



SENSORIKK

MÅLING MED MENNESKELIGE SANSER

Sensorisk StudieGruppe





© Kopinor Pensum AS
3. utgave 2015

ISBN 9788213030762

Det må ikke kopieres fra denne boka i strid med åndsverkloven eller andre avtaler om kopiering inngått med Kopinor, interesseorganisasjon for rettighetshavere til åndsverk. Kopiering i strid med lov eller avtale kan medføre erstatningskrav og inndragning, og kan straffes med bøter eller fengsel.

Boka utgis med støtte fra Kunnskapsdepartementet
v/Lærebokutvalget for høyere utdanning.

Henvendelser om denne boka kan rettes til:

Kopinor Pensum
Postboks 1663 Vikta, 0120 Oslo
www.pensum.no

Faglig koordinator: Marit Rødbotten, tidligere ved Nofima, Ås

Øvrige i bokredaksjonen: Liv Bente U. Strandos, Elopak AS og Tone Eikrem Nyvold,
Orkla Foods Norge

Omslagsdesign og layout: Julie Holst Gjermundsen

Figur 4.6 og 4.7 er brukt med tillatelse fra NCS Natural Colour System®©

Figur 2.3 er brukt med tillatelse fra Bang & Olufsen

Trykk og innbinding: Allkopi AS

Papir: 130 g Galleri Art Silk

Omslag: 300 g Edixion





Sensorisk StudieGruppe takker
Marit Rødbotten
for det hun har betydd for sensorikken
i Norge og for at hun gjorde denne
utgaven av boka mulig.





FORORD

Sensoriske metoder benyttes for bedømmelse av næringsmidler, men også for bedømmelse av andre produkter hvor de menneskelige sansene er nødvendige som måleinstrument. I alle situasjoner hvor nye produkter skal lages eller omskapes må vi bruke minst en av våre sanser (syn, lukt, smak, hørsel og følelse) for å måle produktkvalitet. Det skjer ikke alltid ved bevisst å benytte spesielle metoder, men særlig innen matvareindustrien gjøres det bevisste valg av metoder og måter for gjennomføring av sensoriske analyser.

Sensorisk StudieGruppe (SSG) ble opprettet i Norge i 1972. Pådriverne var personer innen norsk næringsmiddelindustri, forskningsinstitusjoner og læresteder. I løpet av de årene SSG har eksistert har medlemstallet økt og det er fortsatt åpent for nye medlemmer som har en arbeidssituasjon som krever kunnskap om sensoriske analyser. I tillegg til møter med faglig innlegg har SSG arrangert flere kurs for personer som har og/eller skal ha ansvaret for sensoriske bedømmelser i sin bedrift. Tema på disse kursene har vært grunnleggende praktisk arbeid med etablering av sensorisk panel, gjennomføring av analyser med forskjellige sensoriske metoder og resultatbehandling med statistiske analyser.

Sensorisk analyse, Bedømmelse av næringsmidler var den første norske læreboka i sensorisk analyse og den ble utgitt i 1977 med medlemmer i SSG som forfattere. Boka ble revidert i 1997 og nå som tredje utgave i 2015 med en endring av tittelen: Sensorikk, Måling med menneskelige sanser. Årsaken til endringen av boktittelen er at vi vil vise at sensoriske metoder ikke bare er forbeholdt analyser av næringsmidler, men er målemetoder for sensoriske analyser av alle typer produkter.





Innholdet i boka gir kunnskap om de menneskelige sansene som er i aktivitet under sensoriske målinger. Det er nødvendig å vite hvordan sansene fungerer for fullt ut å forstå betydningen av forhold rundt presentasjon av prøver som skal sensorisk analyseres. Siden det er mennesker som utgjør instrumentet er det viktig å tilrettelegge analyseforholdene riktig for å unngå sammenblanding av subjektive og objektive resultater. Problemstillingen som ligger til grunn for en sensorisk analyse avgjør metodevalg, og metoder egnet for analyser med trente dommerpanel og for forbrukertester blir beskrevet. For å trekke riktige konklusjoner fra det til dels store datamaterialet som produseres under analyser er det nødvendig med statistisk behandling, og boka gir innføring i de viktigste statistiske operasjonene. Boka inneholder kapitler fra fagområder som benytter sensorikk i sin virksomhet, både næringsmidler og produkter som ikke skal spises.

Dette er ei lærebok som vil egne seg som undervisningsmateriale ved læresteder som underviser i næringsmiddelfag og ved spesielle kurs som er tilrettelagt for personer som har ansvar for sensorisk kvalitet av forskjellige produkter. I tillegg vil boka være nyttig som oppslagsverk for personer som gjennomfører sensoriske analyser i sin bedrift.

Forfatterne til kapitlene i denne utgaven av boka er personer som i flere år har arbeidet med forskjellige oppgaver relatert til sensorikk, enten i undervisning, forskning, offentlig eller privat bedrift.

Personer som er medansvarlige for utgivelse av denne boka:

Almli, Valerie Lengard, forsker, Nofima, Ås. Førstemanuensis, Inst. for kjemi, bioteknologi og matvitenskap, NMBU

Andersen, Urd Bente, fagsjef sensorikk, Orkla Foods Norge

Bech, Søren, Dr, professor, Head of Research, Bang & Olufsen, Aalborg Universitet, Danmark

Berg, Egil Weie, seniorinspektør, Mattilsynet, region Midt, Trondheim og omland

Bergslien, Helge, fysiolog, Facilitator Kunnskap, NCS Culinary, Stavanger

Carlehög, Mats, prosjektleder sensorikk, Nofima, Ås





Hegna, Margit Anette Bleg, Laboratory- and R&D Manager Home Care R&D Dept., Lilleborg, Orkla

Hersleth, Margrethe, seniorforsker, Nofima, Ås. Førsteamanuensis, Inst. for kjemi, bioteknologi og matvitenskap, NMBU

Gjermundsen, Julie Holst, grafisk designer, Julie Gjermundsen Design

Kraggerud, Hilde, forsker, TINE FoU Måltidets Hus, Stavanger

Larsen, Wenche Emblem, forsker, Møreforskning AS

Lea, Per, forsker statistikk, Nofima, Ås

Mainusch, Yvonne, produktutvikler, Orkla Foods Norge

Merok, Kjell J., grafisk konsulent, tidligere ved Nofima, Ås

Nyrud, Anders Q., avdelingsleder/forskningsleder, Norsk Treteknisk Institutt, Oslo

Nyvold, Tone Eikrem, produktutvikler m/sensorikk, Orkla Foods Norge

Roos, Anders, professor, Institusjon för skogens produkter, SLU, Uppsala, Sverige

Rødbotten, Marit, forsker sensorikk, tidligere ved Nofima, Ås

Simonstad Valle, Elin, konsulent sensorikk/forsøksdesign, TINE FoU

Strandos, Liv Bente U., Manager Sensory Services, Elopak AS

Svenningsen, Mari, fotograf, www.marisvenningsen.no

Waldenstrøm, Lene, overingeniør, Høgskolen i Sør-Trøndelag

Bedrifter som har bidratt økonomisk til utgivelse av denne boka:

A/S Vinmonopolet, Arcus AS, Bama, Coop Norge, Drytech, Elopak AS, Eurofins Food & Agro Testing Norway AS, Mills DA, Møreforskning AS, Nofima AS, Nortura, O.Kavli AS, Orkla Foods Norge, Ringnes A/S, Takeda-Nycomed, Tine, Unil AS.





INNHOOLD

FORORD

- | | |
|---|----|
| 1. INNLEDNING | 11 |
| Marit Rødbotten | |
| 2. SANSENE | 15 |
| Helge Bergslie | |
| 2.1 Introduksjon | |
| 2.2 Opplevelse | |
| 2.3 Den fysiologiske tilstanden | |
| 2.4 Læring og erfaring | |
| 2.5 Helse | |
| 2.6 Sansene | |
| 2.6.1 <i>Syn</i> | |
| 2.6.2 <i>Hørsel</i> | |
| 2.6.3 <i>Smak</i> | |
| 2.6.4 <i>Lukt</i> | |
| 3. RETNINGSLINJER FOR GJENNOMFØRING
AV SENSORISKE ANALYSER | 42 |
| Liv Bente U. Strandos | |
| 3.1 Hovedpunkter for analysen | |
| 3.2 Sensoriske dommere, rekruttering og utvelgelse | |
| 3.3 Trening av sensoriske dommere | |
| 3.4 Panelleders oppgaver | |

7





- 3.5 Prøvepreparering og serveringsprosedyrer
- 3.6 Prøvepresentasjon
- 3.7 Testområdet og bedømmelseslokalet
- 3.8 Faktorer av betydning for sensorisk analyse

4. METODER FOR OBJEKTIV BEDØMMELSE 77

Marit Rødbotten og Mats Carlehög

- 4.1 Skalaer
- 4.2 Forskjellstester
 - 4.2.1 *Triangeltest*
 - 4.2.2 *Partest*
 - 4.2.3 *Duo-trio-test*
 - 4.2.4 *To-av-fem-test*
 - 4.2.5 *A-ikke-A-test*
 - 4.2.6 *Rangeringstest*
- 4.3 Beskrivende tester
 - 4.3.1 *Profilering*
 - 4.3.2 *Kvalitetskontrolltest*
 - 4.3.3 *Indeks metode, QIM*
 - 4.3.4 *Tidsavhengige målinger*

5. METODER FOR FORBRUKERTESTER 117

Margrethe Hersleth og Valerie Lengard Almlie

- 5.1 Introduksjon
- 5.2 Kvalitative studier
- 5.3 Repertory-grid- og free-choice-profilering
- 5.4 Preferanse-testing
- 5.5 Aksept-testing
- 5.6 Nyere sensoriske metoder
- 5.7 Valg av prøver ved conjoint-design
- 5.8 Betydning av testsituasjon (kontekst)
- 5.9 Nett-testing





6. PRODUKTUTVIKLING	136
Urd Bente Andersen, Tone Eikrem Nyvold og Yvonne Mainusch	
6.1 Produktutvikling	
6.2 Prosjektarbeid og prosjektgruppen	
6.3 Produktutviklingsprosessen (PU-prosessen)	
6.4 Sensorikk i produktutvikling	
6.5 Metoder og panel	
7. FORSKNING	148
Margrethe Hersleth	
7.1 Forskning innen sensoriske metoder	
7.2 Forskning basert på sensorisk kunnskap	
8. KVALITETSKONTROLL	153
Elin Simonstad Valle og Hilde Kraggerud	
9. OFFENTLIG TILSYN	158
Egil Weie Berg	
10. KVALITETSSIKRING AV SENSORISKE ANALYSER	165
Urd Bente Andersen	
11. STATISTISK BEHANDLING OG VURDERING AV DATA	173
Per Lea	
11.1 Innledning	
11.2 Gangen i en statistisk test	
11.3 Forskjellstester	
11.3.1 Duo-trio-test	
11.3.2 Triangeltest	
11.3.3 Partest	
11.3.4 To-av-fem-test	





11.4	Beskrivende tester	
11.4.1	<i>Enveis variansanalyse – urealistisk modell</i>	
11.4.2	<i>Toveis variansanalyse uten gjentak</i>	
11.4.3	<i>Rangeringer: Friedman's test</i>	
11.4.4	<i>Multiple sammenligninger</i>	
11.4.5	<i>Toveis variansanalyse med gjentak</i>	
11.4.6	<i>Flerveis variansanalyse</i>	
11.4.7	<i>Manglende verdier</i>	
11.5	Flere egenskaper samtidig: PCA	
11.6	Dyddykking i datamaterialet: PanelCheck	
12.	SENSORISK ANALYSE AV NON-FOOD	215
	Liv Bente U. Strandos	
12.1	Emballasje	
12.2	Hygieneprodukter og vaskemidler	
12.3	TV, -lyd	
12.4	Tregulv	
13.	SENSORISKE NETTVERK	224
14.	NMKL-PROSEDYRER, ISO-STANDARDE, ASTM-STANDARDE	226
15.	REFERANSER OG RELEVANT LITTERATUR	230
16.	VEDLEGG	235
	Per Lea (c. og d.)	
a.	Skjema	
b.	Statistiske tabeller	
c.	Enkel sannsynlighetsregning	
d.	Statistiske begreper	
17.	STIKKORD	264





INNLEDNING

Marit Rødbotten, tidligere ved Nofima

Daglig gjennomføres sensoriske analyser. Mange analyser gjennomføres som systematiske analyser og noen ganger som mer tilfeldige analyser for industrien i alle prosesser ved produktutvikling av spiselige og ikke-spiselige produkter. Men oftest gjennomføres sensorikk av privatpersoner som slett ikke tenker på at de utfører noen sensorisk bedømmelse. De bare gjør det, helt automatisk. De spiser mat som de smaker og lukter på, de sitter i en god stol og lytter til musikk eller de går en tur i skogen og nyter naturen. De utfører sensoriske analyser bare for seg selv i dyp konsentrasjon om opplevelsen eller helt ubevisst om de sansemessige impulsene.

Alt dette faller inn under begrepet sensorikk. Ubevisst eller bevisst registrerer, måler, analyserer og tolker vi reaksjoner vi får gjennom våre sanser. Vi ser, hører, lukter, smaker og føler med våre sanser, enten ved bruk av bare en av sansene eller flere samtidig.

Tidligere ble sensoriske analyser kalt organoleptiske analyser. Betegnelsen betyr analyse som måler stimuli registrert med et sansorgan. For eksempel vil et søtstoff som berører tunga føre til en stimulans i hjernen. Dette kan måles objektivt og bruk av betegnelsen organoleptisk analyse var et valg som bevisst ble tatt for å vise til objektive målinger. Men etter hvert som man fikk mer kunnskap om hvordan vi mennesker tolker slike enkeltstimuli som følge av læring og erfaring ble det klart at en organoleptisk analyse bare gir svar på om våre sanser kan registrere den bestemte stimulus. For å få kunnskap om hvordan den organoleptiske stimulus kan oppleves av mennesker må vi bruke vår evne til å gi uttrykk for den spesielle sansestimuleringen.





Den organoleptiske stimulus i kombinasjon med sansemessig registrering og tolking, gir en sensorisk analyse.

Utvikling av sensorikk som vitenskap

Med handel av matvarer kom behovet for en garanti for riktig kvalitet, og produktkvalitet var og er fortsatt mye av grunnlaget for prisen på varen. Tidlig på 1900-tallet startet en systematisk sensorisk kvalitetsvurdering av matvarer. Meieribransjen og bryggeriene var blant de første som etablerte faste rutiner for kvalitetskontroll av sine produkter. Oslo (Kristiania) fikk en offentlig melkekontroll i 1891 og en offentlig næringsmiddelkontroll i 1898. Hermetikkindustriens laboratorium i Stavanger ble opprettet i 1931, og var det første laboratoriet i Norge som bant annet utførte kvalitetskontroll av hermetiske matvarer. En stor del av fokuset i disse kontrollene var av teknisk og mikrobiologisk art, men prøver ble også sensorisk vurdert dog etter metoder som vi i dag ikke helt vil anerkjenne.

Det har tatt mange år før sensorikk ble godkjent som et vitenskapelig fagområde. Objektive målinger som kan gi reproduerbare resultater er et krav i vitenskapelig sammenheng. Den praktiske nytten av en sensorisk analyse økte etter hvert som metoder med dokumenterte prosedyrer ble etablert. Rose Mary Pangborn (1932–1990) er å betrakte som en pioner for utviklingen av det vi i dag kjenner som vitenskapen sensorisk analyse. Utallige vitenskapelige artikler og bøker har referert til boka «Principles of Sensory Evaluation of Food» skrevet av Amerine, Pangborn og Roessler, 1965. Dette er boka som for alvor satte fart i forskning rundt metoder som underbygger seriositeten av sensorikk som vitenskap. Flere og flere bedrifter har funnet verdien av å bruke sensorikk som et verktøy for å oppnå den produktkvalitet som forbrukeren foretrekker.

Begreper som objektiv og subjektiv er viktige i vitenskapelig sammenheng. Det vises ofte til objektive analyser utført med instrumenter til forskjell for subjektive analyser utført med mennesker som «instrument». Dette er selvsagt riktig, men her må vi også forstå begrepet «intersubjektiv». Intersubjektiv betyr «det som er felles for to eller flere





personer eller som de er enige om». Begrepet intersubjektivitet er viktig innen vitenskapsfilosofien der det hevdes at «intersubjektiv enighet blant kompetente personer regnes ofte som en nødvendig betingelse for objektivitet i vitenskapelig sammenheng» (Store norske leksikon). Et sensorisk panel skal bestå av kompetente personer slik det her er definert. I sensoriske undersøkelser viser vi til forbrukeranalyser hvor personer registrerer sine subjektive opplevelser av et produkt, mens et trent sensorisk panel produserer objektive resultater ut ifra enighet om registrerte sensoriske stimuli.

Allerede på 1940 tallet var det utviklet sensoriske metoder med formål å kunne gjennomføre objektive analyser av matvarer. Flavor Profile®-metoden (FP) ble utviklet av Arthur D. Little og med den metoden kan smak i produkter beskrives. Men denne metoden ble ikke helt godtatt som vitenskapelig objektiv, og flere metoder for beskrivelse av produkter ble etter hvert utviklet. Vi fikk Quantitative Descriptive Analysis®, QDA, en metode som ble utviklet på 1970-tallet av Herbert Stone og Joel Sidel, og Textur Profile (TP) metoden som ble utviklet av G.V. Civille og Liska i General Foods i 1975. Etter å ha arbeidet med både FP og TP en tid utviklet G.V. Civille en ny metode med basis i teknikker fra de to etablerte metodene, nemlig Sensory Spectrum® som hun presenterte i 1979.

Sensorikk som verktøy har etter hvert blitt et naturlig fagområde innen produktutvikling, kvalitetskontroll og forskning også i de Nordiske landene. Et nært samarbeid mellom sensorikere i disse landene har ført til faglig utvikling av metoder for effektiv og detaljert informasjon om både produkter og forbrukeres preferanser og holdninger til produkter. Svensken Bengtson publiserte i 1946 metoden triangeltest, og bryggeriet Carlsberg var blant de første som tok denne metoden i bruk i sin kvalitetskontroll. Metoden som måler sensorisk intensitet av egenskaper over en bestemt tid, kjent som Time-Intensity-metoden (TI) ble utviklet i Sveige på 1950-tallet.

Sensoriske analyser utført i kombinasjon med andre metoder som for eksempel fysiske-, kjemiske- og mikrobiologiske metoder produserer store mengder data. Multivariate metoder som kobler sensorikk til matematikk og statistikk har gitt verdifull innsikt i betydningen av





sensorikk som vitenskap. Norske forskere har hatt en sentral rolle i å vise næringsmiddelbransjen hvordan de kan hente detaljert informasjon om sine produkter nettopp med bruk av multivariate metoder. Behovet for utveksling av ideer og erfaringer med data fra sensorikk og forbrukerforskning førte til etablering av det internasjonale Sensoric Society.

Produsenter av varer har som mål å tilfredsstille forbrukeres krav og ønsker i forhold til produkter. En merkbar dreining av metodeforskningen innen sensorikk har gått fra metoder for objektiv beskrivelse av produkter til metoder egnet for studier av forbrukeres aksept og preferanser, samt holdninger og vaner. Danske MAPP (Centre for research on consumer relations in the food sector) satte på 1990-tallet fokus på metoder egnet for forbrukerforskning hvor psykologi er en viktig del.

Sensorikere fra store deler av verden ønsket å vise sin takknemlighet overfor Rose Mary Pangborn ved å tilegne et internasjonalt symposium i hennes navn. Finske sensorikere hadde et nært samarbeid med henne og det første Rose Mary Pangborn symposium ble arrangert i Finland i 1992. Annethvert år etter dette arrangeres et Pangborn Sensory Science Symposium på forskjellige steder i verden hvor hun har hatt innflytelse på sensorisk forskning.

Arbeidet med å etablere et objektivt sensorisk panel kan være ressurskrevende og tidkrevende og dermed en økonomisk utfordring for både store og små bedrifter. Det er derfor en naturlig konsekvens at det forskes på å fremskaffe instrumenter som kan produsere data med samme informasjon som fra det menneskelige «instrumentet». Forskere fra Island, Sverige, Finland, Danmark og Norge har bidratt til utvikling av «Elektronisk nese» og «Elektronisk munn». Men per i dag er det ikke noe instrument som helt kan gjenskape målinger som de menneskelige sansene kan.





SANSENE

Foto: Julie H. Gjermundsen





SANSENE

Helge Bergslien, NCS Culinary

2.1 Introduksjon

Alle levende organismer har behov for å kunne orientere seg i forhold til sine omgivelser. Dette for å kunne finne næring samt en mulighet for å kunne formere seg. Ulike planter vokser opp i biotoper som er optimale nettopp for dette formålet. Ulike dyrearter har funnet sine plasser på jorda ut i fra tilgang på mat og mulighet for formering. Mennesket er den arten som i størst grad har kunnet etablere seg over hele kloden. Dette takket være unike egenskaper knyttet til sansing, tolking av sanseintrykk, problemløsning og handling. Når temperaturen faller registreres dette av sanseceller som påvirker hjernen slik at vi opplever kulde. På bakgrunn av opplevelsen iverksettes en handling. Vi kler på oss. Vår evne til sansing er knyttet til sanseceller som er organisert i sanseorgan (øyne, øre, nese osv.). Disse organene fanger sanseintrykk som via nervetråder transporteres til hjernen for tolking. Når vi tolker et sett med sanseintrykk skaper dette en opplevelse. Lukt, syn og smak av en matrett gir oss en matopplevelse. Når vi setter oss inn i en ny bil, vil lukter, synsinntrykk, lyder og akselerasjon gi oss en opplevelse av bilen. Opplevelsen skapes i hjernen blant annet på bakgrunn av de sanseintrykk den mottar.

Sansing og persepsjon henger sammen. En kan gjerne si at sansingen er en prosess hvor sanseorganene detekterer stimuli og gjør dem om til nerveimpulser som sendes til hjernen. Persepsjonen er å forstå hva sansene forteller oss. Når vi forstår sanseimpulsene igangsettes en opplevelse av det som sanses.





2.2 Opplevelse

Opplevelsen formes i hjernen og påvirkes av svært mange forhold, som for eksempel erfaring og kunnskap. Mine kunnskaper om biler spiller en stor rolle når jeg opplever gleden av luktene i en ny bil. Måltids-erfaringer spiller en stor rolle for mine matopplevelser. Både kunnskap og erfaringer inngår i det en gjerne omtaler som en persons referanseramme (Sander K. 2014). Det vil si det grunnlaget vi vurderer våre sansestimuli på. Referanserammen omtales ofte som en funksjon av individets kunnskap, erfaring, holdninger, verdier, moral, interesser, livsstil, behov, ønsker, gruppepreferanser, forventninger og motiv. Denne rammen inneholder en rekke begrep som er individuelt formet. Det betyr at referanserammen er individuell eller subjektiv. I og med at denne er avgjørende for vår opplevelse vil også selve opplevelsen være individuell (se kapittel 5.8). På denne måten vil min måltidsopplevelse være min subjektive opplevelse, og ikke nødvendigvis være sammenfallende med mine måltidsvenners opplevelse.

