MANDAR-ORANG | AM-MANDAR-ORANG | HOHES C- O-T-L |

In der Tabelle 18 ist zu sehen, dass das Produkt „AM-Mandarine-Orange“ von den Gruppen Y2, Y3 und Y4 bevorzugt wird. Die Interpretation der Tabelle erfolgt in aufsteigender Reihenfolge. Die Gruppe Y1(20-40%), die das Produkt „Punica“ bevorzugt, lehnt das Produkt AM-Mandarine-Orange ab. „Punica“ wird von Y2, Y4 und Y5 abgelehnt, bzw. findet nur wenig gefallen. Der Countour-Plot mit den deskriptiven Daten der Gruppe 2 wird nicht abgebildet, da dieser eine gespiegelte Variante des Countour-Plots mit den deskriptiven Daten der Gruppe 1 (Abbildung 47) darstellt. Es ist auch zu sehen, dass Produkte wie „AM-Zink plus Vitamin C“, „Hohes C Orange“ keine eindeutige Präferenz aber auch keine Abneigung erhalten. Zudem kann die Aussage erfolgen, dass die Produkte die nahe der 80-100% Marke liegen, einem idealen Produkt entsprechen. „Amecke Mandarine-Orange“ liegt, so wie „Amecke Zink plus Vitamin C“, nah am Grenzbereich. Das wird auch aus der Tabelle 18 deutlich, dass „Amecke Mandarine-Orange“ drei von fünf Mal als präferierendes Produkt ausgezeichnet ist. „Amecke Zink plus Vitamin C“ liegt drei von fünf Mal auf dem zweiten Platz der Präferenz.

5. Diskussion

Mit dieser Studienarbeit muss die Frage geklärt werden, ob ein ungeschultes Panel ähnlich oder gleich vertrauliche und zuverlässige Daten liefert, wie ein geschultes Panel. Ebenso soll der Nutzen bzw. die Handhabung der DIN EN ISO 11132 dargestellt werden. Um diese Fragen zu beantworten, wurde eine Profilierung mit zwei Panels, die in zwei Sitzungen und einem Testdurchlauf auf das Produkt geschult wurden, und ein Verbrauchertest mit ungeschulten Prüfern, durchgeführt.

Die Beliebtheitsfragen des Verbrauchertestes wurden durch die ANOVA ausgewertet. Des Weiteren sind die Beliebtheitsfragen mit in die Auswertung der Penalty-Analyse, sowie der CATA-Fragen eingegangen. Anhand der ANOVA konnte festgestellt werden, dass der Saft „Hohes C Orange“ am besten bewertet wurde. Im Gegensatz dazu schneidet „Punica“ in der Beliebtheit am schlechtesten ab. Eine plausible Erklärung könnte sein, dass der ungeschulte Prüfer sich wohlmöglich einen idealen Saft wie der „Hohes C Orange“ vorstellt. Da dieser Saft aus 100% Orangensaft besteht und keine weiteren Zusätze bzw. andere Fruchtbeimischungen verfügt, wird dieser Saft eher mit dem persönlich „Ideal“-Orangensaft verglichen und somit als positiv befunden. Als Kontrast ist der „Punica“-Saft 100%-Orangensaft und weicht somit sensorisch stark von dem „Hohes C Orange“ ab. Dadurch verlor der „Punica“-Saft stark an Beliebtheit. In der Tabelle 8: „Post hoc Test, Faktor „Produkt“ mit L.S.D.: 0,57“, wird diese Aussage noch einmal bekräftigt. Die Produkte werden in verschiedene Kategorien eingeteilt, sodass das Produkt „Hohes C Orange“ der ersten Gruppe, der Gruppe A und „Punica“ in die letzte Gruppe, die Gruppe C, zugewiesen wird.

Für den Verbrauchertest wurden JAR-Fragen erstellt, mit deren Hilfe herausgefunden werden sollte, in welche Richtung das Produkt optimiert werden darf und welches Produkt dem Ideal entspricht. Dafür wurden die Attribute „Farbintensität“, „orangiger Geschmack“ und „bitterer Geschmack“ ausgewählt. Bei der Wahl der Attribute ist es sehr wichtig, Attribute zu wählen, die der Proband leicht versteht. Außerdem dürfen keine gegensätzlichen Attribute bzw. Attribute, die eine gleiche Aussage treffen, gewählt werden. Mit Hilfe der ANOVA-Auswertung konnte festgestellt werden, dass die Säfte anhand der Attribute mit einer Höchstsignifikanz unterschieden werden konnte. Das zeigt, dass das ungeschulte Panel in der Lage ist die Produkte anhand von der Farbintensität, dem orangigen und bitteren Geschmack, zu differenzieren. Jedoch konnte auch anhand der Höchstsignifikanz der Prüfer schnell ermittelt werden, dass die Prüfer nicht konstant bewerten, was bei einem ungeschulten Panel zu erwarten ist. Der Grund für die Höchstsignifikanz der Prüfer liegt auf der Hand: die ungeschulten Prüfer sind es nicht gewohnt, mit

dem Produkt und den Attributen zu arbeiten. Die Aufgabe des Prüfers bei der Verkostung und Prüfung der JAR-Kategorien bestand darin, sich rein auf die drei Attribute zu konzentrieren. Da der Konsumententest jedoch nicht in einem Raum nach ISO 8589, sondern vor der Mensa der Hochschule Neubrandenburg durchgeführt wurde, bestanden sämtliche Ablenkungsfaktoren wie Lärm, Fremdgerüche und wenig bis keine Diskretion. Eine Möglichkeit die Prüfer-Signifikanz zu reduzieren, wäre mit einer Prüfung in einem Raum nach ISO 8589 eventuell besser gegeben. Ein weiterer Grund kann die persönliche Präferenz des Verbrauchers betreffen. Da es kein Ideal-Produkt gab, musste jeder Proband sich sein persönliches Ideal-Produkt vorstellen und anhand dieser persönlichen Präferenz bewerten. Wäre ein Ideal-Produkt vorgegeben, so könnte sich der Verbraucher anhand diesem orientieren.