Når vi går inn i en situasjon hvor sansene skjerpes, danner vi oss en forventning av hva som kan oppleves. Middagens sammensetning og hvem vi spiser sammen med former en forventning til middagsopplevelsen. Da sanseintrykk fra den aktuelle hendelsen sammenholdes med de skapte forventningene i hjernen vil også opplevelsen være subjektiv. Sanseintrykkene sammen med opplevelsen vil lagres i hjernen som erfaringer, og derved påvirke våre fremtidige opplevelser.

Referanserammen er av svært stor betydning når en trenes opp som dommer i et sensorisk panel. Her skal en benytte sanseapparatet for å gjennomføre en analyse sammen med flere meddommere. Det er derfor viktig at alle som inngår i dommerpanelet har felles referanser knyttet til dommersituasjonen.

En sentral funksjon i all sansing er at denne skal beskytte kroppen. Smak og lukt er gode eksempler på dette. Smaksansen gjør det mulig å skille mellom det som er spiselig og det som ikke er det. Når mat inneholder skadelige komponenter som toksiner vil vår evne til å kjenne for eksempel bitter smak kunne advare oss. På samme måte vil luktesansen kunne fortelle oss om «bedrevet» mat, som ikke bør spises.





2.3 Den fysiologiske tilstanden

For at kroppen skal fungere er den avhengig av at dens mange celler gjennomfører tiltenkte oppgaver på en god måte. Til dette trenger cellene energi som tilføres via mat og drikke. Blant næringsstoffene finnes elementer som brukes i forbrenning, utvikling av nye celler og nytt vev, samt vedlikehold. Den kjemiske forbrenningen i cellene er avhengig av et stabilt kjemisk miljø i kroppen. Det betyr at forholdet mellom salter, syrer og baser i kroppen må opprettholdes. Mat og drikke bidrar til påfylling av en rekke kjemiske komponenter som kroppen er avhengig av, og fine fysiologiske mekanismer regulerer konsentrasjonen og sammensetningen av disse.

Når kroppen har behov for energitilførsel vil hjernen registrere dette og stimulere sultfølelsen. Ved akutt behov for energi vil vi oppleve en sultfølelse for noe som er søtt, som inneholder karbohydrater. Tilsvarende vil vi oppleve salthunger når kroppen er på saltunderskudd etter en langvarig treningsøkt. Saltunderskuddet registreres i hjernen, som så vil stimulere lysten på mat og drikke som inneholder salt. Hjernens tolkning av smaksopplevelse vil i stor grad påvirkes av den fysiologiske tilstanden vi er i. Begge de nevnte eksempler relateres til mat, drikke og smak. Også andre sanseintrykk påvirkes av vår fysiologiske balanse.

En rekke fysiologiske endringer skjer i kroppen under graviditet. Dette er endringer som bidrar til å optimalisere forhold for fosterutvikling. Disse endringene styres av hormoner som også kan gi seg utslag i sanseopplevelsen. De plutselige hormonendringene som skjer i løpet av de første tre månedene av svangerskapet vil forårsake en forkyving i blodets kjemiske sammensetning. Under utviklingen tilføres fosteret næringsstoffer og salter fra morens blod. Dette kan gi et underskudd på enkelte ioner hos moren, noe som kan føre til en øket «lyst» på mat som inneholder disse ionene. Den gravide kvinnen vil endre sensitivitet for enkelte smaker, noe som kan påvirke hennes pålitelighet som sensorisk dommer.

Også normale hormonvariasjoner som oppstår vil påvirke våre sanser. Stress utløser en rekke kroppslige reaksjoner styrt av nerver og





hormoner, som igjen endrer vår oppfatning gjennom sanseapparatet. I en stressituasjon vil kroppen mobiliseres for å unvike eller løse utfordringen den står overfor. Sansene skjerpes nettopp for å mestre situasjonen, da vil andre sanseinntrykk som smaken av god mat eller synet av et flott bilde bli underordnet.

2.4 Læring og erfaring

Det tyske begrepet «fingerspitzgefulen» ble ofte benyttet i forbindelse med bruk av kunnskap koblet til sansing. Det kan være nærliggende å benytte begrepet i forbindelse med blindeskrift. Hvor en bokstavelig talt benytter sanseceller i fingerspissens hud for å kunne lese blindeskrift. Signalene fra sansecellene tolkes i hjernen som tegn for ord og bokstaver, tilsvarende skjer med sanseimpulser fra øyet hos en seende person. Nettopp skrifttegn er et godt eksempel på at tolking av sanseimpulser er basert på læring og at læringen er individuell. I denne boka benyttes det latinske alfabet og bokstavene og bokstavkombinasjonen tolkes av leseren som ord som settes sammen i setninger og som gir mening. For en person som er opplært i f.eks. det hebraiske alfabet vil tolkingen av skriften være umulig, på samme måte som vi ville ha problemer med å tolke ord fra en hebraisk bok, om vi ikke har lært oss språket.

Enkelte personer har utviklet spisskompetanse koblet til sitt fagområde. En som har spesialisert seg på smaksinntrykk fra f.eks. vin kan med stor presisjon fastslå detaljer (type, druesort, produksjonssted etc.) for en vin ved kun å smake på denne. Smaksopplevelsen for denne personen vil garantert inneholde adskillig flere nyanser enn smaksopplevelsen hos en person uten tilsvarende kunnskap. Men selve opplevelsen av vinen kan være like god.

Alle våre tidligere opplevelser lagres i hjernen og faller inn under begrepet erfaringer. Opplevelsen er individuell og det samme vil erfaringene være. En positiv eller negativ opplevelse ved et restaurantbesøk kan lagres som en erfaring nettopp ved denne restauranten. Denne erfaringen vil være en del av vår referanseramme ved senere valg av spisesteder.





Etter en matforgiftning vil en ofte plassere årsaken til denne inn som en negativ erfaring, og i fremtiden forsøke å unngå den aktuelle råvaren eller det aktuelle spisestedet. Den fremtidige smaksopplevelsen knyttet til den aktuelle råvaren vil deretter ofte være negativ.

Hvilke begrensinger har våre sansesystemer

Sansesystemet vårt er avhengig av at sansecellene fungerer godt, at nervebaner til hjernen formidler sanseintrykkene til de deler av hjernen hvor vi kan tolke dem. Sansecellene har sine begrensinger i terskelverdien, deres evne til tilvenning og forholdene rundt sansecellen. Terskelverdien for en sansecelle er som nevnt tidligere den minste mengde stimuli som skal til for at en sansecelle stimuleres.

Når en sansecelle stimuleres over lang tid vil cellen gå tom for energi og ikke lenger kunne sende impulser til hjernen. Noen få sekunder med kraftig stimuli kan være nok til at cellen går tom for energi. Vi merker det når vi tar hånden ned i varmt vann. Det virker da kanskje svært varmt med det første, men allerede etter kort tid kjennes temperaturen ikke lenger så varm. Temperaturfølerene i huden er tilvennet. Tilsvarende er det når vi skal smake oss frem til en «passe salt» saus. Sansecellene for salt på tungen tilvennes og slutter å sende impulser til hjernen, og opplever at det «aldri» blir nok salt. Det er først når vi smaker på sausen siden at vi kjenner saltsmaken. I mellomtiden har sansecellene restituert seg og er klar for nye stimuli.

De fysiologiske forholdene rundt en sansecelle er viktige for at cellen skal fungere optimalt. Dersom det er for varmt eller for kaldt rundt cellen, vil denne ikke fungere optimalt. Når vi spiser iskem direkte fra fryseboksen (-18 grader C) kjenner vi lite smak. Dersom vi lar isen få temperere seg til den er på smeltetgrensen kjenner vi en tydelig smak. Det samme gjelder når vi smaker på for varm mat. Sansecellene reagerer ikke tilstrekkelig når temperaturen rundt dem er ugunstig. Når maten kjøles kommer smaken frem. De kjemiske forholdene rundt cellene vil også påvirke deres reaktivitet. Når de kjemiske forholdene i kroppen avviker fra normalen føler vi oss uvel og syk. Vi vet alle at i en slik situasjon fungerer sansene våre dårligere.





En objektiv smaksbedømming krever full konsentrasjon slik at kun sanseintrykk som er knytte til smaksbedømmingen prioriteres i hjernen. Det er derfor viktig at en ikke blir forstyrret under en slik smaksbedømming. Opplevelser utenfor dommerbordet vil også kunne påvirke oppmerksomheten. En stresset dag på jobb, familiære problemer eller ferieplanlegging vil alt kunne påvirke konsentrasjonsevnen.

I hjernen vil begrensingen være i form av den kunnskap vi har fylt hjernen med. Det være seg kunnskap gjennom læring eller gjennom erfaring. All tolking av sanseintrykk baserer seg på den informasjon som er lagret i hjernen. Når sanseintrykkene gir en ny opplevelse lagres dette som en ny erfaring og trekkes inn ved senere sansepåvirkning.

2.5 Helse

Vår helsetilstand vil åpenbart påvirke våre sanser, sanseintrykk og opplevelser. Smertesansen er et uttrykk for at noe ikke fungerer optimalt i kroppen, og oppleves som ubehagelig. Rent fysiologisk kan vi peke på to ulike former for smerte, overstimulering av sanseceller eller stimulering av egne smertesanseceller.

Ved ekstrem stimulering av sanseceller vil dette oppfattes som smerte. For eksempel vil en overstimulering av temperatursansen vår oppfattes som smerte og gjerne relateres til begreper som sviende eller brennende. Vi har egne smertesanseceller i kroppen som reagerer på unormale forhold, en murrende smerteopplevelse kan være resultat av en betennelsesreaksjon i kroppen.

Sykdom vil også påvirke våre sanseopplevelser. En forkjølelse, influensa eller en omgangssyke vil alle kunne påvirke så vel sansecellenes reaktivitet som hjernens tolkning av sanseintrykket. Hjernen vil i slike tilfeller ofte fokusere på vår egen helsetilstand og utviklingen av denne. Ved kronisk helsesvikt vil sykdomsopplevelsen kunne bli svært dominerende. En syk person kan miste appetitt, interesser og sosialt engasjement, alle faktorer som i ytterste konsekvens kan ytterligere forverre sykdomstilstanden. Da tilførsel av mat og drikke er en forutsetning for restitusjon vil svekket matlyst kunne forverre tilstanden.





Den aktuelle situasjonen

Den situasjonen vi er i til enhver tid vil være av stor betydning for vår opplevelse av sanseintrykk. I en kjøpsituasjon vil alle sanser være rettet mot opplevelsen av kjøpsobjektet og vi vil i liten grad oppleve andre sanseintrykk. De aktuelle sanseintrykkene sammenliknes med referanserammen og på bakgrunn av dette gjøres en beslutning om kjøp eller ikke. Dersom ønsket er sterkt nok vil vi vektlegge positive sanseintrykk sterkere enn de sanseintrykk som ikke stemmer med referanserammen. Ønsket er ofte definert ut fra et behov som kan være så vel fysiologisk (varme klær for vinteren) som sosialt (ny kjole til selskapet). Behovene bygges inn i referanserammen. I liten grad vil vi la oss affektere av at sanseceller formidler behov for mat, eller andre nødvendigheter mens vi er i kjøpsøyeblikket.

Den gode matopplevelsen kan være så mangt. Det er lett å tenke seg til et nydelig anrettet måltid basert på de beste råvarer som en stor matopplevelse. Imidlertid kan også den enkle brødskenen etter langvarig sult assosieres med en stor matopplevelse. Selv om smaksinntrykkene objektivt sett er svært forskjellige i de to situasjonene kan matopplevelsen være lik. Etter en uke på stram diett endres referanserammen, og kanskje skal det ikke mer til enn en enkel brødsken for å fylle denne. Matopplevelsen blir kanskje den beste.

Den sosiale dimensjonen

Våre sosiale omgivelser er med å forme vår referanseramme og derved våre opplevelser basert på sanseintrykk.

I følge Maslows behovspyramide vil vi ha behov for både sosial tilhørighet og anerkjennelse. Behovet for sosial tilhørighet blir en del av vår referanseramme og vil være av stor verdi i sanseopplevelsen. Det handler om å være en del av en sammenheng og ha et sosialt nettverk, det gir anerkjennelse og en god følelse. Foruten tilhørighet i et bestemt miljø vil den sosiale settingen påvirke vår sanseopplevelse. Middagen blir en helt annen sammen med gode venner enn i et miljø bestående av fremmede mennesker.





For å kunne forme en totalopplevelse bearbeider hjernen mange sanseinntrykk på flere nivå. Sanseimpulsene når hjernen via den forlengede marg og passerer her retikulærsubstansen. Denne delen av hjernen foretar en grovsortering av informasjonsstrømmene og kanaliserte prioriterte sanseinntrykk via thalamus til den bevisste delen av hjernen. Andre sanseinntrykk kanaliseres til ubevisste deler av hjernen og vil der være viktige i hypothalamus sin regulering av grunnleggende fysiologiske prosesser.

Thalamus blir av enkelte omtalt som «porten til bevisstheten». I denne delen av hjernen bevisstgjøres sanseinntrykk uten at de tolkes. Det vil si at vi registrerer en hendelse uten å vite hva denne er. I denne fasen er en på et meget lavt bevissthetsnivå. Når vi konsentrerer oss om en krevende oppgave kan vi registrere andre hendelser i rommet (lyder, bevegelser, lukter) uten at vi tolker hva dette er.

De prioriterte sanseinntrykkene formidles videre til tolkningsområder i storehjernen. Her bearbeides inntrykkene til en opplevelse og hjernen iverksetter nødvendige handlinger i forhold til opplevelsen. Når en er på jakt etter den beste tomaten i kurven bevisstgjøres sanseinntrykk som omhandler de kvalitetsparametre vi har definert for en god tomat. Når vi endelig har identifisert hvilken tomat som ser ut til å svare til forventningene iverksettes handlingen som gjør at vi plukker ut denne tomaten fra kurven. Vi har samtidig skapt en forventning (referanseramme) til smaksopplevelsen og vil vurdere den aktuelle smaken opp mot denne.

De sanseinntrykk som ikke prioriteres for bevisstgjøring kanaliseres til de ubevisste delene av hjernen. Sanseinntrykkene bearbeides ubevisst. Behandlingen innebærer at inntrykkene vurderes opp mot et sett standardverdier som kroppen har satt for et fysiologisk variasjonsområde. Disse standardverdier er relatert til viktige fysiologiske prosesser som regulering av blodtrykk, temperatur og ulike kjemiske verdier i blodet. Responsen på avvik iverksettes av vår ubevisste hjerne og vil ikke påvirke vår bevisste opplevelse.





Når sanseinntrykk som behandles ubevisst kommer langt utenfor det normale variasjonsområdet vil hjernen gjøre oss oppmerksom på dette ved at sanseinntrykket kanaliseres til vår bevisste hjerne. Det er dette som skjer når lukten fra en brennende pizza i steikeovnen når oss mens vi er oppslukt av en god film. Vi bevisstgjøres den kraftige lukten av brennende pizza og iverksetter handlinger basert på dette.

2.6 Sansene

Sansing er en forutsetning for å kunne leve. Det er via sansene vi opplever tørst, sult, varme, kulde, smerte og glede. Sansing skjer via spesialiserte celler på kroppen. Alle sansecellene er utstyrt med en «mottakerdel» gjerne kalt dendritter, en cellekropp som bearbeider sansepåvirkningen og en «senderdel», et akson som transporterer sanse-signalet mot hjernen.

Mottakerdelen i sansecellen er spesialisert for å fange opp definerte sansestimuli. Spesialiserte mottakermolekyl (reseptorer) reagerer på stimulerig. I øyet er mottakerdelen formet som «staver» og «tapper» som fanger opp energien i lyset som treffer øyet. I huden er mottakerdelen formet slik at den kan fange opp bevegelser når noen stryker oss på huden, og på tungen er smaksansecellene spesialisert og organisert på en slik måte at en kan fange opp kjemiske komponenter i maten og drikken som tilføres.

Mottakerdelen i sansecellene reagerer på ulike typer stimuli, og sansecellene er derfor ofte gruppert slik at celler med like egenskaper danner et sanseorgan. Alle sansecellene som reagerer på lys er samlet i øynene som er bygget opp på en måte som gjør at lyset kanaliseres gjennom pupillen mot sansecellene som ligger bakerst i øyet. Sanseceller som er spesialisert på å fange opp luktemolekyler fra omgivelsene våre er lokalisert til neshulen. Denne er igjen bygget opp slik at mest mulig av luften vi puster inn stryker langs luktecellene.

Det er ulike energiformer som påvirker sansecellenes dendritter. Elektromagnetisk energi påvirker synscellene, termisk energi påvirker cellene om registrerer temperaturforandringer mens kjemisk energi



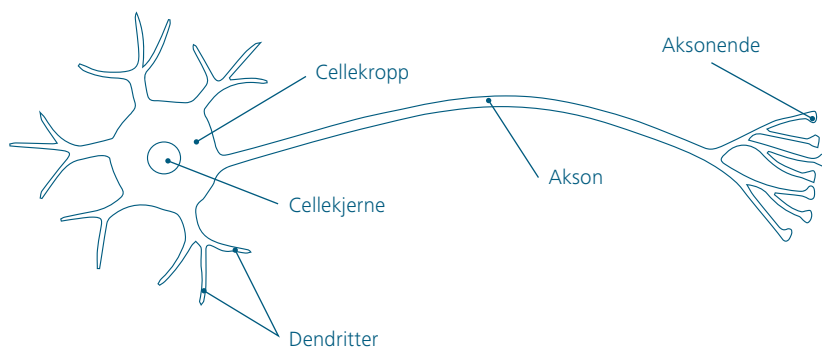


påvirker lukteceller og smaksceller. Vibrasjoner i luftmasser (kinetisk energi) påvirker hørselssansen vår.

Når en sansecelle påvirkes av en energiform oppstår en kjemisk reaksjon i overflaten (membranen) på sansecellen som medfører at det dannes en elektrisk spenning over membranen. Denne spenningen brer seg utover mottakerdelen i form av en liten elektrisk impuls og når etter hvert sansecellekroppen. Ved en sterk stimulering påvirkes dendrittene kraftig og det dannes mange små elektriske pulser. Cellekroppen vil summere mengden av elektriske impulser og dersom disse overskrider en for cellen definert mengde vil cellen generer en nerveimpuls som sendes gjennom aksonet mot hjernen.

Sansecellen har altså satt en grense for hvor sterk stimulering som må til før det genereres en nerveimpuls. Det er dette vi kaller cellens terskelverdi. Rent praktisk kan en tenke seg at en tilfører en liter vann et enkelt sukkorkorn. Når sukkeret løser seg opp er vannet blitt målbart (analyserbart) søtere. Når en drikker dette vannet vil sukkeret utgjøre en svært svak stimulering av smakscellene. Det elektriske signalet som genereres summeres i cellekroppen og da dette er under terskelverdien vil det ikke bli sendt impulser til hjernen og vi kjenner ikke

SANSECELLE



Figur 2.1 Sansecelle

Illustrasjon: Julie Gjermundsen





søtsmaken. Om vi tilfører vannet mer sukker vil stimuleringen av sansecellene øke når vi smaker på vannet. Ved en bestemt sukkerkonsentrasjon blir stimuli sterkt nok til at smakssansecellen sender en impuls til hjernen, og vi opplever en svak sødme. Når vi smaker på ren sukkerlake stimuleres sansecellene så kraftig at de sender mange impulser til hjernen. Vi vil tolke strømmen av impulser og konkludere med en sterk søt smak. På denne måten kan vi via sansecellene oppleve ulike styrker på alle typer stimuli. (Se 3.2, utvelgelse av dommere).

Hvordan kommer sanseintrykkene til vår hjerne?

Fra våre mange sanseceller dannes det en formidabel strøm av nerveimpulser som følger nerver og nervebaner inn mot hjernen. Impulsstrømmen kommer fra ulike sanseceller som registrerer lukt, syn, hørsel, bevegelse samt en rekke forhold inne i vår kropp (blodtrykk, temperatur, blodets kjemiske sammensetning osv.) Utfordringen blir hvordan disse signalene kan mottas og sorteres. Generelt sett kan en si at det er vår vilje og vår bevissthet som avgjør hvilke impulser som skal tolkes til en hver tid. På samme måte gjør vi en bevisst vektlegging av ulike sanseintrykk i våre opplevelser.

Ved inngangen til hjernen (i den forlengede marg) finner vi en stor gruppering nerveceller som påvirkes av vår bevissthet og som også vil kunne påvirke vår bevissthet. Disse cellen utgjør det vi betegner som retikulærsubstansen. Denne delen av hjernen bidrar blant annet til å sortere sanseintrykk samt til å aktivere de deler av hjernen som skal tolke de aktuelle sanseintrykkene. (Moruzzi G. and Maguon H.W. 1949)

Sortering av sanseintrykk styres av vår bevisste hjerne. Når vi bestemmer oss for å lese en bok, vil retikulærsubstansen påvirkes slik at synsinntrykk fra boka prioriteres og kanaliseres til vår bevisste hjerne. Alle andre sanseintrykk vil da nedprioriteres og bli behandlet ubevisst. Det er en av grunnene til at når vi er dypt inne i en spennende bok ikke får med oss hva som skjer rundt oss. Det samme skjer dersom vi konsentrerer oss om en vanskelig oppgave, retikulærsubstansen prioriterer de sanseintrykk som er forbundet med oppgaven og nedprioriterer alle andre sanseintrykk. Vi sier at vi konsentrerer oss.





Å opprettholde konsentrasjonen kan være vanskelig, ofte kommer tankene på «avveier», mens vi forsøker å konsentrere oss. En Med virkende årsak til dette kan være at saken vi forsøker å konsentrere oss om ikke engasjerer oss. Signalene fra vår bevisste hjerne til retikulærsubstansen svekkes. En slik svekkelse kan føre til at andre ikke-prioriterte sanseinntrykk formidles til vår bevissthet og «stjeler» vår oppmerksomhet. Jeg forsøker å konsentrere meg om et tungt dokument som jeg skal ta stilling til. Dokumentet inneholder opplysninger som i liten grad engasjerer meg. Middagslukten fra kjøkkenet derimot når min bevisste hjerne via luktesansen. I en slik situasjon kan fort middagslukten «stjele» hele min oppmerksomhet, mens synsinntrykkene (ordene) fra dokumentet nedprioriteres. Jeg vet at jeg har lest dokumentet, men jeg husker ikke innholdet og mitt beslutningsgrunnlag i forhold til innholdet er derved svekket.

I de to nevnte eksemplene, spenningsboka og det tunge dokumentet engasjeres vi på forskjellig måte. Spenningsboka gir oss en opplevelse vi vil ha mer av og vi forsterker påvirkningen av retikulærsubstansen. Det tunge dokumentet derimot gir oss ikke i samme grad en positiv opplevelse og påvirkningen av retikulærsubstansen svekkes. En kan si at konsentrasjonen forsterkes eller svekkes som følge av den opplevelsen stimuleringen gir.

For dommere i et sensorisk panel vil det være svært viktig å opprettholde konsentrasjonen slik at en kan gi best mulig tilbakemelding under bedømming. Opptrening av dommere er også en trening i å konsentrere seg, og derfor svært viktig når en skal i gang med større bedømminger.

Tilvenning (adapting) er også en funksjon som er felles for sansecellene. Etter langvarig påvirkning vil sansecellen slutte å reagere på en stimulus. En forklaring på dette kan være at cellen har tappet sine energireserver og derfor ikke lenger er i stand til å reagere. Dette er grunnen til at en litt for trang bukse føles trang når en tar den på seg, mens en etter hvert ikke merker hvordan buksa strammer.

I et sensorisk panel for objektiv vurdering av smak og lukt vil tilvenning skje dersom en påvirker sansecellene over for lang tid. Det er





derfor viktig at en smaksprøve ikke blir liggende for lenge i munnhulen eller under nesa. Tilvenningen vil påvirke helhetinntrykket og kunne gi en feil vurdering.

I det følgende skal vi gå gjennom de ulike sansekvaliteter i kombinasjon med de spesifikke sanseorgan.

2.6.1 Syn

Synet er et resultat av lysbølger som treffer sansecellene i øyet. Vår største lyskilde er solen, det er dens energi som representerer lys og varme. Lyset fra solen består av lysbølger som spenner over et stort spekter av bølgelengder. En del av disse bølgelengdene representerer energi som påvirker våre sanseceller i øyet. Det er det vi kaller den synlige delen av lyset. Utenfor dette området stråler solen ut usynlig lys i form av ultrafiolette stråler samt infrarøde stråler.

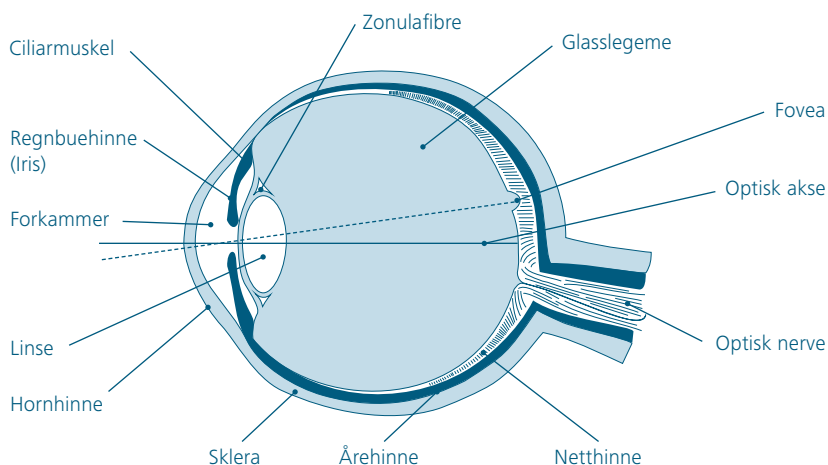
Den synlige delen av lyset vil samlet sett fortone seg som hvitt – og dersom vi ser direkte på solen vil vi oppleve den som hvit. Lyspæren skaper energibølger i det synlige spekteret og vil fortone seg som hvitt lys. Når lyset treffer en hvit flate vil denne reflektere alle lysstrålene og derfor vil vi oppleve flaten som hvit. Dersom vi tilsetter farge på den hvite flaten, en sort tegning på et hvitt ark, vil den sorte fargen absorbere alt lyset. Det betyr at det fra dette området ikke reflekteres lys. Rent praktisk opplever vi det derfor som en sort tegning på et hvitt ark. Dersom vi lager tegningen i rødt vil det fargede området absorbere alle farger unntatt rødt. Det er altså kun de bølgelengdene som representerer rødt som reflekteres og som når øyet. Vi opplever tegningen som rød. Alle fargenyanser vi ser er et resultat av en flate som absorberer deler av det synlige lyset samtidig som en del reflekteres.

Øyet mottar altså lysstråler med ulike bølgelengder som i ulik grad påvirker sansecellene. Sansecellene i øyet er av to typer, de som reagerer på spesifikke bølgelengder (tapper) og de som reagerer på lysstyrke (staver). I godt lys vil øyets pupille slippe gjennom en liten lysmengde som via øyets linse projiseres til den gule flekk bakerst i øyeeplet. I dette området er det en stor konsentrasjon av tapper. Det er tre typer tapper, de som er spesialisert på grønn farg, de som er spesialisert på rød





ØYET



Figur 2.2 Øyet

Illustrasjon: Julie Gjermundsen

farge og de som er spesialisert på blå farge. Alle disse cellene vil reagere på ulike bølgelengder, men sterkest på henholdsvis grønn, rød og blå. Dette betyr at når en gul gjenstand reflekterer lyset vil de ulike tappene stimuleres ulikt, og hjernen mottar ulik stimuli fra celletypene. Basert på stimilmengden vil da hjernen tolke og oppleve den gule fargen. Rød og grønn bølgelengde ligger relativt nær hverandre og de to aktuelle typene tapper reagerer nesten likt på fargene. For enkelte personer kan dette føre til vansker med å skille fargene – en sier gjerne at personene er fargeblinde. Rød/grønn fargeblind er det vanligste. Hos enkelte personer vil de tappene som reagerer på grønn farge og de som reagerer på blå farge ligge nær hverandre i reaksjonsmønster. Dette kan medføre blå/grønn fargeblindhet.

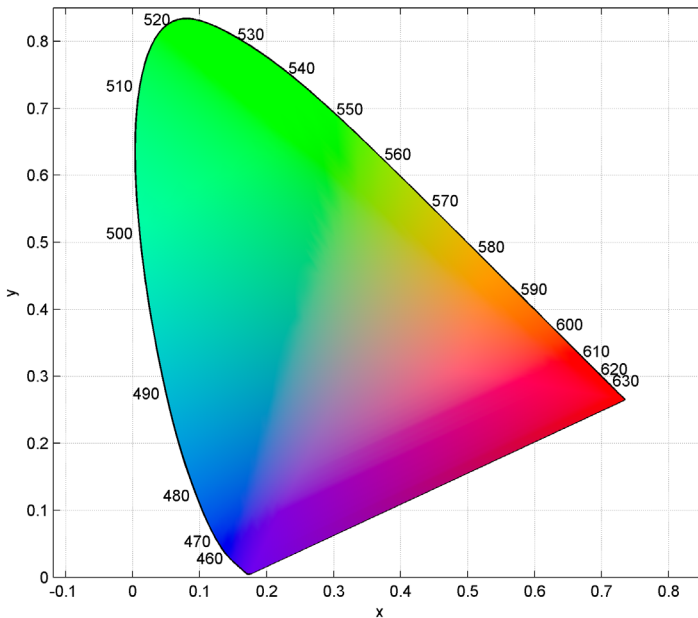
I og med at tappenes reaktivitet styres av våre gener og er personavhengig, kan en forstå at farger oppfattes forskjellig. Nettopp det at en ser farger forskjellig kan bidra til mange gode diskusjoner når fargevalg skal gjøres.





I fravær av lys vil også stimuleringen av sansecellene i øyet opphøre og vi opplever totalt mørke. I dårlig lys vil pupillen utvides for å slippe mest mulig lys inn mot sansecellene. En forholdsvis større andel av lyset vil da treffe «stavene». Disse sansecellene reagerer ikke spesifikt på bestemte farger, men på lys i seg selv og det medfører at vårt fargesyn svekkes i dårlig lys. Som følge av den utvidede pupillen vil vi imidlertid se en større bredde i omgivelsene foran oss.

FARGESPEKTER



Figur 2.3 Fargespekter. Den buede linjen viser bølgelengder i nanometer med de rene spektrale fargene.

Illustrasjon: Bang & Olufsen





Som en følge av at «tappene» som stimuleres av farget lys er konsentrert i et punkt på øyets netthinne, vil alltid fargesynet være best i et lite konsentrert synsfelt. Alt lys som kommer inn fra siden gjennom pupillen vil treffe «stavene» og derved oppfattes i gråtoner. I skumring og mørke vil pupillen åpne seg og mer lys slipper inn. Dette for at vi overhode skal kunne se i dårlig lys. I og med at pupillen er utvidet vil en større mengde av lyset treffe den delen av øyet hvor stavene sitter, dette er grunnen til at vårt fargesyn i dårlig lys er svekket. Tilvarende vil pupillen kontrahere i sterkt lys, og nesten alt lyse treffer den delen av øyet hvor tappene sitter. Derfor er fargesynet vårt best i godt lys.

2.6.2 Hørsel

Lyd er trykkbølger i luft. Når en gjenstand settes i bevegelse vil den danne fine bølger i luften. Luftbølgene treffer kroppen vår og større eksplosjoner gir gjerne kraftige trykkbølger som vi merker på kroppen. Det samme kan skje fra en høyttaler når volumet er høyt.

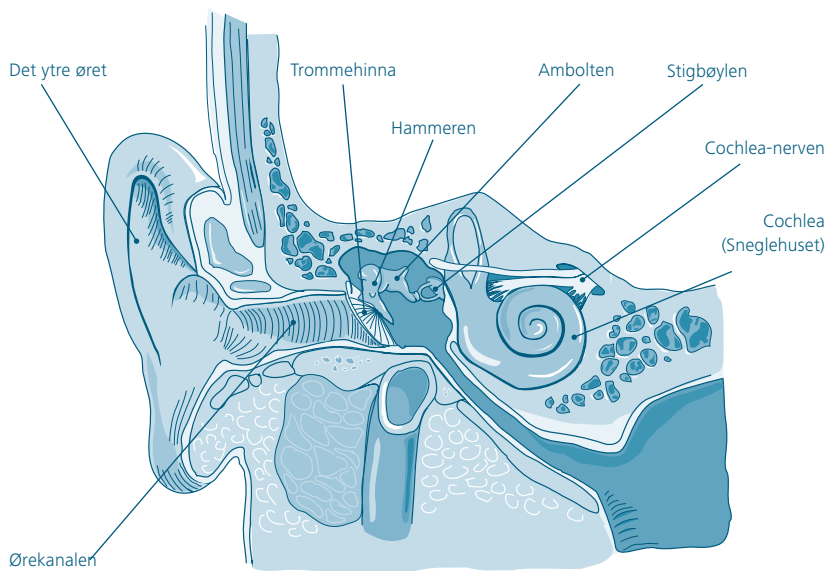
Trykkbølger som treffer kroppen vil også treffe øret og via dette kunne treffe trommehinnen som danner inngangen til mellomøret. Trykkbølgene fører til at trommehinnen som er elastisk svinger i takt med trykkbølgen. Svingningene i trommehinnen forplantes via tre små bein i mellomøret og når på den måten vårt indre øret. Det er her hørselen sitter. Trykkbølgene som forplanter seg gjennom mellomøret fører til at det oppstår «bølger» i væsken i det væskefylte sneglehuset som forplanter seg til hørselcellene. Hørselcellene stimuleres av bølgene som forplanter seg i væskefasen.

En trykkbølge kan beskrives ut fra frekvens og amplitude. Frekvensen beskriver hvor mange svingninger eller bølger i løpet av en tidsenhet. Dess flere svingninger dess «lyser» lyd. På et piano vil tangentene i øverste høyre ende være koblet til små tynne metallstrenger som vibrerer svært fort når vi slår på dem. Dette gir en meget høy frekvens og vi opplever det som en høy diskanttone. I den andre enden av pianoet er tangentene koblet inn mot tykkere og større metallstrenger. Et trykk på tangenten her gir en langsommere vibrasjon og derved færre lydsvingninger pr tidsenhet. Vi opplever lyden som dyp bass. Hver enkelt tone i skalaen representerer en bestemt frekvens. En person med godt



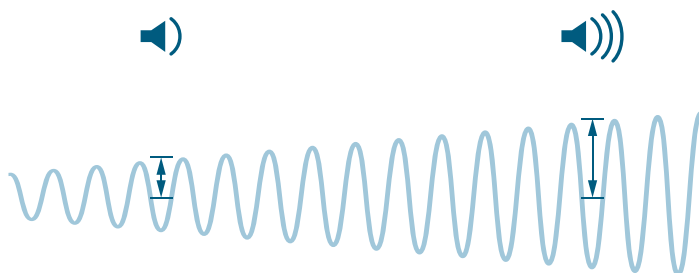


ØRET



Figur 2.4 Snitt av øret

Illustrasjon: Ukjent



Figur 2.5 Lydbølger

Illustrasjon: Julie Gjermundsen





gehør har trent opp øret til å kunne definere de ulike frekvensene i forhold til skalaens toner.

Amplituden beskriver «størrelsen» på lydbølgen eller det vi ofte forbinder med lydvolume. Dersom en slår hardt på en tangent på pianoet vil vi oppleve lyden kraftig, mens et lett trykk på tangenten fører til svingninger som er små i størrelse og vi oppfatter dette som en svak lyd.

Når vi snakker danner vi lydbølger ved hjelp av stemmebåndene. Disse strammes og slakkes for å kunne varierer frekvensen, mens kraften vi presser luften forbi dem på avgjør stemmevolumet. Når vi strammer stemmebåndene blir stemmen lys (høy frekvens), mens slakke stemmebånd fører til lav frekvens og basstemme. Ved å presse luften kraftig forbi stemmebåndene dannes en kraftig lyd (et brøl), mens en forsiktig luftstrøm gjennom stemmebåndene fører til en svak lyd (visking). Lydbølgene som når trommehinnen via det ytre øret medfører bevegelser i hinnen. Disse bevegelsene speiler lydbølgene på den måten.

Svært mange av trykkbølgene som dannes av gjenstander i bevegelse inneholder så lite energi at vi ikke merker det fysiske trykket på kroppen. Vårt ytre øre er formet som en trakt og vil fange opp svakere bølger og trykkbølgefrekvenser. Luftsvingningene kanaliseres gjennom det ytre øret inn mot trommehinnen.

Stimulering av lydcellene i det indre øret (sneglehuset) fører til en endring i den elektriske spenning over sansecellen og et lite elektrisk signal dannes. Cellekroppen summerer de elektriske signalene og dersom disse overstiger cellens terskelverdi sendes en elektrisk impuls til hjernen som tolker lyden, og gir oss en opplevelse av denne. Dess sterkere lyden er (volum) dess kraftigere stimuleres hørselcellen. Når denne stimuleres kraftig dannes en serie elektriske nerveimpulser som når hjernen. Vi tolker da lyden som sterk (høyt volum).

Når vi hører musikk stimuleres flere hørselceller avhengig av tonehøyde. Dette medfører signaler fra ulike hørselceller som når hjernen, som tolker lydene. Vi opplever musikk. Det er vår kjennskap til musikk, komponister og musikkstykker som gjør oss i stand til å tolke musikken. På samme måte vil ord som uttales medføre trykkbølger





i luften, fanges opp av øret, stimulere sansecellene og formidles til hjernen. Vi tolker gjerne ord i den sammenheng det blir sagt, dette kan være grunnen til at samme ord kan ha flere betydninger. Ordet «ved» kan beskrive en plassering (ved bordet). Samme ord «ved» kan beskrive brensel (bjørkeved).

Lyd påvirker i sterk grad også opplevelsen av andre sanseintrykk. Filmskapere benytter seg gjerne av dette når en lydsetter filmen. I spennende scener legges gjerne en lydforsterkende effekt inn. Dersom vi fjerner lyden fra spenningsfilmen vil opplevelsen bli en helt annen.

På samme måte vil lyd forsterke et smaksinntrykk. Når vi tar en bit av et sprøtt eple eller et knekkebrød vil dette forsterke opplevelsen av konsistensen. Tilsvarende lydinntrykk er fraværende når vi spiser loff eller boller. I enkelte restaurantkjøkken spiller en på denne effekten ved å ha et åpent kjøkken inn mot restauranten. Lyden av steking, kutting av grønnsaker og andre kjøkkenlyder forsterker kundens opplevelse av måltidet i likhet med at taffelmusikk har en virkning på måltidsopplevelsen.

2.6.3 Smak

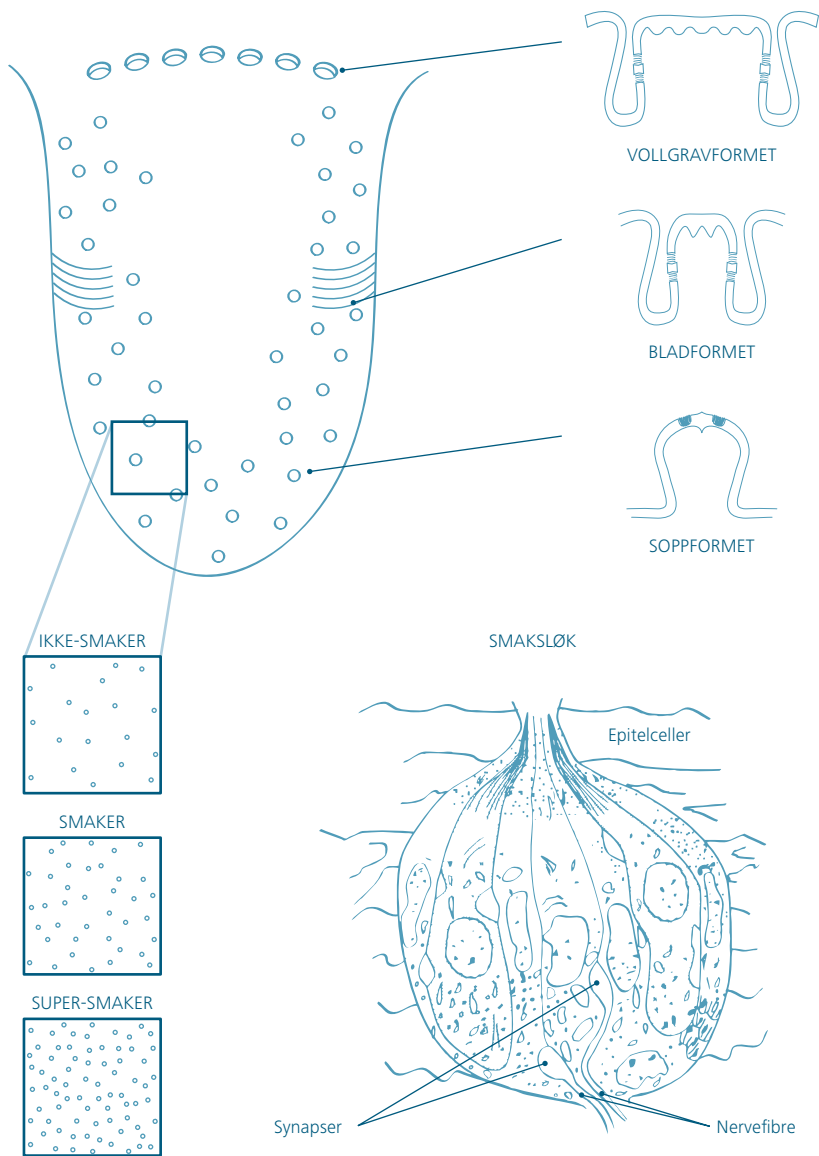
Smakssansen og luktesansen utgjør en kjemisk sanseegenskap. Det betyr at sansecellene stimuleres av kjemiske komponenter (ioner og molekyler). Smakssansen sitter i munnhulen og blir først og fremst påvirket av alt som tilføres munnhulen, mens luktesansen ligger i neshulen og påvirkes av kjemiske komponenter i luften som passerer. Felles for disse sansene er at sansecellene har reseptorer som reagerer på kjemiske forbindelser. Vi kaller gjerne disse reseptorene for kjemoreseptorer.

I internasjonal litteratur benyttes ofte begreper som *flavour*, *taste*, *smell* og *odour* for å beskrive en smak og lukteopplevelse. Det er vaskelig å oversette disse begrepene direkte da begrepet *flavour* omfatter mer enn bare smaken mens begrepet *odour* omfatter mer enn bare lukten. Stevenson påpeker at *flavour* er kombinasjonen av smak (grunnsmaker) «somatosensation» (fysiske sansevariable) og lukt (Stevenson R.J. 2012). *Taste* oversettes til ordet smak og relateres gjerne til grunnsmakene.





TUNGEN MED ULIKE PAPILLER



Figur 2.6 Tungen med papiller og smaksløk

Illustrasjon: Julie Gjermundsen





Smell oversettes med lukt og relateres til luktesansen. Det samme gjør ordet *odour*.

Hva er egentlig smak? Vi bruker begrepet om de sanseintrykk som formidles via munnhulen. Rent fysiologisk vil en da identifisere 5 grunnsmaker samt sanseintrykk som formidler opplevelse av bl.a. temperatur, fuktighet, konsistens og fylde. Mens grunnsmakene kan knyttes til dedikerte kjemoreseptorer på sanseceller, vil de øvrige sansene formidles av celler som reagerer på temperatur, trykk og bevegelse. En overstimulering av disse sansecellene vil ofte oppleves som ubehag eller smerte.

I munnhulen finner vi en rekke kjemoreseptorer på tungen og bak i svelget. På tungen er sansecellene organisert i smaksløker. Smaksløkene består av en gruppe støtteceller som danner små forhøyninger (papiller) på tungen. En kanal på utsiden av støttecellene fører ned til sansecellene som ligger beskyttet like under tungeoverflaten. Kjemiske komponenter fra mat og drikke ledes ned i kanalene og påvirker derved smakscellene. Spyttsekresjonen i munnen bidrar til å løse opp maten og gjøre de kjemiske komponentene tilgjengelige for sansecellene.

Det har i lengre tid hersket en oppfatning av at smakssansen baseres på 5 grunnsmaker; salt, søtt, surt, bittert og umami. Grundige studier av reseptorene på smaksansecellene har bekreftet at disse reagerer nettopp på kjemiske forbindelser som representerer disse smakene. En kan ikke utelukke at videre studier vil identifisere flere grupper reseptorer som reagerer på andre smakskvaliteter. Det diskuteres om mennesket også har egne smaksreceptorer for kalsium (reed D.R. and Knaapila A. 2010).

Vi har ulike typer smaksløker plassert på ulike steder på tungen. Tidligere antok en at de ulike smaksløkene var knyttet til ulike smakskvaliteter. I dag er en av den oppfatning at ulike smakskvaliteter kan registreres over store deler av tungen, og en antar også at flere smakssanseceller finnes i svelget.

Foruten matens kjemiske innhold vil også matens konsistens sanses i munnhulen. Vi kjenner forskjell på flytende og fast konsistens, og vi sanser forskjeller mellom tyntflytende (vann) og seigtflytende (suppe).