Mit Hilfe der Mean Drop- und Penalty-Analyse konnte nun herausgefunden werden, welche der drei Eigenschaften dem Konsumenten am wichtigsten ist bzw. wo es die meisten Abzüge in der Beliebtheit gibt. Diese Analyse-Methode zeigte, dass bei fünf von sechs Säften ein zu milder Geschmack zu einer schlechteren Beliebtheit führen wird. Der Saft „Punica“ wies den höchsten Penalty von 3,5 Beliebtheitspunkten auf, bei einem schwächeren Geschmack. Dieses Ergebnis ist nicht überraschend, da das Produkt im Vergleich zu den anderen Produkten einen sehr milden Orangengeschmack aufweist (näheres dazu später). Somit ist bei einem Verlust des milden Orangengeschmacks mit einem starken Abfall der Beliebtheit als Konsequenz zu rechnen. Daraus kann rückgeschlossen werden, dass ein verbesserter Orangengeschmack die Produktbeliebtheit deutlich anheben würde. Das Produkt „Amecke Mandarine“ besitzt als einziges Produkt einen hohen Penalty (1,925) in der Farbintensität und dem bitteren Geschmack. Somit würde ein zu kräftiger Orangenton, der von 50% der Prüfer als zu viel bewertet wurde, und ein ausgeprägte bittere Note dem Produkt schaden. Durch die Mandarine bekommt das Produkt eine gewisse Süße und einen angenehmen Geschmack, der den Verbrauchern gefällt. Somit akzeptiert der Verbraucher keine bittere Note in dem Produkt. Mit diesen Ergebnissen kann die Produktentwicklung sehen, wie stark die Merkmalseigenschaften empfunden werden und inwieweit die Produkte optimiert werden müssen. Dieses Verfahren eignet sich hervorragend, wenn diverse Zutaten in der Rezeptur verändert werden, z.B. wenn Zucker mit Süßstoff ausgetauscht wird.

Für die JAR-Fragen, aber auch für CATA und die Profilierung werden PCA-Bi-Plots dargestellt. Mit Hilfe dieser Plots kann verstanden werden, welche Produkteigenschaften den Produkten zugeordnet wurden. In dem JAR-Biplot ging es darum, welche Attribute besonders farbintensiv, eine ausgeprägte bittere Note, sowie einen intensiven orangigen Geschmack aufweisen. Es fiel auf, dass „Amecke Mandarine“ eine besonders hohe Farbintensität aufweist, wohingegen „Puni-

ca“ gelblich ist. Im Vergleich zu der CATA-Auswertung war der Saft „Amecke Zink plus Vitamin C“ mit einer Orangenfarbe interpretiert worden und ebenfalls „Punica“, sowie „Tip O-Nektar“. Den besten orangigen Geschmack wies in der JAR-Methode der „Hohes C Orange“, sowie „Amecke Mandarine“ auf, „Punica“ und „Tip O-Nektar“ dahingegen den geringsten. Durch die CATA-Methode konnte ein ähnliches Ergebnis produziert werden, wobei der Fruchtsaft „Hohes C O-T-L“ auch eher einen milden orangigen Geschmack aufweist. Dieses Ergebnis kann dadurch entstanden sein, dass der Saft „Hohes C O-T-L“ kein reiner Orangensaft ist, sondern Beimischungen von Traube und Limette besitzt. Aus diesem Grund wurde der Saft, der über einen intensiven Geschmack verfügt, nicht als ein klarer Orangensaft bzw. der persönliche Ideal-Orangensaft identifiziert. Über eine ausgeprägte bittere Note verfügte in der JAR-Methode der „Hohes C O-T-L“, wohingegen in der CATA-Methode der „Amecke Zink plus Vitamin C“ dominanter war. Beide Methoden wiesen eine schwache bittere Note in dem Produkt „Tip O-Nektar“ auf. Das Problem des Vergleiches der zwei Methoden besteht darin, dass diese unterschiedlich im Aufbau sind. In der JAR-Methode erfolgt eine Intensitätsmessung, wo der Proband entscheiden muss, ob das Produkt „zu intensiv“, „genau richtig“ oder „zu schwach“ ist. Die CATA-Methode beschränkt sich darauf, ob ein Attribut vorhanden ist oder nicht. Die Intensität spielt hier keine Rolle. Trifft eine Merkmaleigenschaft auf das Produkt zu, wird diese angekreuzt. Des Weiteren muss der Proband in der JAR-Methode eine Entscheidung treffen, wohingegen die CATA-Methode darauf vertraut, dass der Proband die Produkteigenschaften identifizieren kann. Es besteht die Wahrscheinlichkeit, dass die Probanden den CATA-Bogen nicht vollständig ausfüllen. Gründe dafür sind, wenn die Prüfer mit vorherigen Prüfmethoden überlassen wurden, zu viele Proben und Attribute vorhanden sind und die Ablenkung durch die Umgebungsgegebenheiten. In der Auswertung der CATA-Fragebögen ist es oft vorgekommen, dass einige Produkte nicht bewertet, also mit „0“-gekennzeichnet wurden, einige Bögen wurden sehr ausführlich behandelt und andere wiederum nur grob angekreuzt. Das PCA-Modell der JAR-Methode wurde mit der Software „Fizz“ und der Anova-Methode produziert. Die Anova fasst die Ergebnisse gemittelt zusammen, d.h. wenn x-hohe Bewertungen abgegeben wurden und genauso viele x-niedrige Bewertungen, kommt eine mittlere Bewertung zum Vorschein (Ellert, 2006). Wohingegen die CATA-Methode mit der Software „Xlstat“ produziert wurde und sich auf die Anzahl der gewählten Attribute, sowie der Beliebtheit, beschränkt. Ein weiterer Unterschied ist, dass CATA auch hedonische Begriffe wie „harmonisch“ und „lecker“ erlaubt. Es wäre auch sinnentleert, wenn die Intensität nach „harmonisch“ und „lecker“ gefragt werden würde. Dies würde zu keinen signifikanten Ergebnissen führen und wäre der Produktentwicklung keine Hilfe.

Die CATA-Ergebnisse zeigten im ersten Moment, wie viele Nennungen der Attribute (in Prozent) durchgeführt wurden. Es konnte festgestellt werden, dass Attribute, die leicht mit dem Produkt in Beziehung gebracht werden konnten, wie z.B. gelb, orangefarbend, süß, sauer, künstlich, orangiger Geschmack und lecker, die meisten Nennungen aufweisen. Nennungen wie „bräunlich“ oder „harmonisch“ sind für den Prüfer schwer zu identifizieren. Es fehlt die Definition zu harmonisch, aus diesem Grund ist es schwer für den Verbraucher mit diesem Begriff zu arbeiten.

Des Weiteren ist es schwer einen Orangensaft mit dem Attribut bräunlich zu belegen, da eine herkömmliche Beschreibung der Farbe eines Orangensaftes „orange“ oder „gelb“ ist. Außerdem wurde ermittelt, dass die Attribute süß, erfrischend, harmonisch, lecker und orangenartig für eine hohe Beliebtheit verantwortlich sind. Also kann davon ausgegangen werden, dass die Produkte, die diese Attribute beinhalten, besonders beliebt sind. Fremdgeschmack und künstlich stellen die Unbeliebtheit dar. Die Produkte „TIP O-Nektar“, „Hohes C 100% Orange“ und „Amecke Mandarine-Orange“ besitzen eine hohe Beliebtheit. Dahingegen können sich „Hohes C O-T-L“ und „Punica“ keiner großen Beliebtheit erfreuen. Außerdem konnte festgestellt werden, dass ein harmonischer, leckerer und orangenartiger Saft als besonders beliebt interpretiert wurde. Diese Attribute sind den Probanden besonders wichtig. Zudem darf der Saft keinen Fremd- und künstlichen Geschmack aufweisen. Dieses Ergebnis erscheint logisch, denn ein künstlicher Orangensaft hat für die meisten Verbraucher eine negative Erscheinung. Nach wie vor gilt: Je näher an der Natur, desto besser. Ein künstlicher Geschmack wird direkt mit „ungesund“ verbunden. Ebenfalls ist ein Fehlaroma negativ in der Wahrnehmung veranlagt und wird oft mit etwas Verdorbenen in Verbindung gebracht (Welle, Datum unbekannt). Künstliche Säfte, die ein Fehlaroma aufweisen, werden als qualitativ minderwertig empfunden, da diese nicht dem persönlichen Idealbild, meist frischgepressten Orangensaft, entsprechen. Wobei eine künstliche Note schon bei einer gewöhnlichen Saftschorle mit ggf. Zuckerzusatz interpretiert wird, da diese sich von einem 100% Fruchtsaft stark differenziert. Des Weiteren werden Fruchtbeimischungen wie bei dem „Hohes C O-T-L“ direkt als Fremdaroma eingestuft, da der typische „orangige“ Geschmack durch Trauben- und Limettensaft überdeckt wird. Dahingegen erfolgt durch die Attribute „harmonisch“, „lecker“ und orangenartig eine positive Interpretation, die ein Gefühl von guter Qualität wiederspiegelt.

Die zwei geschulten Panels bewerteten die Produkte ähnlich und sind somit gut vergleichbar. Weiterhin bewiesen die Prüfer eine gute Unterscheidungsfähigkeit der Produkte anhand der Attribute. Lediglich das Attribut „süß im Geruch“ wies keinen signifikanten Unterschied auf. Das Panel hat dieses Attribut nicht verstanden bzw. wendete es falsch an. Ein möglicher Grund für

diese Diskrepanz besteht darin, dass es olfaktorisch nicht möglich ist eine Süße wahrzunehmen, da Zucker als Kohlenhydrat keinen Geruch besitzt. Lediglich thermisch behandelter Zucker, z.B. in Form von Zuckerwatte, besitzt Aromen. Dieses Beispiel wurde von beiden Gruppen als Referenz- bzw. Ankerpunkt gewählt. Dadurch dass Zuckerwatte nicht viele Gemeinsamkeiten mit Orangensaften aufweist, ist es schwer diese Geruchsnote herauszukristallisieren. Alle anderen Produkteigenschaften wurden, mit einigen Ausnahmen, höchstsignifikant unterschieden. Die Leistungen der einzelnen Prüfer waren unterschiedlich. In beiden Gruppen waren Prüfer mit einer guten Gesamtleistung, aber auch Prüfer, die entweder eine geringe Konsistenz, sowie eine hohe Spannweite der Bewertungen, in den drei Durchläufen aufwiesen. Dieses resultiert z.B. aus einer unzureichenden Schulung oder einem Verständnisproblem des Probanden bei den Attributen. Um diese Frage zu klären, ist oftmals ein Gespräch notwendig. Im Großen und Ganzen sind die Panels gut geschult. Jedoch gibt es einzelne Prüfer, die weitere Schulungen in diversen Attributen benötigen.

Mit Hilfe der Preference-Mapping-Methode können die durch die Konsumenten wahrgenommenen Produktmerkmale und deren Beliebtheitsbewertungen erklären, welche Attribute, aber auch Produkte, für eine besonders hohe Beliebtheit verantwortlich sind (van Kleef et al., 2005). In dieser Arbeit wurde herausgefunden, dass das Konsumentenpanel in fünf Präferenzgruppen eingeteilt werden kann. Das heißt es gibt Prüfer, die den Saft „Amecke Mandarine-Orange“ bevorzugen, jedoch „Punica“ ablehnen. Genauso gibt es umgekehrt Prüfer, die „Punica“ bevorzugen, jedoch nicht „Amecke Mandarine-Orange“. Der Countour-Plot gab zusätzlich die Information, dass 20-40% der Prüfer „Punica“ gut bewerten, jedoch 80-100% auch „Amecke Mandarine-Orange“. Ferner liegen „Amecke Mandarine-Orange“ und „Amecke Zink plus Vitamin C“ näher an einem für den Konsumenten idealen Produkt, als die anderen vier Säfte. Da das Panel aus 98 Personen bestand ist es zu verstehen, dass nicht alle Prüfer die gleiche Präferenz aufweisen. So mit können die 98 Prüfpersonen in fünf Präferenzgruppen eingeteilt werden. Werden die Ergebnisse mit den CATA-Ergebnissen verglichen so wird deutlich, dass „Amecke Mandarine-Orange“ als besonders lecker, orangig und harmonisch betitelt wurde. Diese Attribute sorgen für eine hohe Beliebtheit, was sich auch im Preference Mapping, vor allem dem Countour-Plot, wiederholt. „Punica“ wurde dahingegen als künstlich und gelblich bewertet, was mit einer schlechten Bevorzugung in Verbindung gebracht wurde. Meullenet et al. (2009) erläuterte ebenfalls, dass die CATA-Methode die Ergebnisse des Preference Mapping bestätigt bzw. die Aussagen nahezu identisch sind. Des Weiteren zitiert Meullenet et al. (2009) Lancaster und Foley (2007), dass das Nutzen der CATA-Methode eine geringe Einarbeitung bzw. Einweisung benötigt, sowie leicht und schnell durchzuführen ist.