Sansecellene registrerer også partikler og partikkelstørrelse i maten vi spiser. Informasjon om konsistens og ikke minst temperatur som formidles via temperaturfølere i munnhulen er viktig for vår opplevelse av smak. (Shepherd G.M. 2012)

Grunnsmaken salt er i hovedsak knyttet til at cellene stimuleres av natriumklorid ($\text{Na}+\text{Cl}^-$). Kroppen vår er svært avhengig av å opprettholde en stabil balanse mellom vann- og salt. Dersom balansen forskyves vil en rekke kroppslige prosesser få nedsatt funksjon. Sansecellene på tungen vil holde hjernen orientert om inntak av salt.

Søtsmak er smaken av små karbohydrater, sukkerarter. Karbohydratene er en viktig energikilde for kroppen. Smaken av sødme vil derfor oppfattes positivt for kroppen.

Mat og drikke inneholder syre i ulike mengder. Frukt kan inneholde mye syre, når vi spiser sur frukt sanses dette og hjernen iverksetter en økt spyttsekresjon for å «nøytralisere» syren før denne svelges ned.

Bitre smakskomponenter er ofte knyttet til giftige kjemiske forbindelser. Vi finner den bitre smaken i salater og f.eks. i grapefrukt. Den bitre smaken fungerer som en «advarsel» og små barn vil gjerne brette seg om de smaker noe bittert. Vi lærer oss til å tolerere de bitre komponentene og i mange tilfeller vil vi også sette pris på disse i matretter og drikke.

Umami beskriver «god smak» på japansk. Umami er smaken av proteiner og vi kan oppleve den som smaken av buljong. Det er i særdeleshet glutaminsyre og asparbinsyre fra proteiner som stimulerer denne smaks kvaliteten. Når matvarer som f.eks. tomater og ost modnes, brytes en del proteiner ned og glutaminsyren omdannes til glutamat. Mat med mye glutamat oppfattes som velsmakende, dette er grunnen til at en har tilsatt glutamat i form av (natriumglutamat) til mat.

Smaksblind

Noen har manglende genetisk evne til å kjenne ulike smaker. Spesielt godt beskrevet er manglende evne til å registrere enkelte bitre smaker. Phenylthiocarbamid (PTC) og 6-n-propylthiouracil (PROP) kan gi en





bitter smaksopplevelse. Genetisk er vi mennesker forskjellig utstyrt og denne smaken vil oppleves som spesielt sterk hos noen, men andre kjenner en svak smak og andre igjen ikke kjenner noen bittersmak. En kan si at enkelte mennesker er «smaksblinde» for enkelte smakskomponenter (Bartoshuk L.M. 2000)

Spyttets betydning for frigjøring av smak og luktemolekyl

Spyttets sammensetning er av stor betydning for smak og smaksopplevelse. Dette merker vi svært godt når vi er «tørr» i munnen. Da er det lite smak i maten. Spytt kan grovt sett deles inn i to kategorier, det mukøse spyttet bestående av en stor mengde slim og det serøse spyttet som i liten grad inneholder slim. Det mukøse spyttets funksjon er i første rekke å lette svelgprosessen. Spyttet frigjøres som en følge av tyggebevegelsene og «smører» matbollen som formes i munnhulen. Spyttet vil binde partiklene sammen og på den måten kunne «maskere» vekk en del av smaksstoffene. Spyttet inneholder også enzymer som bidrar til den kjemiske nedbrytingen av næringstoffer. Samtidig vil matens konsistens i stor grad både påvirke sekresjonen av spyttet samt bidra til en opplevelse av konsistensen, som en dimensjon av smaksopplevelsen.

Det serøse spyttet inneholder en rekke salter og vann. I hovedsak kan en gjerne si at dette spyttet bidrar til å fukte maten samt «nøytralisere» den. Ved å tilføre maten fuktighet frigjøres en større andel kjemiske komponenter som i neste omgang vil stimulere smakscellene. Nøytraliseringen av matens innhold vil også påvirke sansingen av surhet i maten. På denne måten vil spyttet og dets sammensetning i stor grad ha innflytelse på smak og lukt.

2.6.4 Lukt

Fra alle våre omgivelser avgis molekyler til luft. Molekylene diffunderer gjennom luften og følger innåndingsluften inn i nesehulen. En rekke molekyler fester seg der til slimhinnene og transporteres gjennom disse til luktecellene. Luktecellene er spesialiserte sanseceller som stimuleres av ulike kjemiske stoffer som finnes i slimhinnen.





Luktesansen tillegges flere funksjoner og lukt er gjennom disse en svært viktig lukteegenskap. Mange åpenbare farer kan identifiseres gjennom luktesansen, røykutvikling og brann er nærliggende å trekke frem. Skadelige gasser og kjemikalier avgir ofte lukt som vi tolker som et faresignal. Lukt kan også bidra til å skille mellom bedrevet og fersk mat, og beskytter oss dermed mot f.eks. matforgiftninger. Mennesket skiller også ut luktstoffer som påvirker oss. I dyreriket ser en gjerne at ulike arter markerer revir ved hjelp av luktestoffer.

Lukter knyttes også sterkt til vårt minne. Lukter kan gi oss mange assosiasjoner knyttet til opplevelser. Lukter av desinfeksjonsmiddel gir oss assosiasjoner til besøk hos tannlege eller lege. At lukt også påvirker appetitt kjenner vi gjennom den påvirkning duften av ferskt bakverk eller annen mat bidrar til å øke sultfølelsen. På samme måte merker vi at ubehagelig og fremmed lukt gjerne bidrar til å redusere appetitten.

Luktesansen er lokalisert til neshulen og en omtaler gjerne den delen av nesen hvor luktesansen sitter som neshulens lukteepitel. Små fine utløpere (dendritter) fra luktesansecellene ligger i slimhinnene beskyttet av et tynt slimlag. Luktemolekylene når reseptorene på dendrittene når de diffunderer gjennom slimlaget. Luktesansen vår avhenger av størrelsen på lukteepitelet, på slimlagets tykkelse og hvilke reseptorer som finnes på sansecellene. Lukteepitelet utgjør rundt 5 cm² hos mennesket, tilsvarende hos en katt er 25 cm². Det betyr at luktesansen vår er dårligere enn hos en katt. Slimlagets tykkelse varierer, når vi er forkjølet eller irriteres av allergene stoffer blir gjerne slimlaget tykkere. Dette for å begrense mengden «støv» som passerer forbi nesen og ned i luftveiene. Når slimlaget er fortykket som under forkjølelse vil en mindre mengde luktestoffer diffundere inn til reseptorene. Lukteinntrykket blir derved svakere enn når vi er friske.

I forbindelse med mat og drikke, vil også ofte førsteinntrykket av lukter nå lukteepitelet via neseborene, en snakker da gjerne om en orthonasal vei. Imidlertid vil en stor andel luktemolekyler nå neshulen via munnen og svelget. Passasjen fra svelget til neshulen (Rhizopharynx) fungerer som transportvei for disse molekylene. Dette skjer hovedsakelig ved utpust. I og med at molekylene beveges fra



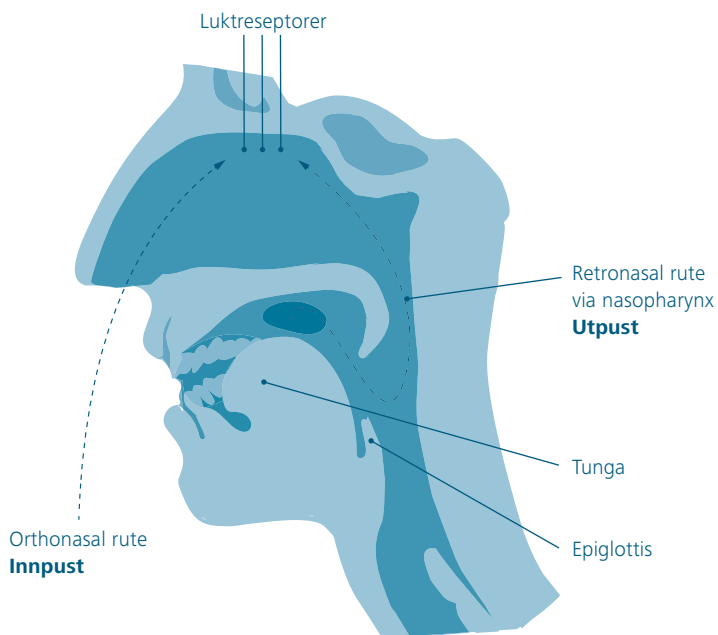


munnhule/svelg i retning nesehule omtales lukteveien som retronasal. (Shepherd G.M. 2012)

Når mat bearbeides i munnhulen vil matens temperatur stabiliseres i forhold til kroppstemperaturen. Varm mat vil kjøles ned og kald mat vil varmes opp. Endringer i temperaturen medfører også frisetting av nye molekyler som utgjør en del av smak- og luktesansingen. Når vin

En regner at vi mennesker kan skille mellom rundt 10000 forskjellige lukter, dette er i sterk kontrast til de fem grunnsmakene knyttet til smakssansen. Selv om vi klarer å skille mellom et stort antall lukter, klarer vi nok bare å identifisere et fåtall med ord. Dette er på samme måte som fargesyn, vi kan se et utall farger uten å ha ord for alle fargenyansene.

NESE OG SVELG



Figur 2.7 Svelget med innpust og utpust

Illustrasjon: Julie Gjermundsen





holdes i munnen varmes denne opp og nye forbindelser går over i en flyktig fase og når derved lukteepitelet retronasalt. Dette bidrar til at nye lukteinntrykk dannes over tid samt en ettersmak.

Lukteblindhet

Manglende evne til å kjenne lukter beskrives ofte som resultat av sykdom eller skader. Det betyr gjerne at «lukteblindheten» kan være midlertidig dersom årsaken er en forbigående sykdom. Ved kronisk sykdom eller skade kan «lukteblindhet» være vedvarende. Også genetiske forhold er avgjørende for luktesansen. Individuelle variasjoner i oppfatning av bl.a. «Rånelukt» tilskrives gjerne genetiske ulikheter. Således er det påvist at luktereseptorer for Androstenone som er en viktig komponent i «rånelukt» opptrer i ulik mengde og aktivitet hos personer. (Wysocki C.J. and Beuchamp G.K. 1984). Det er pågående diskusjoner om en kobling mellom våre omgivelser og genenes aktivisering. Teorier er fremsatt om at luktesansen viser en evolusjonistisk utvikling, noe som kan tyde nettopp på dette. (Reed D.R. and Knaapila A. 2010)

Referanser

- Sander K. 2014, Hvordan tolker vi sanseinntrykkene (Persepsjon og kognisjon). Kunnskapssenteret.com
- Moruzzi G. and Magoun H.W 1949, Brain stem reticular formation and activation of EEG. *Electroencephalography and Clinical Neurophysiology* Vol 1, issue 1 – 4, pp 455 -473
- Shepherd G.M. 2012. *Neurogastronomi*, Columbia University Press, kap 1
- Reed D.R. and Knaapila A. 2010, Genetics of taste and Smell; Poisons and Pleasures. *Prog. Mol Biol Transl Sci* 2010; 94: 213-240.
- Wysocki C.J. and Beuchamp G.K. 1984: Ability to smell androstenone is genetically determined. *Proc. Natl. Acad. Sci* Vol 81 pp 4899-4902
- Bartoshuk L.M. 2000, Comparing Sensory Experiences Across Individuals: Recent Psychophysical Advanced Illuminate Genetic Variation in Taste perception. *Chem Senses* 25: 447-460,
- Stevenson R.J. 2012. The role of attention in flavor perception. *Flavour* 1:2



RETNINGSLINJER FOR GJENNOMFØRING AV SENSORISKE ANALYSER

Liv Bente U. Strandos, Elopak AS

For å sikre konsekvente og pålitelige data må gode standardprosedyrer følges. Når man skal planlegge en sensorisk analyse eller studie, kan det være hensiktsmessig å etablere en sjekkliste. Sjekklisten presentert her er hentet fra *Sensory evaluation of food – Lawless and Heymann (1999)* og kan være til hjelp i planleggingsfasen. Hovedpunktene i denne sjekklisten vil bli gjennomgått i detalj i dette kapittelet.

Tabell 3.1 Sjekkliste

Hovedpunkter for analysen	Viktige resursdetaljer og andre avgjørende detaljer for gjennomføring	Kryss av
Hensikt med forsøket	Eks. produktutvikling, holdbarhetsstudie, kvalitetskontroll	
Forsøksdesign	Antall produkter, intervall, pauser etc.	
Valg av metode	Differansetest, beskrivende analyse, spørreskjema-design	
Dommere / respondenter	Gjennomgang av detaljer rundt dommere	
Rekruttering (hvem, hvordan, når)	Detaljer	
Utvelgelse	Valg av de riktige respondentene i forhold til produkt, bruk etc. om nødvendig	
Trening	Om nødvendig for den aktuelle test	
Prøvepreparering og serveringsprosedyre	Er nødvendige ressurser og utstyr på plass?	
Prøvestørrelse	Detaljer	



Prepareringsmetode	Detaljer	
Temperatur	Fryst, kjølig, romtemperert, oppvarmet?	
Prøvepresentasjon	Detaljer rundt den praktiske gjennomføringen	
Skyllevann og/eller usaltet kjeks for å rense munnhulen	Om nødvendig	
Orientering og instruks til dommerne	Detaljer	
Koding av prøver	Detaljer	
Svar ark eller design av test i relevant software	Detaljer	
Tid og lengde for analysen	Detaljer	
Testfasiliteter	Internt/eksternt	
Lokalitet	Detaljer	
Dommerbåser	Detaljer	
Utstyr i dommerbåsene	Detaljer	
Prepareringsrom	Detaljer	
Temperatur, lys etc.	Detaljer	

Eksempler på hvorfor en bedrift har en sensorisk avdeling er at den:

- Tilbyr kvantitativ informasjon om bedriftens egne, samt konkurrentenes produkter.
- Tilbyr nyttig og verdifull informasjon om produkters sensoriske egenskaper.
- Utvikler både generelle og produktspesifikke metoder.
- Sikrer stabil og god sensorisk kvalitet av bedriftens produkter.





Foto: Kjell Merok





3.1 HOVEDPUNKTER FOR ANALYSEN

Hensikten eller målsetningen med forsøket defineres av de involverte personer tilhørende sensorisk avdeling. Ordlyden skal være klar og forstått av alle og inkludere bakgrunnsinformasjon for å danne kontekst for forsøket/studiet.

Forsøksdesign

Forsøksdesign og serveringsrekkefølge må tas hensyn til i den innledende design av alle studier innen sensorisk analyse.

Nedenfor vises noen av de mest anvendte serveringsdesign som brukes i sensorisk analyse.

Monadisk versus sekvensiell design

- Monadisk serveringsdesign: en respondent evaluerer en prøve.
- Sekvensiell monadisk design: en respondent mottar mer enn en prøve. Prøvene serveres en av gangen, men først etter at den forrige prøven er evaluert. Denne prosessen kan repeteres så mange ganger som nødvendig for å fullføre det eksperimentelle forsøket. Det må vel og merke tas hensyn til faktorer som adaptasjon, tretthet og tid tilgjengelig.

Monadisk design er nyttig når:

- Det er for tidkrevende for respondenter å teste mer enn ett produkt, eller
- Testingen av et spesielt produkt kan ha betydning for oppfatningen av et annet produkt, eller
- Testingen av et produkt kan endre testmiljøet, person eller prøven (f.eks. maling, lakk, personlig hygiene produkter).





Komplett, balansert design

- Alle prøver er servert i hver enkelt posisjon i prøveoppsettet et likt antall ganger. I tillegg er prøvene balansert på tvers av sesjoner og respondenter for å forsikre en minimal rekkefølgeeffekt. I et komplett, balansert forsøk, er antall par også tatt hensyn til og tilstrekkelig balansert. I detalj betyr det at Prøve 1 er servert før og etter Prøve 2 like mange ganger, det samme gjelder Prøve 3 osv.
- Slike design anbefales når respondentene ser alle prøvene, slik som i sekvensiell monadisk design, og når det er repetisjoner.
- Ved bruk av trenet panel, bør prøvene evalueres mer enn en gang av hver dommer (repetisjon) for å fastslå dommerreperbarhet.

Randomisert, komplett design

Prøvene er tilfeldig tildelt hver dommer innenfor en sesjon. Dette designet brukes gjerne når trente dommere ikke kan evaluere alle prøvene i en sesjon; hvilket kan kreve flere sesjoner for å få fullføre alle prøvegentakene.

Ukomplett design

Når det er for mange prøver i et eksperiment for respondentene å håndtere på en fornuftig måte, må man teste et utvalg av prøvene. Dette krever flere respondenter for å fullføre samme antall evalueringer som når man bruker et komplett, balansert forsøk.

Test design blir mer utfyllende omtalt i kapittel 11.

Valg av riktig metode

Valg av riktig metode er avgjørende i planleggingsfasen. Noen ganger kan det være flere målsetninger i et forsøk eller studie. Hvordan velge riktig metode og en grundig innføring i ulike metoder er beskrevet i kapittel 4.





3.2 SENSORISKE DOMMERE – REKRUTTERING OG UTVELGELSE

Den gruppe kvalifiserte dommere som skal utføre en sensorisk analyse kalles et panel. Hvilket panel som skal benyttes, avhenger av hvilke(n) oppgaver som skal løses. Avhengig av behov, kan en bedrift eller organisasjon ha tilgjengelig ett eller flere panel.

Det finnes tre typer panel:

1. Ekspertpanel

Den mest avanserte form for panel består av noen få høyt trene dommere med inngående produktkunnskap. Opplæringstiden er som regel lang og produktspekteret relativt snevert.

2. Laboratoriepanel.

Den mest brukte formen for panel består av fem til tjue dommere, men det mest vanlige er seks til ti dommere. De rekrutteres enten blant egne ansatte (internt panel) eller fra personer utenfra (eksternt panel). En bør unngå å benytte personer som i for sterk grad er knyttet til produksjonen av produktet som skal bedømmes, da de lett kan være forutinntatte. Laboratoriepanel kan benyttes ved de fleste sensoriske analysemetoder, men selvsagt ikke ved forbrukerundersøkelser.

3. Forbrukerpanel

Består av et stort antall utrente dommere som representerer en definert forbrukergruppe. Mer om forbrukeranalyser og forbrukerpanel i kapittel 5.

En sensorisk dommer er en person som utfører en sensorisk analyse.





Rekruttering

Dommere rekrutteres enten fra de ansatte i en bedrift eller organisasjon eller ved bruk av et rekrutteringsbyrå, annonser i blader og aviser, e-mail eller telefonlister eller en kombinasjon av disse.

Interne:

Dersom ansatte skal benyttes, bør de velges basert på interesse, engasjement, tilgjengelighet og evner til å fungere som sensoriske dommere.

En fordel ved å benytte ansatte er at oppgaven ofte kan være relatert til den enkeltes arbeidsaktiviteter. Pausen kan være et velkomment avbrekk i det daglige arbeidet og ha en positiv motivasjonseffekt. En annen fordel er at det er enklere å rekruttere og innkalle dommere som allerede er i hus.

En ulempe kan være at ansatte naturligvis kjenner en bedrifts produkter, så en bias effekt kan oppstå om man f. eks sammenligner bedriftens eget produkt med en konkurrent.

Eksterne:

Eksterne dommere kan bli rekruttert fra lokalsamfunnet og kan lønnes som deltidsansatt eller betales etter avtale. Eksterne dommere kan ofte være mer motiverte enn interne dommere da dette er deres spesielle jobb, ikke bare en «pausejobb» som for de interne.

Utvelgelse

Kriteriene for utvelgelse omtalt i dette avsnittet, gjelder i stor grad dommere som skal bedømme lukt, smak, farge, utseende, taktilt (følelse) eller en kombinasjon av disse, men kan uansett sees på som generell veiledning.

En bør kalle inn 2 til 3 ganger så mange dommerkandidater som det er behov for i det sensoriske panelet.

Kandidater som har sagt seg villige til å være sensoriske dommere, er pålagt å delta i en rekke screeningtester for å kartlegge deres sensoriske ferdigheter.





Utvelgelse av dommere baseres på grunnlag av:

- Interesse og motivasjon
- Sensitivitet for diverse stimuli
- Eventuelle etniske eller etiske motforestillinger knyttet til bedømmelsen av aktuelle produkter
- Generell bakgrunnskunnskap om produksjonsprosess og produkt
- Generell helsetilstand, eventuelle allergier og bruk av medikamenter som kan innvirke på bedømmelsen
- Tilgjengelighet ved deltagelse i et panel, personlighet og evne til å kommunisere muntlig og skriftlig.

Nødvendig informasjon som alder, kjønn, nasjonalitet, sensorisk erfaring, røykevaner osv. noteres.

Dommere som velges ut bør

- Være i alderen ca 20 til 55 år
- Være motivert for oppgaven
- Være tilgjengelig på aktuelle tidspunkt
- Ikke lide av allergi overfor det produktet som skal bedømmes
- Fortrinnsvis ikke røyke
- Ikke være fargeblinde dersom bedømmelse av farge skal bedømmes
- Ha en alminnelig god helsetilstand

Utvelgelsestester

Videre prosedyre for valg av dommere er å teste kandidaten for en rekke egenskaper:

- Terskelverdibestemmelse - grunnsmakstest
- Forskjellstester (differansetester) og identifiseringstester
- Bedømmelse av lukt
- Farge og -smaksblindhet
- Resultatene fra disse testene avgjør hvilke kandidater som er best





egnet til å delta i panelet. Det er alltid en fordel å ta med noen ekstra dommerkandidater i tilfelle noen faller fra eller viser seg å være mindre egnet som dommer.

Individuelle forskjeller og hva som avgjør menneskers evne til å oppfatte og gjenkjenne ulike sensoriske stimuli er omtalt i kapittel 2.

De av dommerkandidatene som ikke viser seg egnede som dommere, bør snarest underrettes om dette.

1.Terskelverdi bestemmelse - Grunnsmakstest

En grunnsmakstest kartlegger ens terskelverdier for grunnsmakene søt, sur, salt, bitter og umami.

Dersom ønskelig, kan en også bestemme terskelverdiene for andre aktuelle smaks- og lukstoffer.

Personer som viser seg å være smaksblinde, det vil si ikke kan identifisere en eller flere av grunnsmakene, eller tar feil av dem, bør utelukkes på dette trinnet av vurderingen. Det bør også testes for lukt- og smaksblindhet for gitte stoffer som kan være aktuelle i de produktene som skal bedømmes, som f.eks. androstenon (rånelukt, rånesmak), eller PROP (propyltioureasil).

En må også skille mellom forskjellige former for terskelverdier:

- Sensitivitetsterskel er den laveste konsentrasjonen av et stoff der dommeren kan oppdage en smak uten nødvendigvis å kunne identifisere den.
- Gjenkjennelsesterskel er den laveste konsentrasjonen av en sensorisk stimulans der et stoff kan gjenkjennes riktig.
- Differanseterskel er den minste forskjell i konsentrasjonen av en smak som behøves for at en slik forskjell skal kunne oppfattes. Det vil si konsentrasjonsforskjell i en og samme smak.