Für den Vergleich des geschulten und ungeschulten Panels können hierfür die Abbildungen 32, 38 und 42 miteinander verglichen werden. Zur Erinnerung: Die Abbildung 32 stellt die Bewertung der Konsumenten (CATA) und die Abbildungen 38 und 42 die des geschulten Panels, der Gruppen 1 und 2, dar. Im Vergleich ist zu sehen, dass die Konsumenten das Produkt „Punica“ als künstlich, gelblich und zitronenartig wahrnehmen. Das geschulte Panel, vor allem Gruppe 2, hat dieses Produkt ebenfalls als künstlich und besonders farbintensiv, d.h. gelb, wahrgenommen. Ebenfalls durch das Attribut „süß“ wurde das Produkt „Tip-O-Nektar“ gleich bewertet. Sowohl das geschulte als auch das ungeschulte Panel haben für „Amecke-Zink-Vit C“ und „Hohes C O-T-L“ die Attribute grapefruitartig, orangefarbend und bitter benutzt. Zusätzlich wurde das „Hohes C O-T-L“ als sauer interpretiert. Ebenfalls bestand eine Korrelation bei dem Produkt „Amecke Mandarine- Orange“. Sowohl das geschulte als auch das ungeschulte Panel war sich einig, dass dieses Produkt orangenartig ist und eine dunkle orangene bis bräunliche Farbe ausstrahlt. Somit stellt sich heraus, dass ein ungeschultes Panel in der Lage ist ähnliche Werte zu erzielen. Der Unterschied zwischen den Abbildungen ist, dass die Abbildung 32 die Produkte und Attribute enger gestreut darstellt. Produkte wie z.B. das „Hohes C- Orange“ sind sehr nahe am Ursprung. Dies resultiert daraus, dass sich das Konsumentenpanel nicht so einig in der Bewertung war wie das geschulte Panel. Es wurden viele verschiedene Bewertungen durchgeführt und dadurch liegen Produkte und Attribute nahe am Ursprung. Außerdem wird durch das geschulte Panel das Produkt „Hohes C- Orange“ als säuerlich im Geruch interpretiert. Dieses Attribut besaß das ungeschulte Panel nicht, jedoch das Attribut sauer, was sich eher auf den Geschmack bezogen hat. Das Attribut sauer im Geschmack korrelierte bei dem geschulten- und ungeschulten Panel. Zu einem ähnlichen Ergebnis ist Ares et al. (2010) gekommen, als er mit Hilfe von zwei ungeschulten und einem geschulten Panel eine ähnliche Aufgabenstellung untersuchte. Hierbei ließ er Schokoladendesserts mit unterschiedlichen Kakao- und Zuckerkonzentrationen testen und mit Hilfe der CATA-, MFA- und Preference Mapping-Methode, sowie mit PCA-Bi-Plots auswerten. In seiner wissenschaftlichen Arbeit kam Ares et al. zu dem Schluss, dass die Konsumenten und das geschulte Panel stark miteinander korrelieren und somit ähnliche Ergebnisse hervorbringen. Ebenfalls stellt er klar, dass die CATA-Methode eine gute Alternative bietet, wenn ein geschultes Panel nicht vorhanden ist oder keine Zeit besteht, ein Panel zu trainieren. Ebenfalls zu einem ähnlichen Ergebnis ist McEwan et al. (2002) bei dem Leistungstest des Panels anhand einer Sorte Apfelsaft mit unterschiedlichen Konzentrationen an Fruktose und Glukose gekommen. Getestet wurden 14 Panels, wovon vier untrainierte und 10 trainierte waren. Herausgekommen ist, dass drei geschulte Panels sehr überdurchschnittlich abschnitten, ein geschultes Panel durchgefallen ist und die weiteren geschulten Panels im Mittelfeld lagen. Genau in diesem

Mittelfeld lag auch die Leistung des ungeschulten Panels. Letztendlich kommt es auf die Effektivität der Schulung an, wie gut ein Panel abschneidet. Sind die Schulungen intensiv und in regelmäßigen Abständen, so ist die Chance höher, dass dieses Panel auch eine gute Leistung erbringt.

Dadurch kann die Annahme getroffen werden, dass ein geschultes Panel sichere Werte liefert, jedoch auch ein ungeschultes Panel in der Lage ist, zuverlässige Aussagen zu treffen. Diese können aber nicht eins zu eins gleichgestellt werden. Dazu fehlt es dem ungeschulten Panel an Sicherheit und Übung. Nichtsdestotrotz ist ein Verbrauchertest günstiger und nicht so zeitintensiv wie eine Profilierung und somit ein gutes Mittel, um Informationen über ein Lebensmittel oder Präferenzen zu liefern. Letztendlich muss jeder Betrieb für sich selber entscheiden, welche Prüfmethode eingesetzt werden soll bzw. ob eine sichere Aussage durch ein geschultes Panel sinnvoll ist, oder ob die Aussage der Konsumenten ausreicht. Die Vorteile des gut trainierten, geschulten Panels liegen auf der Hand: Es ist in der Lage Unterschiede zu erkennen und diese auch zu beschreiben. Ein Konsumentenpanel liefert eher Akzeptanzdaten, wenn z.B. die Frage gestellt wird, wie gut das Produkt bei dem Verbraucher ankommt. Aber sie sind auch in der Lage komplexere Aussagen zu treffen und ähnlich wie ein geschultes Panel abzuschneiden. Wenn der Betrieb viel Wert auf hohe Qualität und genaues Arbeiten legt, sollte ein geschultes Panel gewählt werden. Soll lediglich ein Überblick geschaffen werden, ist ein Konsumentenpanel die richtige Wahl. Es muss lediglich darauf geachtet werden, dass genug Prüfer (mind. 60; Vorgabe der DIN 10974) vorhanden sind, um ein signifikantes Ergebnis zu erzielen. Ein weiterer Vorteil wäre eine Prüfkabine nach der ISO 8559, wodurch die Konzentration der Konsumenten verbessert wird und gegebenenfalls noch präzisere Ergebnisse erlangt werden können. Die DIN EN ISO 11132 gibt einen guten Leitfaden zur Bewältigung der Panel- und Prüfungsleistung. Jedoch werden einige Zwischenrechnungen benötigt und die Auswertung erfolgt anhand von mehreren Tabellen und Abbildungen. Die von Lepage et al. (2014) an die ANOVA basierende, durchgeführte „Temporal Dominance of Sensation“ bietet eine noch unkompliziertere Methode, die Panel- und Einzelprüferleistung anhand von einer Tabelle darzustellen. Die Interpretation der TDS-Methode wird einfach gehalten mit „ok“ (erfolgreiche Unterscheidung und Zustimmung), „A!“ (erfolgreiche Unterscheidung aber keine Zustimmung) und „D!“ (keine Unterscheidung). Des Weiteren ergibt sich eine einzige übersichtliche Tabelle, anhand der die Interpretation erfolgt. So kann schon auf den ersten Blick ermittelt werden, welche Prüfer nicht mit dem Panel einig sind und somit Schulungsbedarf benötigen, ohne in die Rohdaten schauen zu müssen. Dabei werden die „p-values“ nicht wie üblich durch den Fisher-Test, sondern durch einen Permutation-Test berechnet. Im Gegensatz zu der gemischten linearen Methode, die 2005 von Latreille et al. durch

die SAS-Software berechnet wurde, eignet sich die DIN EN ISO 11132 auch für größere Panels, mehrere Sitzungen oder mehrere zu testende Produkte. Latreille et al. empfiehlt für die gemischte lineare Methode weniger als fünf Sitzungen und möglichst wenige Produkte (Latreille et al. testeten anhand von drei Lippenstiften), ansonsten gestaltet sich die Auswertung und Interpretation schwierig. In dieser Arbeit wurde mit sechs Produkten á drei Sitzungen und mit 13 und 14 Prüfern gearbeitet. Die Auswertung und Interpretation der Ergebnisse gestaltete sich unkompliziert, da die DIN EN ISO 11132 einen roten Faden in der Auswertung der Leistung bietet.