Grunnsmakstest er ofte en god indikator på dommerkandidatens generelle sensoriske ferdighet.





Absolutt terskel er konsentrasjonsnivået for et stoff der 50 % av svarene er riktige. Absolutt terskel har en mer generell betydning enn sensitivitetsterskelen og gjenkjennelsesterskelen, og kan benyttes til å fastsette begge de andre tersklene. I praksis er det imidlertid blitt vanlig å bruke absolutt terskel til å bestemme sensitivitetsterskelen. Denne praksis har medført at absolutt terskel og sensitivitetsterskel brukes om hverandre.

Flere ulike tester vil være nødvendig for å vurdere dommerkandidaten. Det er også gunstig å teste repeterbarhet. Det at en ikke har alle svar i en test riktig, utelukker ikke noen som dommere. At en dommer har høyere terskelverdi enn gjennomsnittet, vil heller ikke automatisk bety at han eller hun er uegnet som dommer. En dommerkandidat som har lave terskelverdier og evne til repeterbarhet på grunnsmakene, er heller ikke automatisk en god dommer på sammensatte produkter. De ulike grunnsmakene kan oppfattes noe annerledes i ferdige produkter enn de oppfattes i en ren grunnsmakstest.

ISO 3972:2011 Sensory analysis – Method of investigating sensitivity of taste omtaler flere objektive tester for testing av sensitivitet, blant annet grunnsmaker.

En pannelleder anbefales derfor å teste dommerkandidatenes terskelverdi for ulike stoffer i ett eller flere næringsmidler. I de fleste tilfeller synes terskelverdien å øke (det vil si høyere konsentrasjon for å gjenkjenne stoffet) når stoffene settes til forskjellige næringsmidler men det motsatte kan også noen ganger være tilfelle, for eksempel kan sukrose i krem gjenkjennes ved lavere konsentrasjon enn i vann.

I praksis vil en ofte foretrekke dommerkandidater som har evnen til å karakterisere forskjellige lukt- og smaksstoffer i forhold til de produktene som er aktuelle ved en sensorisk bedømmelse. Det bør i denne forbindelse presiseres at en velger en middels vanskelighetsgrad ved de enkelte testene, slik at de best egnede dommerne peker seg ut blant kandidatene. Er testene for krevende, risikerer en at ingen er brukbare som dommere. Er testene for lette, vil de kanskje ikke skille mellom kandidatene.





Grunnsmakene som benyttes ved terskelverdiestemmelse

For utprøving av grunnsmakene lager en vandige løsninger av:

- Sukrose (søt)
- Natriumklorid (salt)
- Sitronsyre (sur)
- Koffein (bitter)
- Monosodium glutamat, MSG (umami)

Disse løsningene lages i forskjellige konsentrasjoner. Konsentrasjonene i tabellen nedenfor samsvarer ikke helt med de nevnt i ISO-standarden, men er justert for en enklere tilnærming. Kjemikaliene og konsentrasjonene har en årrekke vært brukt i det sensoriske miljøet i Norge.

I tillegg til kjemikaliene, lages en prøve med vann (samme vann som brukes til å løse opp kjemikaliene). Det brukes vanligvis tre konsentrasjoner av hver prøve. Alle kjemikalier som brukes, må være absolutt rene (pro analyse eller renere).

Tabell 3.2 Grunnsmaker

Grunnsmak	Konsentrasjoner (% vekt/volum)		
Sukrose (søt)	0,2	0,4	0,6
Natriumklorid, NaCl (salt)	0,01	0,03	0,06
Sitronsyre (surt)	0,005	0,01	0,015
Koffein (bittert)	0,006	0,014	0,027
Monosodium glutamat, MSG (umami)	0,017	0,034	0,070

Testen kan utføres på to måter:

1. Kandidaten får på forhånd vite hvilken smak som skal testes
2. Kandidaten får ikke oppgitt hvilken smak som skal testes

Ved 1) bestemmer en sensitivitetsterskelen og ved 2) både sensitivitetsterskelen og gjenkjennelsesterskelen. Gjenkjennelsesterskelen er alltid lik eller høyere enn sensitivitetsterskelen. Det settes klart større krav til kandidatene dersom de på forhånd ikke får oppgitt hvilke grunnsmaker de skal teste.





Det finnes flere måter å gi oppgavene til kandidatene på. Det vanligste er å bestemme deres sensitivitetsterskel og gjenkjennesterskel (se tabell ovenfor). Det gjør en enten ved hjelp av den absolutte terskel (50 % riktige avgjørelser), eller som det laveste konsentrasjonsnivå som med sikkerhet kan fastsettes av kandidaten.

Metoden kan benyttes som en del av en prosess der en vurderer dommerkandidatens egnethet til ulike tester. Vurdering av terskelverdier ved hjelp av grunnsmakstest kan også benyttes til periodisk testing av dommere som allerede er med i et dommerpanel.

Eksempel på utførelse av en terskelverditest på grunnsmakene

Som nevnt ovenfor kan en grunnsmakstest benyttes i utvelgelse av dommere og til periodisk testing av dommere som er med i et panel.

Prosedyrer ved tilberedning og servering av sensoriske grunnsmaker:

- Testen utføres helst i to omganger og gjøres i løpet av en-to dager. Hvis dette ikke er mulig, må det lages nye smaks løsninger før en ny test kan finne sted ved en senere anledning fordi smaks løsningene endres og er ustabile over tid. Hvis begge omgangene blir gjennomført samme dag, skal det gå minst to timer mellom avsluttet første omgang til start på andre omgang. Prøvene skal serveres ved romtemperatur (20 °C) og bør derfor lages dagen i forveien.
- Skyllevannet som benyttes ved grunnsmakstesten, må være fra samme kilde som prøvene løses i. Det anbefales å bruke flaskevann kjøpt i dagligvarebutikker.
- Svarskjemaet som skal brukes ved testen, gjennomgås muntlig av panelleder for dommerne, og eventuelle uklarheter må oppklares. Det må poengteres at det ikke er tillatt å forandre svarene. Når en har smakt ferdig en prøve, skal den spyttes ut, og en skal på ny skylle munnen med vann som også spyttes ut. Dette gjøres for å unngå smaksoverføring til neste prøve og hindre at dommerne blir smakstrette og mister noe av konsentrasjonen. Det er viktig å bruke god tid på testen, anslagsvis ca 30 minutter.





Grunnløsninger		
Søt	10 % sukrose	50g sukrose/500 ml vann
Salt	10 % salt	10g NaCl/100 ml vann
Sur	10 % sitronsyre	10g sitronsyre/100 ml vann
Bitter	0,1 % koffein	1g koffein/1000 ml vann
Umami	0,2 % monosodium glutamat	2g monosodium glutamat/1000 ml vann

Fortynninger: x antall ml av grunnløsningene fortynnes opp til 1000 ml		
Sukrose	Salt	Sitronsyre
20 ml (0,2 %)	1 ml (0,01 %)	0,5 ml (0,005 %)
40 ml (0,4 %)	3 ml (0,03 %)	1,0 ml (0,010 %)
*60 ml (0,6 %)	*6 ml (0,06 %)	*1,5 ml (0,015 %)
Koffein	Monosodium glutamat	
60 ml (0,006 %)	84 ml (0,017 %)	
140 ml (0,014 %)	172 ml (0,034 %)	
*270 ml (0,027 %)	350 ml (0,070 %)	

De sterkeste løsningene, merket med * i tabellen, tjener til to formål. De er ukjente prøver i selve testen, og de skal brukes som referanseprøver på grunnsmakene søt, salt, sur, bitter og umami, det vil si de prøvene som dommerne skal smake på først for å bli kjent med de forskjellige smakene, i tillegg kommer en prøve med vann slik at det totalt er seks ukjente smaker i prøvene.

Deretter skal de smake på 16 ukjente prøver, derav er 15 fortynningene fra tabellen ovenfor, pluss vann. De ukjente prøvene kodes med tresifrede, tilfeldige tall. Alle de 16 ukjente og seks kjente prøvene serveres samtidig. Prøvene skal stå i sammen rekkefølge som på svarskjemaet, med de kjente prøvene først.

Eksempel på beregnet prøvemengde

Hver dommer skal ha 25 ml prøve av hver løsning. Dette medfører at en trenger følgende mengde av hver løsning for et panel på 8 dommere:





8 dommere * 25 ml/prøve * 2 omganger = 400 ml prøve.

Det er alltid lurt å lage litt ekstra, og i dette eksemplet er det 1000 ml til rådighet. Husk at det trengs dobbel mengde av den sterkeste konsentrasjonen av de ulike løsningene på grunn av at de blir servert både som kjente og ukjente løsninger.

2. Forskjellstester (differansetester) og identifiseringstester

Forskjellstester og identifiseringstester kan i de fleste tilfeller være nyttig å benytte i en utvelgelsesprosess av sensoriske dommere. I tillegg til disse metodene kan man teste dommerne i relevante produkter og deres sensoriske egenskaper.

Metodene blir omtalt i kapittel 4.

3. Bedømmelse av lukt

Menneskets evne til å oppfatte og gjenkjenne smak er i hovedsak enklere å kartlegge og trene enn tilsvarende for lukt. Det finnes flere metoder for å vurdere følsomheten av luktorganet, men det vanligste er å presentere en stimulus i stigende konsentrasjon. En begynner med en konsentrasjon som ligger under en bestemt verdi (terskel) og fortsetter oppover til en når en konsentrasjon som oppfattes i 75 % av tilfellene. Det vil si midt i mellom 50 %, som er en tilfeldig deteksjon, og 100 %, som er en sikker deteksjon.

Dersom analyse av lukt er viktig for en bedrift bør potensielle nye dommere i en utvelgelsesfase presenteres for relevante produkter. I en luktbedømmelse må for eksempel en dommer kunne detektere og skille mellom konsentrasjoner av vanilje, bitre mandler, roser, sjasmin, mint, sitron og lignende, samt andre aromatiske løsninger.

4. Farge- og smaksblindhet

I utvelgelsesprosessen av potensielle nye dommere, er det viktig å kartlegge blindhet for enkelte sensoriske stimuli dersom dette spiller en sentral rolle for bedriften og de produkter som skal bedømmes.





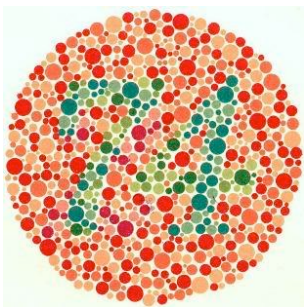
Fargeblindhet

Et klart og skarpt syn er nødvendig dersom dommerne visuelt skal bedømme ulike produkter, både næringsmidler så vel som non-food, det være seg bedømmelse av form, farge, mønster, linjer, foto eller TV-bilder.

Hvorvidt en dommer er fargeblind eller ikke, er avgjørende dersom farge på for eksempel kjøtt, vin eller bær skal bedømmes, men av mindre betydning dersom dommeren skal bedømme overflatestruktur på trepanel, parfymelukt eller smak av ulike aromatisetninger.

Fargeblindhet er nedsatt fargesyn, altså svekket fargesans, hos mennesker. Dette betyr i praksis at visse farger oppfattes annerledes. Tilstanden er som regel medfødt og kan ikke helbredes. Se kapittel 2.

Det finnes ulike tester for å kartlegge fargeblindhet, den mest anvendte er Ishihara's test. Denne testen ble introdusert tidlig på 1900-tallet og avdekker ulike grader av fargeblindhet.



Figur 3.1 Eksempel fra Ishihara's test

Visste du at:

8% av menn og 0,3% av kvinner er fargeblinde?

Eller at:

Betydelig flere kvinner enn menn er supersmakere?





Smaksblindhet

En liten del av befolkningen »lider» av spesifikk smaksblindhet, en manglende evne til å oppfatte bitre kjemikalier som PROP (6-n-propylthioureasil). Se kapittel 2.

En såkalt PROP test kartlegger hvorvidt dommerkandidater er «ikke-smakere», «smakere» eller «supersmakere», av særlig av bitre smaker.

Det er en sammenheng mellom evnen til å smake PROP og antall smaksløker på tungen. Ca 70 % av hvite mennesker har denne evnen og kalles «smakere» (45 %) og «supersmakere» (25 %). Resterende 30 prosent smaker PROP svakt eller oppfatter det som smakløst og blir kalt «ikke-smakere». Evnen til å smake PROP er arvelig, men kvinner er oftere «supersmakere».

3.3 Trening av sensoriske dommere

Dommerne er det viktigste instrumentet ved en sensorisk analyse. For å bli en dyktig og pålitelig dommer, trengs kunnskap og mye trening. Den som formidler denne kunnskapen (ofte panellederen) har et ansvar, ikke bare overfor nye dommerkandidater, men også overfor etablerte dommere. Dommerne trenger repetisjon og oppfriskning en gang i blant for å opprettholde motivasjon og konsentrasjonsevne.

Trening

Dersom det er mulig, er det vanlig å starte opplæringsperioden med flere dommerkandidater enn det som er ønskelig i det endelige panelet. Det er av stor betydning for motivasjon og forståelse for det sensoriske arbeidet at kandidatene får en grundig opplæring i faget, både teoretisk og praktisk. De bør få en innføring i sansenes oppbygning og virkemåte, samt forstå betydningen av korrekt prosedyre for testingen.

Praktisk gjennomføring

Hvordan den sensoriske analysen bør utføres vil variere ut fra hvilken metode som er valgt og hvilket prøvemateriale som skal bedømmes.





Det kan være til god hjelp å utarbeide en arbeidsbeskrivelse for dommerne hvordan de skal utføre analysen.

Dersom næringsmidler skal vurderes, vil en vanligvis starte med å vurdere prøvens utseende og farge, deretter foreta en luktanalyse og til slutt smaken og teksturen.

Smaksløkene for de ulike grunnsmakene er fordelt over hele tunga, men konkrete områder domineres av en grunnsmak (se kapittel 2).

Viktige huskereglene for dommerne:

- fordel prøven rundt i hele munnen ved hjelp av tunga.
- bruk god tid på hver prøve
- spytt ut og skylle mellom hver prøve

I en læringsperiode kan det være hensiktsmessig å servere prøver med høy og lav intensitet for å trene dommerne på ytterpunktene og for å bli fortrolig med de ulike sensoriske egenskapene.

For eksempel vil det være av interesse for en bedrift innen kjøttindustrien å få trent sine dommerkandidaters evne til å karakterisere ulike krydder, mens det på et bryggeri vil være hensiktsmessig å trene dem i ulike estere, høyere alkoholer og smørlukt (diacetyl).

Metoder

Den praktiske treningen bør konsentreres om de metoder som panelet kommer til å benytte senere. Det er viktig at dommerne leser nøye gjennom arbeidsbeskrivelsen og følger denne gjennom testen. Gjennom praktisk trening får dommerne økt innsikt i sensorikk og samtidig utvikles evnene til å finne, gjenkjenne og beskrive sensoriske stimuli. De ulike metodene er grundig beskrevet i kapittel 4.

Produktbeskrivelse og -kvalitet

Etter å ha tilegnet seg basiskunnskaper om sensorisk analyse og sensoriske metoder, følger opplæring og trening i de ulike produktene som panelet skal arbeide med. Dommerne må lære å skille gode og dårlige





produkter – eller «riktig» og «ikke-riktig» kvalitet. Til å bedømme den sensoriske kvaliteten er vanligvis poengbedømmelse den mest anvendte metoden (se kapittel 4).

Etter at treningsperioden er avsluttet, er det nyttig at dommerne fra tid til annen får repetert metoder og analyseprosedyrer. Mer om dette blir omtalt i neste delkapittel «Panellederes oppgaver».

3.4 Panellederes oppgaver

Det er avgjørende for et profesjonelt sensorisk laboratorium at det blir ledet av en faglig sterk panelleder. I tillegg til teoretisk kunnskap om sensorisk analyse og aktuelle testmetoder, er det viktig at panellederen fungerer bra som et bindeledd mellom ledelsen i bedriften på den ene siden og fagpersonale, laboratorieansatte og sensoriske dommere på den andre siden. Dette kapitlet vil omhandle nettopp dette.

En av panellederens oppgaver er å planlegge laboratoriets daglige aktiviteter og oppgaver, og vedkommende skal ikke være med i dommerpanelet.

Det er helt nødvendig at vedkommende har god kjennskap til praktisk paneldrift. Dette inkluderer prinsippene for god drift av et sensorisk laboratorium samt gjennomføring av sensorisk analyse.

Panellederen er også ansvarlig for utvelgelse og trening av dommere. At vedkommende av den grunn har en viss kunnskap innenfor fysiologi er en nødvendighet for å kunne gi dommere og eventuelt laboratoriepersonale innføring i sansenes funksjoner.

I tillegg til daglig drift, deles panellederens oppgaver inn på følgende måte:

- Kontaktperson inn mot ledelsen
- Ansvarlig for de ansatte på sensorisk avdeling
- Oppfølging av dommerne og dommerpanelet





Kontaktperson inn mot ledelsen

Panellederen bør ha god kontakt inn mot ledelsen i bedriften ettersom vedkommende ofte har administrative, vitenskapelige og tekniske oppgaver. I visse sammenhenger vil det være en fordel å ha tekniske kunnskaper om de produkter som skal bedømmes.

I bedrifter der produktutvikling og eventuelt forbrukertester gjennomføres er det veldig viktig at panellederen har et så nært samarbeid med markedsavdelingen som mulig. Mange panelledere kan oppleve det som et problem at markedsbeslutninger blir tatt uten at det sensoriske laboratoriet har i stor nok grad vært involvert. Forbrukertester blir omtalt i kapittel 5.

Ansvarlig for de ansatte på sensorisk avdeling

I tillegg til panellederen, består et sensorisk laboratorium ofte av en eller flere fagpersoner og laboratoriarbeidere. Panellederen fastsetter minimumskrav for medarbeidere i de ulike stillingene og sørger for nødvendig opplæring og trening.

Fagpersonalet bør ha relevant faglig utdanning og produktkunnskap innenfor det aktuelle arbeidsområdet. Videre er kunnskap innen statistikk en stor fordel, samt evnen til å planlegge, gjennomføre og vurdere resultater fra sensoriske analyser.

Fagpersonalet vil i mange tilfeller være de som til daglig er ansvarlig for dommerpanelet, som innebefatter trening og testing, utvelgelse av testprosedyrer og overoppsyn med at testprosedyrene blir fulgt. Videre bør fagpersonalet også kunne være ansvarlig for bearbeiding, vurdering og rapportering av analyseresultater.

På laboratoriet kan det være personer ansatt som kun har sine daglige oppgaver knyttet til sensorisk arbeid, eller det kan være personer som rekrutteres internt i bedriften ved behov. Det stilles ikke nødvendigvis krav til formell utdanning, men de bør kjenne til de viktigste metodene innenfor sensorisk analyse og hvordan disse utføres.

At medarbeiderne er motiverte og føler ansvar for det arbeidet de utfører er av stor betydning. Ansvarsområdene kan for eksempel være





å holde laboratoriet i orden, klargjøre utstyr og prøver til servering, servere prøver og registrere resultater.

Oppfølging av dommerne og dommerpanelet

Selv om det i mange tilfeller er fagpersonen(e) som til daglig er ansvarlig for trening, oppfølging og kontroll av dommerpanelet, er det panellederen som bærer hovedansvaret og må legge til rette for dette.

Periodisk trening

Dersom tilegnede ferdigheter ikke regelmessig brukes, glemmes de. Det er derfor viktig med periodisk trening. Dette kan omfatte terskelverdibestemmelser på aktuelle produkter eller grunnsmaker, bruk av nomenklatur og skalabruk.

Kontroll av enkeltdommere og panel

Det er også av betydning at panellederen til enhver tid har oversikt over dommernes prestasjoner, slik at eventuelle avvik i bedømmelsen raskt kan justeres. I praksis innebærer dette at det bør lages egne kontrollprogram for den enkelte dommer. Før et kontrollprogram opprettes, bør det vurderes nøye hvilke opplysninger som er relevante å legge inn. Lagres sensitiv informasjon om enkeltpersoner, kan det være nødvendig med konsesjon fra offentlige myndigheter for å opprette og drive et slikt register.

Flere av software-løsningene for sensorisk analyse som er på markedet i dag har inkludert en funksjon hvor man enklere kan følge opp dommerprestasjoner. PanelCheck (www.panelcheck.com) er vel og merke en gratis software som er et spesielt utviklet verktøy som gjør det enklere å overvåke enkeltdommeres og panelets prestasjoner (se kapittel 11).

Med jevne mellomrom bør panelet kontrolleres i sine bedømmelser. Kontrollen kan enten være åpen, det vil si at dommerne informeres at de skal kontrolleres, eller den kan være lukket. Kontrollen kan gå ut på at dommerne får to like prøver til bedømmelse, eller at en «ikke riktig» prøve legges inn i en prøveserie. Dommerne skal etter hver kontroll ha en tilbakemelding, og eventuelle uklarheter og misforståelser skal kunne oppklares.





Deltagelse i sensoriske interkalibreringer, såkalte sammenlignende laboratorieprøvinger (SLP), eller ringtester kan være en nyttig måte å trene dommerne på. Dette er også en god mulighet til å vurdere nye dommerkandidater.

Disse testene vil også kunne gi personalet og medarbeiderne en tilbakemelding på hvordan de arbeider, slik at de lettere kan kvalitetssikre sitt eget arbeid. Kvalitetssikring av sensoriske analyser blir ytterligere omtalt i kapittel 10.

Dommerne må alt fra starten av informeres om at de kan risikere å bli utelukket fra panelet dersom de ikke oppfyller ønskede krav. Dersom panellederen ser at avvikene over tid er for store, bør man i første omgang ta en samtale med den dommeren det gjelder for å avklare om det er noen spesielle forhold som kan forklare avviket. Er dommeren interessert i å forbedre sine prestasjoner? Aksepterer dommeren å gå igjennom en opptreningsperiode? En god toveis kommunikasjon mellom panel og leder er nødvendig for å oppnå gode resultater.

Motivasjon

Viktige faktorer for å opprettholde dommernes motivasjon og konsentrasjonsevne er:

- Oppfriskning av metoder og analyseprosedyrer
- Tilbakemelding på testresultater
- Anerkjennelse fra ledelsen om hvilken viktig jobb de utfører
- Deltagelse i diskusjonsgrupper for å fastsette sensorisk profil eller andre kvalitetskriterier
- Kursing, bedriftsbesøk, temadager med deltagelse av andre sensoriske panel

3.5 PRØVEPREPARERING OG SERVERINGSPROSEDYRER

En sensorisk analyse gjennomføres ofte på en prøve som skal representere et større vareparti, en større produksjon eller en periodes produksjon. Antall prøver som tas ut i en slik sammenheng, eventuelt når de





tas ut og hvilken størrelse eller mengde som tas ut, bør derfor vurderes nøye. Det samme gjelder betingelsene under eventuell oppbevaringstid og -temperatur fram til analysetidspunkt.