6. Zusammenfassung

In dieser Bachelorarbeit wurde das Arbeiten mit der DIN EN ISO 11132 beschrieben und in der Handhabung geprüft. Des Weiteren fand eine Überprüfung, sowie ein Vergleich der Daten eines geschulten und ungeschulten Panels, statt. Dadurch wurde die Panelleistung ermittelt die in der Sensorik von besonderer Bedeutung ist. Mit der Panelleistung kann überprüft werden, wie gut z.B. ein geschultes Panel trainiert wurde, aber auch Attribute herauskristallisiert werden, die für ein schlechtes Verständnis sorgen.

Für die Bestimmung der Panelleistung nach DIN EN ISO 11132 wurden zwei geschulte Panels mit der Größe von 13 und 14 Probanden aufgebaut. Diese wurden in zwei Sitzungen á 90 Minuten auf das Produkt, diverse Orangensaft aus dem Handel, geschult. In einem Übungsdurchlauf wurde das Verständnis des Panels überprüft und anschließend in drei Sitzungen eine Profilierung nach DIN 10967 durchgeführt. Die Auswertung erfolgte mit Hilfe der Software „FIZZ“, wo eine Varianzanalyse durchgeführt wurde. Parallel fand ein Verbrauchertest nach DIN 10974, mit 98 ungeschulten Prüfern vor der Mensa der Hochschule Neubrandenburg, statt. Die Aufgabe der Verbraucher bestand aus der Angabe der Beliebtheit auf einer 9-Punkt-hedonischen Skala nach Jones et al. (1955), Peryam und Pilgrim (1957). Außerdem erfolgte eine Intensitätsangabe anhand von „Just-About-Right“-Fragen, wo die Intensität der Farbe, des bitteren und orangigen Geschmackes gefragt war. Zudem kam ein „Check-All-That-Apply“-Fragebogen, wo überprüft werden sollte, welche Attribute auf die Produkte zutreffen. Die Auswertung des Verbrauchertests wurde teilweise (Beliebtheit und JAR) durch die Software „FIZZ“ und ebenfalls der Varianzanalyse, sowie mit Hilfe der Software „Xlstat“ (CATA-Analyse), durchgeführt. Es konnte festgehalten werden, dass beide geschulte Panels ähnliche Bewertungen abgegeben haben. Die Panelleistung erwies sich als gut, da die Produktunterscheidung anhand von allen Attributen größtenteils höchstsignifikant verlief. Lediglich das Arbeiten mit „süß im Geruch“ zeigte bei beiden Panels Probleme auf, da die Produkte anhand des Attributes nicht signifikant unterschieden werden konnten. Die Prüfereinzelleistung zeigt, dass beide Panels sowohl gute Prüfer, aber auch Prüfer die eine weitere Schulung benötigen, aufwies. Insgesamt konnte die Gruppe 2 besser abschneiden als Gruppe 1. Der Vergleich des geschulten und ungeschulten Panels stellte dar, dass das ungeschulte Panel ähnliche Ergebnisse liefert und somit gut mit dem geschulten Panel korrelierte. Es zeigte aber auch, dass das geschulte Panel konstanter in der Bewertung lag, als das ungeschulte Panel. Es konnte auch festgestellt werden, dass in fünf von sechs Säften der orangige Geschmack von großer Bedeutung war. Eine Reduzierung des orangigen Geschmacks würde also auch zu einer Beliebtheitsabnahme führen. Ebenfalls wurde ersichtlich, dass das Produkt

„Amecke Mandarine-Orange“ eine hohe Beliebtheit aufwies und „Punica“ eine niedrige. Insgesamt konnte das Verbraucherpanel in fünf Beliebtheitskategorien aufgeteilt werden, wobei die Kategorie der Amecke-Produkte die meisten Anhänger fand.

Die DIN EN ISO 11132 erweist sich als eine gute und anerkannte Methode zur Bestimmung der Panelleistung. Das Arbeiten mit dieser internationalen Norm ist simple und die Ergebnisinterpretation unkompliziert. Des Weiteren ergibt sich die Erkenntnis, dass mit einem ungeschulten Panel ebenfalls gearbeitet werden kann, wenn ein geschultes Panel nicht zur Verfügung steht oder finanzielle Mittel und Zeit für eine Schulung nicht gegeben sind.

Anhang 1.1

+

F.C. 0038 O.N. 0057 PAGE 00:1/004

+

V e r b r a u c h e r t e s t z u r B e u r t e i l u n g
v o n O r a n g e n s a f t

A n g a b e n z u r P e r s o n :

G e s c h l e c h t :

| | |
|----------|--------------------------|
| weiblich | <input type="checkbox"/> |
| männlich | <input type="checkbox"/> |

A l t e r :

| | |
|---------------|--------------------------|
| 18-25 Jahre | <input type="checkbox"/> |
| 26-35 Jahre | <input type="checkbox"/> |
| 36-45 Jahre | <input type="checkbox"/> |
| über 45 Jahre | <input type="checkbox"/> |

W i e h ä u f i g t r i n k e n S i e O r a n g e n s a f t ?

| | |
|-----------------|--------------------------|
| täglich | <input type="checkbox"/> |
| 1 x pro Woche | <input type="checkbox"/> |
| 2-3 x pro Monat | <input type="checkbox"/> |
| selten | <input type="checkbox"/> |

B e u r t e i l u n g d e r e i n z e l n e n P r o d u k t e :

Bewerten Sie bitte die Beliebtheit, d.h. wie gut Ihnen das Produkt insgesamt gefällt, auf einer Skala von 1 („missfällt außerordentlich“) über 5 („weder gefällt noch missfällt“) bis 9 („gefällt außerordentlich“).