De generelle sensoriske metodene har i liten grad konkrete anbefalinger vedrørende disse parameterne. Enkelte parametere kan være dekket i bransjestandarder for utførelse av sensorisk analyse, men det henvises oftest til mer generelle standarder for statistiske stikkprøveplaner og prøveuttak. En slik standard er NMKL Prosedyre Nr. 12, 2013: Håndbok i prøvetakning av næringsmidler.

Når man planlegger en sensorisk test, er det mange viktige detaljer å ta i betraktning.

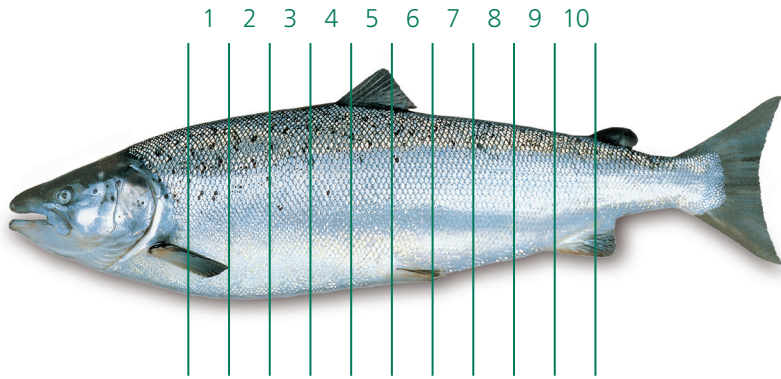
- prøvepreparering
- holdetid
- størrelse og form på prøven
- serveringstemperatur
- serveringsbeholder
- antall produkter i hver sesjon
- skylle-/renseprosedyre til dommerne

Prøvepreparering og holdetid

Så langt det er mulig, skal prøver som skal bedømmes i en og samme test prepareres med identisk utstyr (samme merke og modell) som har blitt kalibrert for prøve preparering. Serveringstemperaturen skal kontrolleres nøye og være lik for alle prøvene som skal bedømmes i samme test.

Temperaturen for enkelte produkter kan være vanskeligere å kontrollere enn for andre, for eksempel kjøtt. Andre produkter, som vin eller meieriprodukter har gjerne en høyere serveringstemperatur enn anbefalt lagringstemperatur for å forsterke oppfatningen av lukt og smak. Serveringstemperatur skal vel å merke være trygg for konsum og stemme overens med konsumentenes forventninger.





Figur 3.2 Fisk med markering for de enkelte dommeres stykker

Foto: Nofima



Figur 3.3 Ulike serveringsbeholdere

Foto: Kjell Merok





Prøvestørrelse

Prøvene innenfor en test skal være så identiske i størrelse og form som mulig når de serveres dommerne. Størrelse og form kan påvirke både utseende og tekstur. Dersom et sensorisk panel eksempelvis skal analysere en fisk, kan dette by på utfordringer ettersom et ryggstykke er betydelig tykkere enn et halestykke. For rettmessig å kunne sammenligne flere fisk (som for eksempel er lagret ved ulik temperatur), er det viktig at hver dommer får bedømme samme anatomiske del av alle fiskene for at sammenligningen mellom prøvene skal bli riktig. For eksempel dommer 1 bedømmer skive 1 på alle fiskene og dommer 2 bedømmer skive 2 på alle fiskene, osv. (se figur 3.2). Størrelsen bør videre stemme overens med hensikten med forsøket og være tilstrekkelig for å fullføre bedømmelsen.

Serveringsbeholder

Serveringsbeholderne (glass, fat, skåler etc.) må være forhåndsgodkjent av sensorisk avdeling for ikke å avgi uønsket smak og/eller lukt til produkter som skal bedømmes. For noen laboratorier er det mest gunstig å bruke engangsbeholdere mens for andre er det best å bruke vaskbare beholdere. Dersom engangsbeholdere brukes, må de innkjøpte partier regelmessig kontrolleres ettersom råmaterialene som brukes raskt endres.

Utstyr og redskap skal være rent og ikke forurensende. Eksempelvis kan såperester gi lukt og smak og plast og papir kan avgi lukt. Serveringsbeholderne må testes og tåle å bli brukt ved den temperaturen som prøvene normalt skal serveres ved. For eksempel, en plastbeholder som ikke avgir lukt og smak ved romtemperatur, kan godt gjøre det dersom den blir varmebehandlet eller at det blir servert f. eks. varm suppe i den.

3.6 Prøvepresentasjon

Rensing av munnhulen

Rensing av munnhulen mellom prøvene som skal testes er viktig for å fjerne prøverester. Vann og usaltet kjeks er mest brukt til det formålet





i sensoriske tester så vel som forbrukertester. Ved bedømmelse av fettholdige næringsmidler, anbefales bruk av varmt vann. Mange ulike produkter som te, melk, gulrot og brødbiter har vært forsøkt brukt som rensmiddel ved bedømmelse av sterkt krydret mat. Ulempen med dette er selvsagt at man tilfører nye smaker og teksturer som forstyrrer de sensoriske egenskapene til produktene som skal testes. Derfor anbefaler gjerne profesjonelle sensorikere at munnhulen renses med usaltet kjeks og vann.

Spytting og svelging

Spytting versus svelging av en prøve er et ofte debattert tema. Mange profesjonelle sensorikere anbefaler at prøven spyttes ut fordi det reduserer en mulig overføring fra en prøve til neste, en såkalt «carryover effect».

Svelging gir også en metthetsfølelse som kan påvirke evnen til å bedømme mange prøver. Det som på den annen side taler for å svelge prøver er at det også sitter smaksløker bak i munnen og i svelget som kan være viktigere for bedømmelse av noen produkter enn andre. Dette inkluderer hvordan svelget oppfatter inntak av varm mat samt ettersmaken av aromatiske produkter som øl, vin og sjokolade.

Man må være klar over at jo lenger unna en «normal» konsument situasjon man velger å teste produktene i, dess større er muligheten for at man mister går glipp av sensoriske egenskaper som er avgjørende for konsumentene.

Dommerne kan også være ukomfortable med å spytte, hvilket igjen kan påvirke deres bedømmelse. Der det er et tema, bør det være tilrettelagt for at prøver både kan svelges og spyttes. Det er viktig at dommerne er konsekvente i sitt valg ettersom både skylling og svelging kan endre sensitiviteten.

Orientering og instruks til dommerne

I forkant av en analyse bør dommerne få en instruksjon og muligheten til å stille spørsmål. Instruksjonen til dommerne skal være klar og presis. I noen tilfeller er det nødvendig at informasjon gis muntlig ettersom skrevne instruksjoner kan misforstås. En god regel kan være å gi en





muntlig instruks i forkant av testen samt å tilby en skriftlig instruks de kan benytte under analysen.

Det er vel og merke viktig at det ikke gis noen informasjon rundt om selve produktet som skal testes.

Det er også viktig at dommerne er kjent med mer generelle instruksjoner som har betydning for hver enkelt dommer, samt meddommere. Noen av instruksene nedenfor er generelle, mens noen er spesielt beregnet på dommere som skal bedømme lukt og smak.

- Møt til rett tid
- Unngå å forstyrre andre dommere med prat og telefon
- Unngå sterk parfyme og aftershave
- Unngå å bruke sterk parfymert såpe og håndkrem
- Unnlate å spise eller drikke annet enn vann inntil 30 min før bedømmelsen

Koding av prøver

Alle prøver skal være kodet med en unik, tresifret kode. Grunnen til at man unngår å bruke koder med bokstaver eller ett eller to siffer, er at disse lett kan gi assosiasjoner og forventninger som ikke tresifrede koder gir. På den måten unngår man i størst mulig grad at dommerne velger prøver hvor koden har en bevisst eller ubevisst betydning.

Det anbefales at kodene som brukes i enhver test er tilfeldige valgt uten noe form for system.

Utforming av svarskjema

Enten den sensoriske analysen besvares med penn og papir, på nettbrett eller på PC med spesialtilpasset software, er det mange viktige hensyn å ta ved utformingen av svarskjema.

I tillegg til en informativ instruks er det viktig at metoden som brukes er kjent for dommerne. Dette gjelder både differansetester, så vel som beskrivende analyse med ulike skalaer osv.





Hvordan et svarark korrekt bør utformes vil variere med metode og hva som er målet med analysen, hvilket ikke gjør det hensiktsmessig å generalisere det i dette avsnittet.

Det er allikevel viktig å understreke viktigheten med å kvalitetssikre oppbyggingen av et svarark for å unngå unødvendige mistolkinger eller misforståelser blant dommerne som kan resultere i at oppgaven blir ulikt løst innad i panelet.

Tid og lengde for analysen

Ved gjennomføring av sensorisk analyse, er det viktig at det avsettes tilstrekkelig tid. Dette omfatter både forberedelser i form av trening, og at panellederen sørger for at dommerne bruker tilstrekkelig tid på selve bedømmelsen. En generell regel er at det er bedre å teste få prøver ordentlig enn mange prøver dårlig.

Dommerne må ikke være for mette eller for sultne ved bedømmelsen. Tidspunktet for bedømmelsen bør derfor være mellom to måltider, for eksempel mellom frokost og lunsj. Har dommerne andre arbeidsoppgaver, er det viktig å ta hensyn til dette når en tilpasser tidspunktet for bedømmelsen. Benyttes et internt panel, bør en unngå at det blir konflikt mellom det daglige arbeidet og gjennomføringen av de sensoriske analysene.

3.7 Testområdet og bedømmelseslokalet

Testområdet bør være sentralt plassert slik at det er mest mulig praktisk for dommerne som skal delta. Det bør ikke være i nærheten av områder med mye lukt og/eller lyd.

Avhengig av bruk, kan det være hensiktsmessig med et diskusjonsrom og et venterom i umiddelbar nærhet av selve bedømmelseslokalet.

Et prepareringsrom bør ha direkte tilgang til bedømmelseslokalet. Innredning av dette er avhengig av bruk og hvilke prøver som skal analyseres.





Dommerne må ha tilgang til bedømmelseslokalet uten å gå igjennom prepareringsrommet for å unngå fysisk eller visuell tilgang til informasjon som muligens vil kunne påvirke bedømmelsen deres.

ISO 8589:2007 behandler anbefalinger til selve bedømmelseslokalet. Denne standarden er relativ generell og innholdet bør tilpasses det aktuelle laboratoriet og de metodene som brukes der.

Et laboratorium som skal bedømme kjeks eller suppe vil nødvendigvis se annerledes ut enn et laboratorium som skal bedømme parfyme, såpepulver eller lyd. Uansett, er det en del generelle prinsipper som er verdt å vurdere. Lokalet bør være innredet slik at det er lett å rengjøre og byggematerialene som benyttes må ikke avgi lukt.

Forholdene må være slik at dommerne ikke forstyrres under selve bedømmelsen. Årsaker til forstyrrelser kan være forskjellige typer lyd, støy fra samtaler, telefon eller at uvedkommende kommer inn i lokalet.

Farge og lys

Fargen på en prøve vil oppfattes forskjellig avhengig av lyskilden, og er det dagslys i rommet, vil lyset variere avhengig om det er solskinn eller overskyet. Er det vinduer i lokalet, bør disse blendes.

Fargen på vegger, skillevegger og bordplater bør være nøytral og matt, offwhite eller lys grå er anbefalt.

Lokalet bør være belyst av en lyskilde med en fargetemperatur på 6 500 K («dagslys») og en best mulig fargegjengivelse, hvis det ikke er helt spesifikke krav til belysning.

Belysningsstyrken på bordplaten bør være konstant. Den anbefalte lysstyrken er ca 1000 lux. Belysningsstyrken er den totale lysytelsen (hvor mye lys totalt stråler ut fra lyskilden) som treffer en flate delt på flatens størrelse angitt i m².

Det bør videre være mulig å bruke farget lys (rødt eller grønt) slik at prøvene kan maskeres i de tilfellene der farge skal utelukkes fra bedømmelsen.



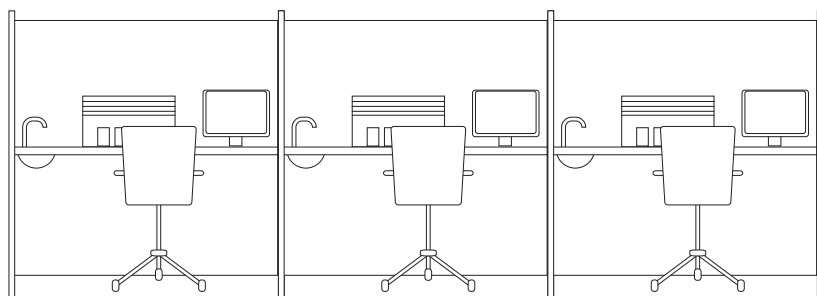


Temperatur, luftfuktighet og luftkvalitet

Temperaturen i lokalet skal være tilpasset hvilke produkter som skal bedømmes og dommerens komfort (e.lign.). Generelt er temperaturen mellom 20 og 22 °C behagelig. Luftfuktigheten bør også vurderes. I prinsippet bør den være behagelig for dommerne, stabil og kontrollerbar. Dersom prøvene krever testing i en spesiell temperatur må det selvsagt tas hensyn til dette. Luften i rommet bør være HEPA filtrert. Et HEPA-filter (High Efficiency Particulate Air) har som funksjon å fange opp fine støvpartikler. Det bør også være et svakt overtrykk i rommet for å forhindre uønsket lukt å trenge inn.

Innredning

Hvilken innredning man skal velge er avhenging av flere faktorer, blant annet om det er et nybygg eller et eksisterende rom som må tilpasses i forhold til størrelse og ikke minst økonomi. Det er ikke nødvendig med dyre innredninger og utstyr for å etablere rutiner som kan fungere. Dommerne bør vel å merke normalt være adskilt med skillevegger for å kunne analysere uavhengig av hverandre. Lokalet bør generelt være utstyr slik at dommerne føler seg mest mulig komfortable og kan gjøre en best mulig jobb.



Figur 3.4 Eksempel på innredning av bedømmelseslokale

Illustrasjon: Julie Gjermundsen





Når man planlegger innredningen i et bedømmelseslokale, er det i første omgang følgende hensyn som må tas:

- hvor mange dommere som skal delta
- frekvens på bedømmelser
- hva slags prøver som skal bedømmes
- antall prøver som skal bedømmes samtidig
- hvorvidt det er nødvendig med strømuttak og vask integrert i hver dommerbås

3.8 Faktorer av betydning for sensorisk analyse

Ved gjennomføring av en sensorisk analyse, er det en rekke fysiologiske og psykologiske faktorer som påvirker resultatet. Noen av faktorene, som skylling av munnen, tidsfaktoren, tidspunkt for bedømmelsen, egenskaper ved prøvene, terskelverdier, helsetilstand og motivasjon er allerede omtalt i dette kapittelet. Men det er også andre viktige faktorer som vi må vektlegges.

Som ved all annen eksperimentell virksomhet, gjelder det først å undersøke hvilke faktorer som kan påvirke resultatene, for deretter å ta hensyn til disse ved planleggingen av forsøkene. Når det gjelder «å ta hensyn til», innebærer dette at en på forhånd bestemmer seg for enten å standardisere variabelen, det vil si holde den konstant under hele forsøket, eller å gjøre den tilfeldig, det vil si å sørge for at effekten blir lik for alle dommerne og alle prøvene.

Fysiologiske faktorer

Adapsjon

Adapsjon vil si at sensitiviteten reduseres for en bestemt stimulus som et resultat av gjentatt eksponering for en bestemt impuls(er) som er tilnærmet like. I forbindelse med sensoriske tester er dette en viktig og uønsket kilde til variasjoner i resultatene. Er skyllevannet som benyttes surt, vil dommernes evne til å bedømme hvor surt et produkt er bli redusert. Motsatt effekt vil oppstå hvis skyllevannet er tilsatt sukker,





og dommerne skal bedømme intensiteten av salt. Dommerne vil da bli mer følsom for salt enn om det blir brukt nøytralt skyllevann.

Forsterking eller kamouflering

Stimuli kan bli forsterket eller kamuffert avhengig av om de blir presentert hver for seg eller i en blanding. Dommerens evne til å bedømme lukt vil reduseres hvis prøvene avgir lukt i bedømmelseslokalet. Derfor er det viktig med god ventilasjon i rommet.

Synergieffekt er når intensiteten av flere stoffer i en blanding er større enn summen av intensiteten av stoffene hver for seg.

Psykologiske faktorer

Forventningseffekt

Denne effekten oppstår når dommeren assosierer to eller flere egenskaper i en prøve til en bestemt egenskap og får en forutinntatt holdning til prøven. Et eksempel er at vi antar at en lys rød bringebærsaft er mindre søt enn en mørkere rød saft.

Antall tester og prøver

Hvor mange tester eller prøver en dommer klarer å vurdere i en omgang kan avgjøres empirisk. Det avhenger blant annet av faktorer som type produkt, frekvensen på bedømmelsene over lengre tid, samt dommernes motivasjon og trening. Erfaringene har vist at en kan klare flere bedømmelser av prøver med mild smak enn sterk smak, og at en kan klare flere prøver om antall spørsmål om hver prøve reduseres (partiell profilering). Det er ikke mulig å gi absolutte anbefalinger for hvor mange tester eller prøver som et panel kan klare i en omgang, men det er åpenbart at en erfaren dommer kan klare flere prøver enn en dommer helt uten trening.

Individuelle forskjeller hos dommerne

Enkeltdommeres personlighet og samspillet mellom dommerne kan være av stor betydning. Enkelte mennesker er svært forsiktede, mens andre er mer dominerende. Det er viktig at dominerende personer ikke påvirker de andre dommerne under eller etter bedømmelsen.





Enkeltdommeres ulike terskelverdier, helsetilstand, alder og motivasjon er som tidligere nevnt selvsagt også avgjørende. Dette er faktorer som til en hver tid må vurderes og tas hensyn til.

Forhold ved prøvene

Forhold som prøvemengde, prøvenes temperatur og prøvepresentasjon er omtalt i delkapittel 3.5 og 3.6.

I tillegg er det flere andre viktige forhold som må belyses.

Kontrast- og konvergensfeil

Dette er trolig den vanskeligste type feilkilde å eliminere da disse feilkildene er knyttet til produktet og oppstår ofte samtidig. Kontrastfeil oppstår når to prøver som bedømmes er veldig forskjellige samtidig som forskjellen mellom prøvene blir mye større enn forventet. Dette kan oppstå når det serveres en prøve av antatt riktig kvalitet før en prøve av gal/feil kvalitet. Kontrasteffekten oppstår ved at den dårlige prøven får en dårligere vurdering enn om rekkefølgen hadde vært motsatt. Forskjellen blir overdrevet og den dårlige prøven blir oppfattet som bedre enn om den hadde blitt servert sammen med en prøve som er mer lik i kvalitet.

Konvergens (sammenfallende) er den motsatte effekten. Kontrast mellom to eller flere prøver overskygger mindre forskjeller mellom en av prøvene og de øvrige prøvene i en test. I tester som omfatter tre eller flere prøver vil som regel begge effektene gjøre seg gjeldene. Disse to effektene er vist i figur 3.5. Det øverste eksemplet viser innvirkningen på skalering og det nederste eksemplet viser en preferansetest. Det er klart at når forskjellene blir unaturlig stor mellom noen prøver, kan dette medføre at små forskjeller på andre prøver i prøveserien blir oversett.

Forventning til prøvene

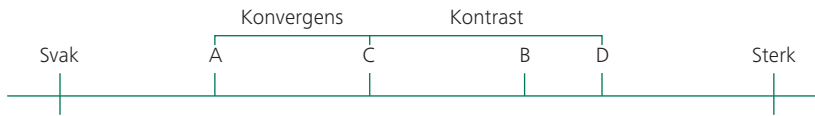
Dette er en effekt som er til stede når dommeren har en bestemt forventning til prøven. Det er derfor viktig at dommerne kun har den informasjonen om prøvene som er nødvendig for å gjennomføre bedømmelsen. Dette er en av grunnene til at panellederen ikke skal





Skalering

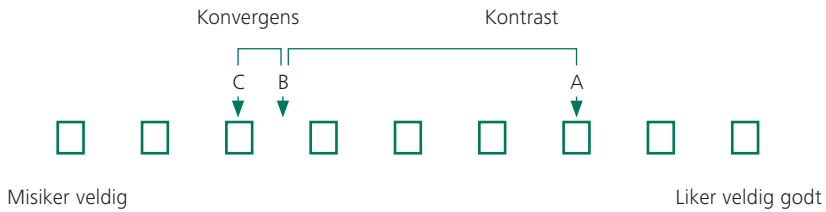
Intensitet av salt smak



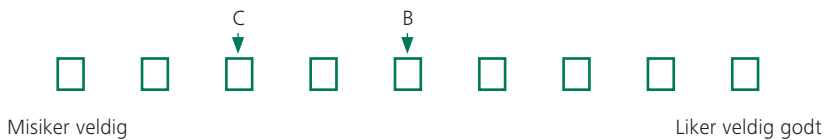
Resultat uten prøve A



Preferansetest



Resultat uten prøve A



Figur 3.5 Eksempel på måling av intensitet (skalering) og preferansmåling som viser kontrasteffekt og konvergenseffekt





delta i panelet. Eksempelvis ved bedømmelse av jordbærsyltetøy vil de fleste oppfatte et mørkt rødt syltetøy som tydeligere på smak enn det syltetøyet som er blitt brunfarget. Derfor vil det være hensiktsmessig å la en slik bedømmelse bli gjennomført under rød belysning som kamuflerer den riktige fargen på prøven. Se kapittel 2.

Serveringsrekkefølgen

Det er flere forhold ved serveringsrekkefølgen som påvirker resultatene. Feil som oppstår som en følge av serveringsrekkefølgen kan reduseres ved at forsøket er balansert og randomisert (se delkapittel 3.1, Forsøksdesign). Et annet tiltak som kan redusere effekten, kan være å la dommerne »varme opp» før selve bedømmelsen. Med oppvarming menes at dommerne starter en omgang med å bedømme en eller flere prøver som er ganske lik(e) de reelle prøvene, før selve bedømmelsen starter.

Haloeffekten eller glorieeffekten

Når flere egenskaper ved en prøve skal bedømmes kan resultatet bli annerledes enn om hver egenskap blir bedømt hver for seg. Dersom en prøve har et lite innbydende utseende, kan dette gi en forventningseffekt som kan påvirke vurderingen av andre egenskaper ved prøven. Dette kan ha stor betydning for eksempel i en forbrukerundersøkelse der man gjerne spør om forbrukerens generelle oppfatning av produktet. Spør man deretter om spesielle egenskaper i produktet, vil resultatet bli påvirket av om personen liker produktet eller ikke. Dette kalles haloeffekten eller glorieeffekten. Er det egenskaper ved prøven som er spesielt viktige, kan denne effekten reduseres ved å servere prøven i flere omganger og vurdere én egenskap hver gang.

Systemeffekten

Serveres prøvene i en bestemt rekkefølge i forhold til økende eller avtagende intensitet vil dommerne fort oppdage dette. Dette vil naturlig nok påvirke bedømmelsen, for eksempel søt-søtere-søtest.