Bitte beurteilen Sie anschließend bei jeder einzelnen Probe die Intensität der Farb- und Geschmackseigenschaften, d.h. wie stark oder schwach Sie diese beim Produkt wahrnehmen. Die Einordnung erfolgt hier zwischen „viel zu schwach“ (1) über „genau richtig“ (3) bis „viel zu intensiv“ (5).

+

+

Anhang 1.2

+ F.C. 0038 O.N. 0057 PAGE 002 / 004 +

Verbrauchertest zur Beurteilung
von Orangensaft

453

| | | | | | | | | | |
|----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| Gesamteindruck | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| | <input type="checkbox"/> |
| | mißfällt | sehr | | | mittelmäßig | | | gefällt | sehr |

453

| | | | | | |
|-----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | 1 | | 3 | | 5 |
| Farbintensität | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |
| | 1 | | 3 | | 5 |
| orangenartig GS | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |
| | 1 | | 3 | | 5 |
| bitterer GS | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |

872

| | | | | | | | | | | |
|----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|------|
| Gesamteindruck | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| | <input type="checkbox"/> | |
| | mißfällt | sehr | | | | mittelmäßig | | | gefällt | sehr |

872

| | | | | | |
|-----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | 1 | | 3 | | 5 |
| Farbintensität | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |
| | 1 | | 3 | | 5 |
| orangenartig GS | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |
| | 1 | | 3 | | 5 |
| bitterer GS | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |

+ +

Anhang 1.3

+
F.C. 0038 O.N. 0057 PAGE 003 / 004 +

Verbrauchertest zur Beurteilung
von Orangensaft

710

| | | | | | | | | | | |
|----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|------|
| Gesamteindruck | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| | <input type="checkbox"/> | |
| | mißfällt | sehr | | | | mittelmäßig | | | gefällt | sehr |

710

| | | | | | |
|-----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | 1 | | 3 | | 5 |
| Farbintensität | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |
| | 1 | | 3 | | 5 |
| orangenartig GS | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |
| | 1 | | 3 | | 5 |
| bitterer GS | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |

548

| | | | | | | | | | |
|----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| Gesamteindruck | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| | <input type="checkbox"/> |
| | mißfällt | sehr | | | mittelmäßig | | | gefällt | sehr |

548

| | | | | | |
|-----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | 1 | | 3 | | 5 |
| Farbintensität | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |
| orangenartig GS | 1 | | 3 | | 5 |
| | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |
| bitterer GS | 1 | | 3 | | 5 |
| | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |

+

Anhang 1.4

+ F.C. 0038 O.N. 0057 PAGE 004 / 004 +

Verbrauchertest zur Beurteilung
von Orangensaft

129

| | | | | | | | | | |
|----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| Gesamteindruck | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| | <input type="checkbox"/> |
| | mißfällt | sehr | | | mittelmäßig | | | gefällt | sehr |

129

| | | | | | |
|-----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | 1 | | 3 | | 5 |
| Farbintensität | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |
| orangenartig GS | <input type="checkbox"/> |
| | 1 | | 3 | | 5 |
| | viel zu schwach | | genau richtig | | viel zu intensiv |
| bitterer GS | <input type="checkbox"/> |
| | 1 | | 3 | | 5 |
| | viel zu schwach | | genau richtig | | viel zu intensiv |

386

| | | | | | | | | | | |
|----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|------|
| Gesamteindruck | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| | <input type="checkbox"/> | |
| | mißfällt | sehr | | | | mittelmäßig | | | gefällt | sehr |

386

| | | | | | |
|-----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | 1 | | 3 | | 5 |
| Farbintensität | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |
| | 1 | | 3 | | 5 |
| orangenartig GS | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |
| | 1 | | 3 | | 5 |
| bitterer GS | <input type="checkbox"/> |
| | viel zu schwach | | genau richtig | | viel zu intensiv |

+ #

Anhang 1.5

Sie dürfen die 6 gereichten Proben erneut probieren. Tragen Sie in den grauen Bereich, die verschiedenen 3-stelligen Probenzahlen ein. Achten Sie beim Probieren hierbei auf die Begriffe in der linken Spalte. Trifft ein Merkmal auf ein Produkt zu, so kreuzen sie das Feld mit der Probe und dem Merkmal an.

z.B. Welche Probe erscheint Ihnen künstlich? Sollte ein oder mehrere Merkmale auf die Produkte nicht zutreffen, so lassen Sie die Felder frei.

| gelblich | | | | | | |
|------------------------|--|--|--|--|--|--|
| orangefarben | | | | | | |
| bräunlich | | | | | | |
| süß | | | | | | |
| sauer | | | | | | |
| bitter | | | | | | |
| orangenartig | | | | | | |
| zitronenartig | | | | | | |
| grapefruitartig | | | | | | |
| Fremdgeschmack | | | | | | |
| künstlich | | | | | | |
| lecker | | | | | | |
| harmonisch | | | | | | |
| erfrischend | | | | | | |

Literaturverzeichnis

Ares, G., Barreiro, C., Deliza, R. et al.: Application of Check-That-All-Apply question to the development of chocolate milk desserts. 28.04.2010. [Zugriff am: 28.11.17], Verfügbar unter: <http://onlinelibrary.wiley.com/doi/10.1111/j.1745-459X.2010.00290.x/full>

Beeren, C., Schneider-Häder, B.: Sensory Claims: Methodische Vorgehensweise zur Entwicklung und Untermauerung. 15/2015. [Zugriff am: 08.11.17], Verfügbar unter: <http://www.dlg.org/sensoryclaims.html>

Bongartz, A., Mürset, U.: Statistische Methoden in der Sensorik (Teil 2): „Verbrauchertest“. 06/2011. [Zugriff am 21.11.17], Verfügbar unter: http://www.dlg.org/fileadmin/downloads/fachinfos/sensorik/Downloads/Statistische_Methoden/Statistik2_AB_Sensorik_2011_06.pdf

Carr, B.T., Civille, G.V., Meilgaard, M.C.: Sensory Evaluation Techniques. 5. Auflage. Boca Raton: CRC Press, 2015.

Chambers, D.H., Chambers IV, E., Lotong, V. Categorization of Commercial Orange Juices Based on Flavor Characteristics. Journal of Food-Science, Nr. 2, Vol. 68, Institute of Food Technologists, United States, 2003.

Derndorfer, E., Schneider-Häder, B.: Sensorische Analyse: Methodenüberblick und Einsatzbereich. 05/2016. [Zugriff am: 08.11.17], Verfügbar unter: http://www.dlg.org/sensorische_analyse.html

DIN 10967-2: Profilprüfung, 10/2000.

DIN 10974: Sensorische Analyse- Verbrauchertests, 12/2008.

DIN EN ISO 11132: Sensorische Analyse – Methodologie – Leitlinien zur Überprüfung der Leistungsfähigkeit eines quantitativen sensorischen Panels, Deutsche Fassung, 2017.

DLG-Verbraucher-Information: Warenkund Fruchtsäfte- ein Überblick. [Zugriff am: 05.12.17], Verfügbar unter: <http://www.dlg-verbraucher.info/de/lebensmittelwissen/warenkunde/fruchtsaefte-ein-ueberblick.html>

Ellert, J.: Verbraucherakzeptanz und Preference Driving Factors bei Zuchtlachsen(...) unterschiedlicher Herkunft. Hamburg, 18.07.2006.

Fliedner, I., Wilhelmi, F.: Grundlagen und Prüfverfahren der Lebensmittelsensorik, 2. Auflage. Behr's Verlag. Hamburg, 1995.

Fruchtsaft- und Erfrischungsgetränkeverordnung (FrSaftErfrischGetrV) vom 24.05.2004. Stand: 5.17.17, [Zugriff am: 05.12.17], Verfügbar unter: https://www.gesetze-im-internet.de/frsaftv_2004/FrSaftErfrischGetrV.pdf

Giechau, S.: Auswirkung des Reifegrades von Bananen auf die Just-About-Right- und Akzeptanzinstufung durch Konsumenten. Hamburg, 06.07.2009.

Latreille, J., Mauger, E., Ambroisine, L., Tenenhaus, M., Vincent, M., Navarro, S., Guinot, C.: Measurement of reliability of sensory panel performances. Food Quality and Preference, Ausgabe 17/5. Frankreich, 2005.

Lawless, H.T., Heymann, H.: Sensory Evaluation of Food, 2. Auflage. New York: Springer, 2010.

Lepage, M., Neville, T., Rytz, A., Schlich, P., Martin, N., Pineau, N. Panel performance for Temporal Dominance of Sensations. Food Quality and Preference, Ausgabe 38. Schweiz und Frankreich, 2014.

Meullenet, J.-F., Dooley, L., Lee, Y.-S.: The application of CATA consumer profiling to preference mapping of vanilla ice cream and its comparison to classic external preference mapping. Food Quality and Preference, Ausgabe 21/4. Fayetteville, 2009.

Meullenet, J.-F., Xiong, R. Findlay, C.J.: Multivariate and Probabilistic Analyses of Sensory Science Problems. IFT Press and Blackwell Publishing. New York, 2007.

Meier, J.: Methoden zur Qualitätssicherung für das Arbeiten mit deskriptiven Panels in der Lebensmittelsensorik. Hochschule Neubrandenburg, 2006.

Moskowitz, H.R., Muñoz, A.M., Gacula, M.C.: Viewpoints and controversies in sensory science and consumer product testing, Connecticut USA: Food & Nutrition press, inc., 2003.

Nießen, M., Thölking, S.: Sensorische Prüfverfahren, 1. Auflage. Behr's Verlag. Hamburg, 2007.

Schüssler, K., Paul D., Rummel, C., aus: Busch-Stockfisch M.: Praxishandbuch Sensorik in der Produktentwicklung und Qualitätssicherung. Behr's Verlag. Hamburg, 08/2002.

Schröder, K.: Darstellung der Konsumerakzeptanz unterschiedlich gegarter Fleischsorten mittels Internal Preference Mapping und Penalty Analyse (DIPLOMARBEIT). Hamburg, 29.09.2008.

Statista.com: Pro-Kopf-Konsum von Orangensaft in Deutschland in den Jahren 2005 bis 2016 (in Liter). 2017. [Zugriff am 05.12.2017], Verfügbar unter:

<https://de.statista.com/statistik/daten/studie/167146/umfrage/pro-kopf-verbrauch-von-orangensaft-in-deutschland-seit-2002/>

Statista.com: Pro-Kopf-Konsum von Apfelsaft in Deutschland in den Jahren 2005-2016 (in Litern). 2017. [Zugriff am 05.12.2017], Verfügbar unter:

<https://de.statista.com/statistik/daten/studie/167145/umfrage/pro-kopf-verbrauch-von-apfelsaft-in-deutschland-seit-2002/>

Tsalma, P.: Assessing sensory panel performance using generalizability theory. Food Quality and Preference, Ausgabe 47/A. Großbritannien, 2015.

McEwan, J.A., Heiniö, R.-L., Hunter, E.A., Lea, P.: Proficiency testing for sensory ranking panels: measuring panel performance. Food Quality and Preference, Ausgabe 14/3. Finnland/Großbritannien/Norwegen, 2002.

McEwan, J.A., Hunter, E.A., van Gemert, L.J., Lea, P.: Proficiency testing for sensory profile panels: measuring panel performance. Food Quality and Preference, Ausgabe 13/3. Großbritannien/Niederlande/Norwegen, 2002.

Van Kleef, E., van Trijp, H.C.M., Luning, P.: Internal versus external preference analysis: An exploratory study on end user-evaluation. Food Quality and Preference, Ausgabe 17/5. Wageningen, Niederlande, 2006.

Welle, F.: Der Geschmack entscheidet – Verpackungssensorik – Kombination aus Analytik und menschlichen Sinnen. Fraunhofer Institut für Verfahrenstechnik und Verpackung, Freising. [Zugriff am 16.01.2018], Verfügbar unter:

http://ernaehrungsdenkwerkstatt.de/fileadmin/user_upload/EDWText/TextElemente/Lebensmittel/Sensorik-Geschmack_01/Sensorische_Bewertung_Methoden_Infos_OLT_04_04_09.pdf

Xlstat.com: Cata-Analysis in Excel Tutorial. [Zugriff am 28.12.2017], Verfügbar unter:

https://help.xlstat.com/customer/en/portal/articles/2062449-running-a-cata-check-all-that-apply-analysis-in-xlstat?b_id=9283

Xlstat.com: Internes Präferenzmapping. [Zugriff am 08.01.2018], Verfügbar unter:
<https://www.xlstat.com/de/loesungen/eigenschaften/internes-präferenzmapping>

Xlstat.com: Penalty analysis in Excel Tutorial. 20.10.2017. [Zugriff am 03.12.2017], Verfügbar unter: https://help.xlstat.com/customer/en/portal/articles/2062259-penalty-analysis-in-excel-tutorial?b_id=9283

Xlstat.com: Preference Mapping in Excel Tutorial. 20.10.2017 [Zugriff am 08.12.2017], Verfügbar unter: https://help.xlstat.com/customer/en/portal/articles/2062247-running-a-preference-mapping-with-xlstat?b_id=9283