Sentrumseffekten

Det er en psykologisk effekt at man helst unngår ytterpunkter og derfor velger prøver i midten av prøveserien. Prøver som blir servert





midtveis i en prøveserie, har også en tendens til å bli vurdert bedre enn om de serveres mot slutten. Og man ser at i triangeltester vil prøven som er forskjellig, oftere oppdages når den serveres i midten. Denne effekten kan elimineres ved å randomisere serveringsrekkefølgen.

Rekkefølgeeffekten

Dommernes holdning til prøvene forandrer seg gjennom en prøveserie. Når de starter på en serie har de visse forventninger som kan gi utslag i at den første prøven blir vurdert som veldig bra (eller dårlig). Mot slutten av serien blir dommeren ofte sliten og kan få en negativ eller likeglad holdning til prøvene, noe som også kan påvirke bedømmelsen. Dommerne kan også venne seg til en dårlig smak og bedømmer de siste prøvene bedre enn de første.

Referanser

NS-EN ISO 8586:2014 Sensorisk analyse - Generelle retningslinjer for valg, opplæring og overvåking av utvalgte dommere og ekspertdommere (ISO 8586:2012)

ISO 3972:2011 Sensory analysis – Method of investigating sensitivity of taste

ISO 5496:2006 – Sensory analysis - Initiation and training of assessors in the detection and recognition of odours

ASTM STP758 – Guidelines for the selection and training of sensory panel members

ISO 11132:2012 – Sensory analysis – Guidelines for monitoring the performance of a quantitative sensory panel

ISO 13300-1:2006 – Sensory analysis – General guidance for the staff of a sensory evaluation laboratory — Part 1: Staff responsibilities

ISO 13300-2:2006 – Sensory analysis – General guidance for the staff of a sensory evaluation laboratory – Part 2: Recruitment and training of panel leaders

ISO 8589:2010 - Sensory analysis – General guidance for the design of test rooms

NMKL Prosedyre Nr. 12, 2013: Håndbok i prøvetakning av næringsmidler





METODER FOR OBJEKTIV BEDØMMELSE



Foto: Mari Svenningsen





METODER FOR OBJEKTIV BEDØMMELSE

Marit Rødbotten og Mats Carlehög, Nofima

Det er etter hvert blitt utviklet mange metoder som er mer eller mindre standardiserte og som gir svar på bestemte problemstillinger vedrørende produkters sensoriske egenskaper. Som for kjemiske og mikrobiologiske analyser brukes det forskjellige metoder også for sensorisk analyse, alt etter hva som skal undersøkes. En kjemiker velger bestemte kjemiske analysemetoder for å undersøke surhetsgrad i et produkt, og en sensoriker velger en bestemt sensorisk metode for å analysere om det er en sensorisk målbar forskjell i surhet mellom to produkter.

I dette kapitlet skal vi gjennomgå de metodene som oftest blir brukt for analyse av objektive egenskaper ved et produkt.

Det finnes forskjellige måter å klassifisere sensoriske metoder på, som for eksempel etter bruksområde, paneltyper, metodetype. Vi kan her klassifisere dem i følgende tre hovedgrupper:

- forskjellstester
- kvantitative tester
- forbrukertester

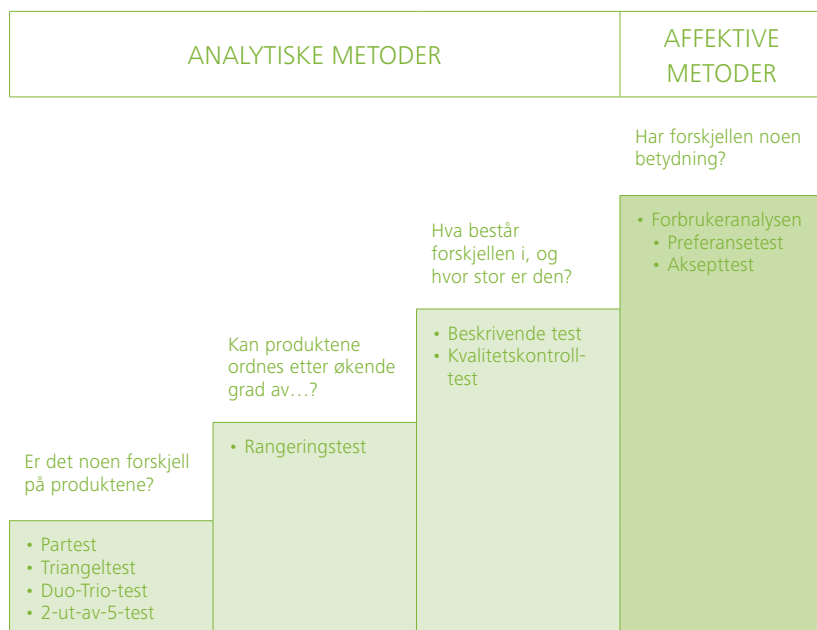
Forskjellstester og kvantitative metoder kalles også laboratoriemetoder eller analytiske metoder, og er objektive metoder som utføres av et trent panel (se kapittel 3.2.1). Metoder for forbrukeranalyser kalles affektive (= følelsesmessig) eller hedoniske (= lykkefølelse) metoder og er subjektive analyser utført av forbrukere som skal gi informasjon om hvilken prøve de liker best eller på annen måte gi uttrykk for sine subjektive meninger (se kapittel 5).





Vi finner ofte at ordene «test» og «metode» brukes om hverandre. «Metode» blir brukt om å undersøke eller teste noe. Det betyr at når vi for eksempel snakker om en forskjellstest burde vi riktigere snakke om en forskjellsmetode eller kanskje en forskjellstestmetode. På engelsk heter det en «sensory difference test», og det er antakelig forklaringen på at den norske oversettelsen er blitt «sensorisk forskjellstest». Vi velger i denne boka å fortsette bruken av ordene test og metode slik det er blitt en vane innen vitenskapen sensorikk, vel vitende om den språklige unøyaktigheten.

Det er hva en ønsker å finne ut om produkters sensoriske kvaliteter som avgjør hvilken metode en skal velge. Metoden avgjør hva slags panel en skal bruke for å få riktig analyseresultat. Analytiske metoder krever som oftest et trent dommerpanel mens utrente dommere (forbrukere) gjennomfører analyser med affektive metoder, se figur 4.1.



Figur 4.1 Sensoriske problemstillinger løst med bestemte metoder





4.1. Skala

Når en eller flere av våre sanser får et inntrykk eller en stimulanse (for eksempel vi spiser et eple) vil det bli oppdaget som en følelse eller sansing (vi kjenner sprøhet, søt smak). Personen som spiser eplet må bruke en metode for å gi uttrykk for sin oppfatning av denne følelsen, en respons på følelsen, og ofte angi grad eller intensitet av den stimulus som er opprinnelsen til følelsen.

Stimulans — følelse — respons

Det er først når observasjoner kan relateres til en skala at vi kan snakke om måling. Data kan forekomme på fire forskjellige skalaer: nominal-, ordinal-, intervall- eller ratioskala.

- Den enkleste måten å anvende tall på er til gruppering av informasjon ved opptelling av antall tilfeller i bestemte kategorier. Kategoriene kan være for eksempel ja/nei, mann/kvinne eller eple/pære/banan. Selv om dette ikke representerer noen skala og derfor ikke er noen måling slik vi vanligvis forstår ordet, snakker vi likevel om måling med nominalskala (nomen=navn). Eksempel på sensorisk analyse på nominalnivå er forskjellstester.
- Måling forutsetter at objektene lar seg ordne etter en rekkefølge fra «minst» til «mest» med hensyn på en definert egenskap. Hvis vi antar at det dreier seg om en naturlig ordning mellom kategoriene, kalles det en ordinalskala. Slik måling kan for eksempel være en konsistensmåling av majones som «fast», «fastere», «fastest». Det dreier seg da om en rangordning eller en rangering. Dersom vi bruker tall for å betegne de forskjellige kategoriene (fast=1, fastere=2, fastest=3), er dette bare en betegnelse for rangordning. Tallene sier ikke noe om innbyrdes avstand mellom kategoriene, og beregning av gjennomsnittsverdier er uten mening.

Det er først når observasjoner kan relateres til en skala at vi kan snakke om måling.





- Målinger som er utført slik at tallverdiene formidler informasjon om avstanden mellom punktene på en skala med hensyn på en bestemt egenskap, kalles måling på intervallnivå eller måling med intervallskala. I tillegg til rangordning kan det her også defineres meningsfylte differanser. Differansen mellom 3 og 6 er den samme som mellom 5 og 8. Ved bruk av intervallskala er det meningsfylt å beregne gjennomsnittsverdier av målingene. Tester som beskrivende test og kvalitetskontrolltest forutsetter data på intervallnivå.
- Ved bruk av fysiske målemetoder får vi data på ratioskala (meter, liter, gram). Her har vi måleenheter med absolutt nullpunkt, og først på dette nivået kan vi snakke om forhold mellom størrelser, for eksempel at 20 gram er dobbelt så mye som 10 gram.

Innen forbrukeranalyser brukes måling av preferanse og aksept. Ved måling av preferanse angis valg mellom to eller flere prøver, men her får vi ikke svar på om prøvene likes eller mislikes, bare rekkefølgen i liking av prøvene. En akseptmåling viser om produktet blir oppfattet som godt eller ikke, og i hvor stor grad prøven likes eller mislikes. Forbrukere kan altså foretrekke (preferere) produkt A fremfor produkt B, men mislike begge produktene (lav aksept av begge to).

En forbruker kan også bli presentert forskjellige påstander relatert til produkter og bli bedt om å angi grad av enighet i påstandene. En slik skala blir kalt en Likert-skala og den kan deles inn som en tallskala (for eksempel 5-, 7-, 9-punkt) med endepunkter enig og ikke enig.

Kroppsspråk og ansiktsuttrykk avslører ofte om vi liker eller misliker produkter og påstander. Dette ligger til grunn for en skala som ofte brukes når barn er forbrukere i sensoriske tester. En skala som viser blide og sure ansikter er hensiktsmessig for barn som ikke kan lese verbale skjema (figur 4.2).

Poengbedømmelser

Til å bedømme den sensoriske kvaliteten eller deler av den, er poengbedømmelse (ofte kalt numerisk gradering eller scoring) mest brukt.





Figur 4.2 Skala med ansikter hensiktsmessig ved analyser hvor barn er forbrukere

Egenskapene ved varen som skal bedømmes blir karakterisert ved hjelp av en tallskala (poengskala). Dette er måling på intervallnivå.

Ved kvalitetskontrolltest er det vanlig at den ene enden av skalaen angir riktig kvalitet og den andre enden angir avvikende kvalitet. Det er med andre ord et karaktersystem. Metoden er illustrerende og anvendelig, men den krever at dommerne må ha en klart definert spesifika-sjon å arbeide etter.

Ved beskrivende test blir intensiteten av enkelte egenskaper målt ved at den ene enden (oftest venstre siden) av skalaen betegner «ingen» eller «lite» av den angitte egenskapen, mens den andre enden (høyre siden) av skalaen betegner «mye» eller «tydelig» intensitet av den angitte egenskapen.

De ord som velges i bedømmelse av et produkt må være beskrivende for produktet, entydige og klart definerte.

Unipolar eller dipolar skala?

På en intensitetsskala måles styrken av en gitt egenskap. En unipolar skala går fra «ingen» til «mye» av den enkelte egenskapen, for eksempel søthet. En dipolar skala får vi ved å sette sammen to unipolare skalaer for egenskaper som vi tror er de absolutte motpoler til hverandre, for eksempel tørr og saftig.

De ord som velges i bedømmelse av et produkt må være beskrivende for produktet, entydige og klart definerte. For svært mange egenskaper er det vanskelig å si med sikkerhet hvilke som er nøyaktige motsetninger til hverandre, og det er en viktig grunn for å bruke unipolare skalaer. Både de sensoriske dommerne og leserne av resultatene får en enklere



oppgave på denne måten. Dersom en prøve er for eksempel tydelig tørr, vil den ikke samtidig være saftig. Det kan derfor være bortkastet å bruke begge disse egenskapene ved beskrivelsen av et produkt, velg det ordet som er mest naturlig å bruke for produktet (for eksempel saftighet i kjøtt og tørrhet i kjeks).

Det er vanlig å merke endene av skalaen med ord som for eksempel «ingen» og «tydelig» eller «noe» og «mye». For mange egenskaper kan det være vanskelig å tenke seg en prøve som inneholder «ingen» intensitet, eksempelvis gjelder dette egenskaper som har med tekstur eller konsistens å gjøre. Det er enklere å tenke seg egenskaper som dreier seg om farge, lukt og smak med «ingen» intensitet i en prøve. Det blir en trenings- og definisjonssak å forholde seg til ordene på skalaen. Skalaen må i alle fall trenes inn i forhold til de prøvene som skal bedømmes.

Det kan knyttes beskrivende ord til alle punktene på skalaen. Problemet er da at det kan diskuteres om betydningen av ordene gir en skala som har lik avstand mellom punktene. Eksempel på en bipolar skala med ni punkter ved måling av mørhet eller seighet:

9. ekstremt mør
8. meget mør
7. moderat mør
6. litt mør
5. verken seig eller mør
4. litt seig
3. moderat seig
2. meget seig
1. ekstremt seig





Eksempel på en ikke-symmetrisk intensitetsskala som også er unipolar:

- 0 ingen søt smak
- 1 merkbar søt smak
- 2,5 svært svak søt smak
- 5 svak søt smak
- 7,5 svak/moderat søt smak
- 10 moderat søt smak
- 12,5 tydelig søt smak
- 15 svært tydelig søt smak

Skalaer av denne typen kan både forlenges og forkortes. I disse to eksemplene er det brukt adjektiver som er logiske ved beskrivelse av økende/minkende intensitet av egenskaper. Men er det like stor opplevd avstand mellom adjektivene?

Ved bruk av disse skalatypene kan data vanligvis bli analysert med parametriske statistiske teknikker (se kapittel 11) dersom vi kan anta at lik numerisk avstand tilsvarer lik sensorisk forskjell. Hvis vi ikke har like intervaller mellom punktene på skalaen, bør dataene behandles som data fra kategorien ordinalskalet og dermed bli analysert med ikke-parametriske teknikker. Det er vanskelig å påstå at vi kan måle nøyaktige avstander av sensoriske stimuli. I praksis er likevel en variansanalyse såpass robust at vi får et resultat som vi kan stole på.

De fleste skalaer går fra null til et positivt tall. I enkelte tilfeller benyttes skalaer med nullpunkt på midten av skalaen og for eksempel fem positive og fem negative tallverdier. Skalaen kan være anvendelig når en har en referanseprøve (standardprøve) og skal måle om de andre prøvene har mer eller mindre intensitet av de angitte egenskapene. Skal en bedømme egenskapen «salt smak» i et produkt, kan en avmerke referanseprøven (R) ved null-punktet, og analysere om den andre prøven er mer eller mindre salt enn denne referansen (R).





Skalaen kan da se slik ut:

-5 -4 -3 -2 -1 0 1 2 3 4 5

ikke salt R mye salt

Det samme kan uttrykkes ved en skala fra 0 til for eksempel 11, hvor 0=ikke salt og 11=mye salt. Det kan da opplyses til dommerne om at referanseprøven har poeng 6. En fordel ved denne skalaen er at den ikke har negative tall som kompliserer behandlingen av tallmaterialet.

Ved en forbrukertest der aksept for det ene eller det andre produktet skal måles mot en referanse (for eksempel det markedsledende produktet), kan det være hensiktsmessig å velge en skalatype der referansen er plassert på midten av skalaen. Det blir da målt om en prøve er bedre eller dårligere enn referanseprøven (R), Just-about-right eller JAR-skala (kapittel 5).

Kontinuerlig eller diskret skala?

Når vi skal konstruere en skala for måling av en sensorisk egenskap, kan vi velge mellom en kontinuerlig skala som er en linje uten avmerkinger, eller en diskret skala som er en linje med oppmerkede delstreker.

En kontinuerlig skala (A):

En diskret skala (B):



Fordelen med den kontinuerlige skalaen (A) er at dommerne kan avmerke prøven hvor som helst på skalaen, de er ikke bundet av noen merker eller tall på skalalinjen.





Mange av data-innregistreringssystemene som finnes på markedet i dag, gir muligheten til å bruke kontinuerlig skala. Datamaskinen omregner markeringene til tallverdier for videre databehandling. Ved manuell lesing av bedømmelse utført på kontinuerlig skala, kan en benytte linjal med millimeteravlesning for korrekt registrering.

Når en benytter den diskrete skalaen (B), kan en risikere å miste informasjon om forskjeller mellom prøver som er svært like. Ofte vil dommerne ønske å gi karakterer mellom de vertikale linjene eller boksene, men det medfører selvsagt at panellederen burde utarbeide en annen skala med flere vertikale streker eller bokser på bedømmelseslinjen. Det betyr altså en forlengelse av skalaen. I slike tilfeller ser en da ofte at dommerne bare benytter deler av skalaen (den ene halvdel eller bare midten), slik at i realiteten ingenting blir vunnet ved å «utvide» skalaen.

Det blir brukt skalaer av forskjellig lengde, fra 3 til 100 skalaenheter. Målet er at skalaen skal være stor nok til å gi mulighet for å skille mellom små nyanser i prøvematerialet, men ikke så stor at dommerne bare benytter deler av skalaen. Det har vist seg at ved bruk av skala på 100 enheter, blir skalaen redusert ved at dommerne praktisk omarbeider skalaen ved bare å bruke hver femte eller tiende del. I dag brukes mest 5-, 7-, 9- og 15-punkts skala i sensorisk analyse.

4.2. Forskjellstester

Lene Waldenstrøm, Høgskolen i Sør-Trøndelag

Forskjellstester brukes for å undersøke om det er sensoriske forskjeller mellom prøver. Noen tester egner seg hvis den forventede forskjellen antas å være svært liten mens andre igjen benyttes der man ønsker å bekrefte en antatt forskjell eller finne eventuelle preferanseforskjeller.

Forskjellstester, eller differansetester kan deles inn i to grupper, generelle forskjellstester og spesifikke forskjellstester.

Det man kan få svar på ved bruk av de generelle forskjellstestene triangeltest, duo-trio-test og to-av-fem-test er om det beviselig er forskjell





mellom *to* prøver, gjerne kalt prøve A og prøve B. Hvor mange ganger den enkelte prøven serveres bestemmes av metoden. Med de generelle forskjellstestene stilles spørsmålet: Eksisterer det en forskjell? De oftest brukte metodene er triangeltest, duo-trio-test, to-av-fem-test og A- ikke A-test. Disse testene gir bare et generelt svar på om det er sensorisk forskjell mellom prøvene. Ved mindre endringer i ingredienser eller produksjonsmetoder er disse testene gode å bruke. Man vet altså at det for eksempel er en teknologisk eller kjemisk forskjell mellom prøvene, men lurer på om forskjellen kan oppdages sensorisk. Det kan en generell forskjellstest gi svar på.

Spesifikke forskjellstester brukes når det er spørsmål om forskjell mellom prøvene for en konkret egenskap, eller om preferanse for en enkelt prøve. Metoder som brukes er partest og rangeringstest. Disse testene gir oss svar på om det er en merkbar forskjell i én angitt egenskap, eller om prøvene kan rangeres etter økende grad av preferanse.

Disse sensoriske metodene brukes også ved utvelgelse og trening av dommere i den hensikt å finne ut om de aktuelle personene har sensorisk evne til å skille mellom prøver som beviselig bare er litt forskjellige. Da vil panelleder alltid ha en fasit, og dommerens svar vil være enten feil eller riktig.

Felles prinsipper for metodene

Tvunget valg

Dommerne må avgi svar. Det er ikke lov å svare “kjenner ikke forskjell”. På denne måten vil selv forskjeller som er så små at dommeren nesten ikke registrerer det selv, bli fanget opp. Statistikken som ligger bak tabellene som er utarbeidet for beregning av sikkerhet i analysesvaret, tar høyde for at dommerne må velge en av prøvene selv om han eller hun ikke egentlig kjenner forskjell.

Variert servering

Det er viktig å sikre at serveringsrekkefølgen av prøvene blir variert mellom dommerne. De ulike metodene krever ulike oppsett som styres av panelleder ved hjelp av en serveringsplan (se 16. vedlegg, a). Hvordan serveringsplanene skal utformes avhenger av metode og antall





serveringsomganger. Dette er nærmere beskrevet under den enkelte metode. For å sikre at dommerne vurderer prøvene i den rekkefølgen serveringsplanen forteller så må det opplyses om rekkefølgen for bedømming (vanligvis fra venstre mot høyre) på bedømmelseskjemaet.

Koding

Forskjellstestene utføres i all hovedsak som blindtester der dommerne skal vite minst mulig om prøven eller produktet. Dette sikres ved å kode prøvene. Det er vanlig å benytte nøytrale tretalls-koder (se kapittel 3). Hvis man utfører tester i flere omganger med de samme dommerne så må prøvene få nye kodesett til hver omgang.

Studier viser at det å servere en eller to prøver som eksempler på de prøvene som skal inngå i analysen, altså en oppvarmingsprøve (eng. warm-up-sample), vil styrke resultatet av analysen ved at dommerne blir mer fokuserte på de aktuelle egenskapene ved prøvene. Det noteres ikke resultater fra en slik oppvarmingsprøve, den inngår ikke i den reelle analysen.

Hypotesetesting

Alle metoder baserer seg på hypotesetesting. Ved fastsettelse av 0-hypotesen, skal man gå ut fra at panelet ikke klarer å skille prøvene fra hverandre generelt eller for en angitt egenskap. Resultatet av testen vil gi oss svar på om 0-hypotesen skal forkastes eller ikke. For statistisk behandling av de ulike forskjellstester, se kapittel 11

Statistisk signifikans

Et signifikansnivå gir uttrykk for hvor stor sjanse man ønsker å ta for feilaktig å forkaste nullhypotesen. Vanligvis brukes 5 %, 1 % eller 0,1 % signifikansnivå. Ved signifikansnivå 5 % kan vi være 95 % sikre på at vi ikke forkaster nullhypotesen på feil grunnlag. Hvis vi ikke får forkastet nullhypotesen, betyr ikke det at nullhypotesen er sann. Dommernes sensitivitet for eksempel kan være avgjørende, lite sensitive dommere finner kanskje ikke forskjellen mellom prøvene, mens mer sensitive dommere finner en forskjell. Forkaster vi nullhypotesen, kan vi si at alternativ hypotese er sann med 95 % sikkerhet (ved 5 % signifikansnivå).





4.2.1. Triangeltest

Triangeltest brukes for å påvise om det er forskjell mellom to produkter, og er en av de mest benyttede forskjellstestene. Testen kan også brukes til å teste likhet, men det kommer vi ikke inn på i dette kapitlet. Testen brukes i de tilfeller hvor vi mener at forskjellen mellom produktene er svært liten og kanskje ikke er merkbar. Antall testomganger er begrenset fordi metoden kan være mentalt og sensorisk utmattende å utføre.