Abbildungsverzeichnis

| | |
|--|----|
| Abbildung 1: Beispiel einer Intensitätsauswertung des gesamten Panels (Meier, 2006)..... | 17 |
| Abbildung 2: Beispiel eines CATA-Prüfformulars (Derndorfer und Schneider-Häder, 2016) | 21 |
| Abbildung 3: Beispiel JAR-Bewertungen (Xlstat,2017)..... | 25 |
| Abbildung 4: Beispiel der Mean Drops (Xlstat, 2017) | 26 |
| Abbildung 5: Beispiel Penalties (Xlstat, 2017) | 27 |
| Abbildung 6: Beispiel zweidimensionaler Biplot eines internal pref.maps. (Xlstat, 2017)..... | 29 |
| Abbildung 7: Beispiel external preference mapping Biplot (Ellert, 2006) | 31 |
| Abbildung 8: Beispiel eines Countour Plots (Xlstat, 2017) | 32 |
| Abbildung 9: Probenpräsentation..... | 35 |
| Abbildung 10: Altersaufteilung der Probanden | 37 |
| Abbildung 11: Häufigkeit des Orangensaftverzehrs | 38 |
| Abbildung 12: PCA-Plot der JAR-Attribute (Achse 1 + Achse 2 =96,6%)..... | 42 |
| Abbildung 13: JAR-Bewertungen Punica | 43 |
| Abbildung 14: Mean Drops des Punica Saftes..... | 44 |
| Abbildung 15: Penalties des Punica Saftes | 45 |
| Abbildung 16: JAR-Bewertung TIP Nektar..... | 46 |
| Abbildung 17: Mean Drops des TIP-Nektars..... | 46 |
| Abbildung 18: Penalties des TIP-Nektars | 47 |
| Abbildung 19: JAR-Bewertung Amecke plus Zink und Vitamin C | 48 |
| Abbildung 20: Mean Drops des Amecke plus Zink und Vitamin C Saftes | 49 |
| Abbildung 21: Penalties des Amecke plus Zink und Vitamin C Saftes..... | 50 |
| Abbildung 22: JAR-Bewertung Hohes C 100% Orangensaft | 51 |
| Abbildung 23: Mean Drops des Hohes C 100% Orangensaftes | 52 |
| Abbildung 24: Penalties des Hohes C 100% Orangensaftes..... | 52 |
| Abbildung 25: JAR-Bewertung Amecke Mandarine und Orange | 53 |
| Abbildung 26: Mean Drops des Amecke Mandarine und Orange Saftes | 54 |
| Abbildung 27: Penalties des Amecke Mandarine und Orange Saftes..... | 55 |
| Abbildung 28: JAR-Bewertung Hohes C O-T-L | 56 |
| Abbildung 29: Mean Drops des Hohes C O-T-L Saftes | 57 |
| Abbildung 30: Penalties des Hohes C O-T-L Saftes..... | 57 |
| Abbildung 31: Nutzung der CATA-Attribute in Prozent..... | 59 |
| Abbildung 32: Beliebtheitsüberblick anhand eines PCA-Plots..... | 60 |
| Abbildung 33: PCA Plot mit Produkten und Attributen | 61 |

| | |
|--|----|
| Abbildung 34: Mean Impact | 62 |
| Abbildung 35: Anova-Graph der Gruppe 1..... | 63 |
| Abbildung 36: Durchschnittsbewertungen des Attributes „süßlich Geruch“ | 65 |
| Abbildung 37: Durchschnittsbewertungen anhand des Attributes „gelblich/bräunlich Aussehen“ | 66 |
| Abbildung 38: PCA-Plot Gruppe 1 | 67 |
| Abbildung 39: Anova-Graph Gruppe 2..... | 71 |
| Abbildung 40: Durchschnittsbewertung des Attributes „süßlich im Geruch“ | 72 |
| Abbildung 41: Durchschnittsbewertungen des Attributes „gelb-orange Aussehen“ | 73 |
| Abbildung 42: PCA-Plot Gruppe 2 | 74 |
| Abbildung 43: Platzierung der Säfte durch das Konsumentenpanel..... | 77 |
| Abbildung 44: Produktpräferenz durch das Konsumentenpanel..... | 78 |
| Abbildung 45: PCA-Bi-Plot des geschulten Panels Gruppe 1 anhand des Internal Preference Mapping | 79 |
| Abbildung 46: PCA-Bi-Plot des geschulten Panels Gruppe 2 anhand des Internal Preference Mapping | 80 |
| Abbildung 47: Darstellung der Konsumentengruppen anhand einer Preference Map..... | 81 |
| Abbildung 48: Countour Plot | 82 |

Tabellenverzeichnis

| | |
|--|----|
| Tabelle 1: Rohdatenbeispiel der ANOVA-Tabelle | 16 |
| Tabelle 2: Beispiel einer Varianzanalyse | 18 |
| Tabelle 3: 9-Punkt hedonische Skala | 20 |
| Tabelle 4: JAR-Kategorien anhand des Attributes „sauer“ | 23 |
| Tabelle 5: Produkte für die Untersuchung | 33 |
| Tabelle 6: Durchschnittsbewertung mit Standardabweichung..... | 38 |
| Tabelle 7: ANOVA Beliebtheitsprüfung..... | 39 |
| Tabelle 8: Post hoc Test, Faktor „Produkt“ mit L.S.D.: 0,57..... | 39 |
| Tabelle 9: ANOVA der 3 JAR-Attribute | 40 |
| Tabelle 10: Post hoc-Test Farbintensität mit LSD: 0,24..... | 41 |
| Tabelle 11: Post hoc-Test orangenartiger GS mit LSD: 0,27 | 41 |
| Tabelle 12: Post hoc-Test bitter GS mit LSD: 0,28 | 41 |
| Tabelle 13: Cochran's Q-Test..... | 59 |
| Tabelle 14: Anova „süßlich Geruch“ Gruppe 1 | 64 |
| Tabelle 15: Leistungsprüfung der Einzelpersonen Gruppe 1 | 69 |
| Tabelle 16: Anova-Tabelle „süßlicher Geruch“ Gruppe 2..... | 71 |
| Tabelle 17: Leistungsprüfung der Einzelpersonen Gruppe 2..... | 75 |
| Tabelle 18: Rangfolge der Beliebtheit der Gruppe 1 | 82 |

Erklärung über die selbstständige Anfertigung der Arbeit

Hiermit versichere ich, dass die vorliegende Arbeit selbstständig angefertigt wurde und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt haben.

Ich erkläre weiterhin, dass die abgegebene elektronischen Fassungen mit der eingereichten Arbeit identisch sind.

Ort, Datum

Unterschriften

Neumann, Alexander