I praksis vil vi ofte ha flere enn to produkter vi vil teste. I slike tilfeller er det relativt ressurskrevende å gjennomføre en triangeltest, fordi det kreves mange triangler. Noen produkter skal kanskje smakes og luktes på, mens andre bare skal sees på eller føles på. De forskjellige sansene våre trenger varierende tid for restitusjon (se kapittel 2, sansene) så avhengig av produkt er det derfor en forskjell i hvor mange prøver som kan analyseres i løpet av en gitt tid. Matvarer skal som oftest smakes på og da vil dommerne fort kunne bli smakstrette, noe som ødelegger testen. Ofte kan det da lønne seg å bruke andre tester, som for eksempel poengbedømmelse.

Prinsipp og gjennomføring

Hver dommer presenteres for tre prøver samtidig (et triangel), hvorav to er identiske og én er forskjellig fra de to andre. Hver prøve er gitt tilfeldige koder, se eksempel i kapittel 16, vedlegg a) hvor prøve A har fått to forskjellige koder og prøve B har to andre tilfeldige koder. Dommerens oppgave er å peke ut den ulike prøven. Hvis dommeren ikke merker noen forskjell, må han likevel foreta et valg. Dette betyr at dommeren da må velge én av prøvene. Sannsynligheten for å velge den ulike prøven selv om man egentlig ikke merker forskjell på de tre prøvene, er $p=1/3$ (33,33 %).

Siden forskjellen mellom prøvene er svært liten kan det være fristende for en dommer å svare «ingen forskjell» dersom han er litt usikker og slett ikke ønsker å svare «feil». Tabellene som brukes for å se om dommerpanelets svar viser at prøvene er merkbart forskjellige eller ikke, er utarbeidet med utgangspunkt i at dommerne må velge en prøve som den forskjellige (se kapittel 11, statistikk). Framgangsmåten for





gjennomføring av triangeltest hvor dommerne kan få mene «ingen forskjell» er imidlertid beskrevet i ISO-standarden for triangeltest (ISO 4120:2004) med den statistiske beregningen som da kreves for rett konklusjon.

Når man skal lage triangler med de to prøvene A og B, vil det være seks mulige kombinasjoner:

ABB BAB BBA BAA ABA AAB

Disse seks trianglene utgjør ett kombinasjonssett med ulik prøvepresentasjon til 6 dommere. Hvis dommerantallet ikke er delelig med 6, kreves en tilpasset framgangsmåte. Når dommerantallet for eksempel er 16, tenker vi oss tre kombinasjonssett, altså 18 prøver. To sett (12

Eksempel på gjennomføring av triangeltest:

1. Er det sensorisk forskjell mellom A (lagret tomatpuré) og B (ny type tomatpuré)?
2. Panelet består av 24 dommere. Det lages 4 hele sett med de 6 serveringskombinasjonene ABB, BAB, BBA, BAA, ABA, AAB. I denne testen er det kun en prøveomgang noe som gjør at vi kan dele de 24 dommerne i to grupper à 12, der den ene gruppen får prøve A (ordinær tomatpure) som den ulike prøven, mens den andre gruppen får B (ny type) som ulike prøve. Vurderingsrekkefølgen sikres ved at det lages individuelle bedømmelseskjema til hver dommer.

Hvis vi her hadde ønsket flere prøveomganger med de samme dommerne måtte vi ha sikret tilfeldig fordeling av prøvene da det ikke må være noen form for systematikk i serveringsrekkefølgen.

3. Antall riktige besvarelser (de dommerne som har klart å plukke ut den ulike prøven) summeres. Ved å gå inn i tabell for triangeltest for 24 dommere (kapittel 16 Vedlegg b, del a binomisk fordeling) finner vi det antall besvarelser i den ulike prøvens favør, som er nødvendig for å kunne forkaste nullhypotesen ved de ulike signifikansnivåene.

Resultatet fra en triangeltest kan vise at det er en forskjell mellom de to analyserte prøvene. Ofte er det et tilfredsstillende svar, men ikke dersom det også er behov for å vite hvilken egenskap som er årsak til forskjellen. Vi kan for eksempel benytte en partest og spørre panelet hvilken prøve som er mest.....(søt, rød, likt) eller en beskrivende test som gir en intensitetsbeskrivelse av flere av produktets egenskaper (se 4.3.).





triangler) lages og deles ut i sin helhet til 12 dommere, fra det tredje settet velges fire prøver, to med A som ulik og to med B som ulik.

Det er viktig å sikre at prøvene bedømmes i angitt rekkefølge (ABB/BAB osv.) fordi den ene prøven ikke skal ha noen fordel eller ulempe av å bli bedømt som første, andre eller tredje prøve. A-prøven og B-prøven skal serveres som singelprøve like mange ganger. Dersom produktet har en sterk smak vil det kanskje være en mulighet at dommeren ikke har klart å skylle ut smaken fra en prøve før den neste smakes på (carry-over effect). Det skal instrueres i nødvendigheten av å skylle munnen mellom hver prøve og vente med å smake på neste prøve til smaken fra den forrige er helt borte.

Uavhengig av om vi benytter papir eller databasert besvarelse kan vi sikre en variert og kontrollert serveringsrekkefølge enten ved å:

a) lage individuelle bedømmelsesskjema med koder og angitt vurderingsrekkefølge. Det er viktig å passe på at prøvene som blir servert står i samme rekkefølge som på bedømmelsesskjemaet. Denne fremgangsmåten bør velges ved bruk av utrente dommere og ved utvelgelse og trening av dommere.

b) skrive på skjemaet at dommeren skal smake fra venstre mot høyre og sikre at rekkefølgen blir riktig for hver enkelt dommer idet vi serverer prøvene. Denne fremgangsmåten bør kun benyttes ved bruk av trent panel.

Samtidig må vi, ved flere prøveomganger med de samme dommerne, sikre randomisert servering. Det må ikke være noen form for systematikk i serveringsrekkefølgen.

Eksempel på bedømmelsesskjema vises i kapittel 16. vedlegg a. Det er lov å be dommerne kommentere prøvoforskjellen på skjemaet, men ikke stille ledende spørsmål. Imidlertid er det ingen verdi i de kommentarene som blir gitt av en dommer som ikke har klart å plukke ut den single prøven i triannglet. Kommentarer gir ikke grunnlag for statistisk behandling.





Bearbeiding av resultatet:

Først må man undersøke om noen av skjemaene er feil og mangelfullt utfylt. Det kan skje når dommerne gir sine svar på papirskjema, men svært sjelden når besvarelsen registreres elektronisk fordi slike programmer sikrer fullstendig besvarelse før dommeren kan avslutte bedømmelsen.

Når vi må plukke ut et skjema på grunn av feil utfylling, kan det bli skjeheter i fordelingen av prøve A og B. Dette kan rettes opp ved å forkaste et tilfeldig valgt riktig utfylt skjema med motsatt serveringsrekkefølge. Ut fra de riktig utfylte skjemaene teller man opp hvor mange som har klart å plukke ut den ulike prøven. Deretter går man inn i tabellen for triangeltester i kapittel 16 b, statistiske tabeller, del



Figur 4.3 Presentasjon av kodete prøver til henholdsvis partest, triangeltest, duo-triotest og rangering.

Foto: Kjell Merok





a binomisk fordeling. I rubrikken for antall dommere finner vi fram det antallet dommere som var med i vår test. Antall riktig utfylte skjema tilsvarende antall dommere. Tabellen viser det minimum antall bedømmelser som har gitt det forventede svaret i den ulike prøvens favør. Vi kan finne ut om vi med en gitt sikkerhet kan påstå at prøvene er forskjellige.

4.2.2 Partest

Partest brukes for å undersøke om det er merkbare forskjeller mellom to prøver (A og B) når det gjelder en bestemt sensorisk egenskap. Dommerens oppgave er å velge den ene prøven framfor den andre på grunnlag av nærmere angitt kriterium, f.eks. lydstyrke, fargenyanse, fruktig lukt, syrlighet, hardhet eller sprøhet. Her er det svært viktig at panellederen forsikrer seg om at alle i dommerpanelet forstår hvilken egenskap prøvene skal bedømmes etter og at dommerne har samme egenskapsforståelse. Metoden kan også brukes til å avgjøre preferanse for den ene prøven framfor den andre, innenfor en gitt egenskap, eller generelt.

Testen brukes også til å velge ut og trene dommere. Da benyttes prøver med kjent styrkeforskjell av angitt egenskap, f.eks. søttest. Testen har altså en fasit, og panelleder vet hvilken prøve som er søttest og om dommeren har svart riktig eller feil.

Prinsipp og gjennomføring

Dommeren presenteres for to prøver og stilles spørsmålet: Hvilken av disse to prøvene oppfatter du som for eksempel søttest/rødest/surest/sprøest/best?

Hver dommer serveres to kodete prøver samtidig (et par). Dommerens oppgave er å peke ut den prøven han mener har mest av den egenskapen det spørres etter. Hvis dommeren ikke kjenner noen forskjell når det gjelder angitt egenskap, må han likevel avgi et svar. Sannsynligheten for å velge den ene prøven framfor den andre hvis man ikke kjenner forskjell på de to prøvene når det gjelder intensitet av angitt egenskap, er $p=1/2$ (50 %).





Vi tilbereder likt antall prøvesett av kombinasjonen AB og BA. Disse fordeles tilfeldig mellom dommerne. Antall sluttbedømmelser må være delelig med 2, og like mange AB som BA skal serveres. Det er også her viktig å sikre at dommerne bedømmer prøvene i riktig rekkefølge (se også triangeltest). Samtidig må vi, ved flere prøveomganger med de samme dommerne, sikre variert servering. Det må ikke være noen form for systematikk i serveringsrekkefølgen. Bare én egenskap kan prøves om gangen – det er altså ikke lov å be dommeren avgi svar på både søthet og konsistens på samme skjema og med samme prøver.

Bearbeiding av resultatet:

Man skiller mellom ensidig og tosidig partest avhengig av kunnskap om prøvene og problemstillingen for undersøkelsen (se kapittel 11.3). Det er kun i bearbeidingen av resultatene de to typene partest skiller seg fra hverandre. Både bedømmelsesskjemaer og måten man gjennomfører testen på er lik.

Ut fra de riktig utfylte bedømmelsesskjemaene finner man den prøven som flest dommere mener inneholder mest av egenskapen det spørres etter. Kapittel 16b viser tabell for ensidig- og tosidig partest. Tabellene forteller oss det minimum antall bedømmelser som kreves for at vi med en bestemt sikkerhet kan si at det er forskjell mellom prøvene, altså om 0-hypotesen kan forkastes ved de ulike signifikansnivåene.

4.2.3 Duo-trio-test

Duo-trio-testen kan sies å være en mellomting mellom partest og triangeltest. Her blir dommerne presenterte for tre prøver, hvorav to er like og den tredje er ulik. Her blir en av de to like prøvene merket som en standard eller referanse (R), og de to andre prøvene kodes etter de samme prinsipper som nevnt i kapittel 3. Her er det naturlig å tenke seg at bedriftens vanlige produkt bør fungere som referanseprøve, altså merkes R. Dommerens oppgave er å finne hvilken av de to kodede prøvene som er ulik standarden/referansen.

Ved bruk av duo-trio-metoden er det mulig å servere prøvene på fire forskjellige måter i forhold til hverandre:





AR-A-B (Referansen er en prøve fra A)

AR-B-A (Referansen er en prøve fra A)

BR-A-B (Referansen er en prøve fra B)

BR-B-A (Referansen er en prøve fra B)

Det er to måter å organisere denne testen på; ved en balansert referanseteknikk eller en konstant referanseteknikk. De to prøvene A og B kan begge inngå som referanseprøve (balansert referanseteknikk) eller bare den ene prøven kan inngå som referanse (konstant referanseteknikk).

Ved konstant referanseteknikk er den ene prøven alltid referanseprøve, og det er bare to mulige serveringskombinasjoner (AR-A-B eller AR-B-A, der A er kontrollprøven eller altså referanseprøven). Denne situasjonen er hensiktsmessig ved kvalitetskontroll der et trent panel er kjent med bedriftens standardprodukt, og dommernes oppgave er for eksempel å avgjøre om den løpende produksjonen avviker fra denne standarden eller referansen. I de tilfeller der det er et lite produksjonsparti (for eksempel ny resept) av den ene prøven i forhold til den andre, er det hensiktsmessig å velge en duo-trio-test med en konstant referanse som metode. Den prøven som det finnes mest av, velges følgelig som referanseprøve (R).

Ved balansert referanseteknikk sikrer vi lik serveringsbehandling for begge prøvene. Har vi et panel bestående av tolv personer, vil hver av de fire serveringskombinasjonene bli servert tre ganger. Dersom panelet består av ti personer, skal et sett der A er referanse og et sett der B er referanse, fjernes fra det tilsvarende oppsettet til tolv dommere.

En duo-trio-test er ikke egnet som metode for å måle preferanse eller akseptanse. Den er heller ikke egnet for å analysere hvilken egenskap som skiller prøvene eller intensiteten av den egenskapen som varierer mellom de målte prøvene.

Resultatbehandling se kapittel 11





4.2.4. To-av-fem-test

En annen metode som kan brukes for å sammenlikne to prøver er to-av-fem-test. Dommerne blir presentert for fem prøver der henholdsvis tre (A) og to (B) er like. Prøvene er kodet etter prinsipper for sensorisk analyse, slik at de serveres for dommeren som fem forskjellige prøver. Dommernes oppgave her er å avgjøre hvilke tre prøver som er like og følgelig hvilke to prøver som er like. Prosedyren og gjennomføringen av en denne analysen er i hovedsak lik det som er skrevet om triangeltesten. Den eneste forskjellen er at de to prøvene A og B som skal analyseres ved to-av-fem metoden må deles i henholdsvis to og tre prøver mens ved triangelmetoden er det bare den ene av A eller B som må deles i to prøver.

Denne metoden er mer krevende for dommerne enn en test med triangelmetoden i og med at i alt fem prøver må vurderes for å kunne gi svar på hvilke prøver som hører til i hver gruppe. Fordelen med metoden sammenlignet med en triangeltest er at det kreves færre dommere som deltagere i analysen for å oppnå et signifikant sikkert svar på oppgaven. Panellederer bør likevel være klar over at denne testen er vanskelig for dommerne, særlig dersom det er snakk om prøver med sterk lukt eller smak. Metoden egner seg kanskje best ved testing av produkter som ikke skal smakes på, men for eksempel bare fargebedømmes.

Denne testen kan serveres med 20 forskjellige prøvekombinasjoner:

AAABB BBBAA AABAB BBABA ABAAB.
BABBA BAAAB ABBBA AABBA BBAAB
ABABA BABAB BAABA ABBAB ABBAA
BAABB BABAA ABABB BBAAA AABBB

Består panelet av færre enn 20 dommere, bør serveringsordningen velges tilfeldig blant de 20 mulige kombinasjonene, men det må sørges for like mange triple A-prøver som triple B-prøver.

Sannsynligheten for å velge en av prøvene fremfor en annen hvis man ikke kjenner forskjell på de fem prøvene er $p=1/10$ (10%).





4.2.5 A – ikke A test

Dommerne får presentert to prøver for trening i å gjenkjenne de sensoriske kvalitetene i hver av dem. Deretter skal de ukjente prøvene bedømmes. Det er altså to prøver; A og B, som her kalles hhv. A og ikke-A.

Det er flere muligheter for serveringsprosedyre ved A- ikke A metoden. Enten kan vi servere en og en prøve som skal sammenlignes med den innøvde A-prøven, eller vi kan servere to prøver (et par) som skal sammenlignes med A-prøven.

Den første varianten kan sammenlignes med en partest, med den forskjell at den ene prøven her bare eksisterer i hukommelsen til dommeren. Flere prøver som er kodet, noen som er lik «A-prøven» og noen som er lik «ikke-A-prøven», blir presentert for sammenligning med den innlærte «A-prøven». Dommernes oppgave er å avgjøre om de ukjente prøvene er lik den innlærte A-prøven.

I den andre varianten, som er den mest brukte kan vi se en likhet med en duo-trio-test, men her med den forskjellen at A-prøven som er den kjente prøven bare eksisterer i hukommelsen mens i duo-trio-test er den merket som Standard eller Referanse. Et par (2) kodete prøver skal bedømmes for å avdekke om en eller begge eller ingen i paret er lik med hukommelsen om A-prøven. Det er fire forskjellige serveringskombinasjoner, AA, BB, AB, BA. Disse parkombinasjonene skal serveres i tilfeldig rekkefølge til dommerne.

Statistisk analyse av data følger samme prinsipper som for partest.

A – ikke A metoden brukes gjerne når det ikke er mulig å presentere helt identiske kopier av en prøve, for eksempel to spekesinker med litt forskjellig fettmarmorering eller to treplater med litt forskjellig kvistmønster som kan forårsake prøvevariasjon. Forskjellen som følge av variasjon i utseende eller fargenyanse må imidlertid være så liten at den kun kan oppdages når prøvene plasseres ved siden av hverandre. Derfor er det i denne metoden vesentlig at A-prøven skal studeres nøye og lagres i hukommelsen før den fjernes.





4.1.6 Rangering

Når flere prøver skal sammenlignes for intensitet av en spesiell egenskap og man ønsker å vite rekkefølgen av prøvene med tanke på denne egenskapen, er det hensiktsmessig å velge rangeringsmetode. Metoden brukes også som en sortering av prøver før anvendelse av andre sensoriske metoder. Dessuten brukes rangeringstester til å velge ut og trene dommere.

Prinsipp og gjennomføring

Hver dommer presenteres for minst tre prøver, som skal rangeres etter et bestemt kriterium, f.eks. utseende, søthet, konsistens, farge eller preferanse. Antall prøver som skal rangeres bør ikke bli for stort fordi dommerne fort kan bli «smakstrette». Dommerne må informeres om hvilken grad av egenskapen som skal settes som nr. 1 (f.eks. minst søt) og hvilken som skal settes sist (f.eks. mest søt). Alle prøver skal rangeres, ingen får altså settes lik hverandre.

Det er viktig å unngå at prøvene serveres alle dommerne i samme rekkefølge. Dette sikres ved å lage individuelle bedømmelsesskjema med koder og angitt vurderingsrekkefølge (eksempel i kapittel 16 a). Samtidig må vi, ved flere prøveomganger med de samme dommerne, sikre at prøvene blir tilfeldig fordelt. Det må ikke være noen form for systematikk i serveringsrekkefølgen. Bare én egenskap kan prøves om gangen.

Beregning av resultater fra rangeringstest er vist i kapittel 11

4.3. Beskrivende tester

Marit Rødbotten og Mats Carlehög, Nofima

Kvantitative metoder som også kan kalles beskrivende – eller deskriptive metoder, er metoder som beskriver produktets sensoriske profil. Vi kan bedømme produktet som helhet eller bare deler av det. Hvor detaljert en skal gjøre undersøkelsen er avhengig av problemstillingen.

Dersom det er forventet at det kan avsløres en sensorisk forskjell, vil det være mest hensiktsmessig å benytte en kvantitativ beskrivende





metode istedenfor en forskjellsmetode. Ved en forskjellsmetode får vi svar på om to prøver kan oppleves som forskjellige. Dersom resultatet sier at det er forskjell mellom prøvene, er det rimelig å forvente at oppdragsgiveren vil ønske å få svar på hvilken egenskap, eventuelt hvilke egenskaper som varierer mellom prøvene og om forskjellen er betydelig eller minimal. Tid og penger kan spares ved at panellederen får detaljert informasjon om prøvene og undersøkelsens problemstilling, slik at riktig metode blir valgt med en gang.

Det er en selvfølge at det sensoriske personalet har taushetsplikt angående prøvenes problemstilling og eventuelt hvem som er oppdragsgiver. Ikke noe av denne informasjonen blir gitt videre til dommerpanelet fordi informasjon om problemstillingen som ligger bak undersøkelsen kan ubevisst eller bevisst påvirke resultatet fra dommerne.

Metoder som oftest blir brukt er kvalitetskontrolltest og beskrivende test (forskjellige profileringsmetoder). Andre metoder som ikke så ofte brukes, men som er velegnet i spesielle situasjoner er kvalitetsindeksmetode (QIM) og metoder hvor intensitet av sensoriske egenskaper måles over en bestemt tidsperiode, som for eksempel time-intensity metoden (TI) og Temporal Dominance of Sensation (TDS).

4.3.1 Beskrivende test (profilering)

Innen den sensoriske vitenskapen er beskrivende tester de som gir mest informasjon om det analyserte produktet. Her kan vi få detaljert informasjon om blant annet produktegenskaper, om ingredienser i produktet kan identifiseres eller om hvilke egenskaper som er av betydning for forbrukervalg.

Dette kapitlet vil først omhandle faktorer som er viktige for et best mulig resultat av en profileringsmetode som på engelsk ofte blir betegnet som en DA-metode, en Descriptive Analysis metode. Fra ca 1940 er det utarbeidet flere varianter av beskrivende tester som alle er i bruk i dag, og noen av disse blir beskrevet senere i dette kapitlet. Selv med noen vesentlige forskjeller er de alle opptatt av å gi en detaljert beskrivelse av alle eller enkelte av produkters sensoriske egenskaper.



ISO 13299:2003 beskriver trinnvis prosedyren ved profilering. Her vises de forskjellige trinnene som i etterkant blir kommentert.

Tabell 4.1 ISO 13299:2003. Trinnvis prosedyre ved profilering.

Trinn nr	Oppgave	Relevant ISO-dokument
1: Innrede sensorisk lokale	Innrede bedømmelsesbåser, prøveprepareringsrom, osv.	8589
2: Velge prøver som illustrerer de egenskapene som skal bedømmes	En eller to eksperter (produktspert, sensorisk leder) velger egnede prøver.	11035
3: Velge og trene dommere for prosjektet	Panellederen velger en egnet gruppe mennesker, trener dem ved bruk av valgte produkter.	8586 5496
4: Velge egenskaper egnet for beskrivelse av de aktuelle prøvene. Kan kombineres med trinn 3.	Panellederen velger blant eksisterende egenskaper, eller dommerne bedømmer produkter fra trinn 2 og foreslår egenskaper. Valget av egenskaper gjøres enten av panellederen eller i fellesskap med dommerne. Egnede referanseprøver velges der egenskapene blir demonstrert.	5492 6564 11035 11036
5: Bestemme rekkefølgen av de valgte egenskapene	Panellederen arbeider sammen med dommerne under trinn 3 og 4.	11035
6: Velge skala for måling av intensitet av de valgte egenskapene.	Panellederen velger den mest aktuelle skalaen.	4121 11056
7: Gjennomføre testen.	Dommerne bedømmer prøvene.	6658 6564
8: Rapportere resultatene.	Analysere resultatene ved hjelp av statistiske metoder. Presentere resultatene i form av tabeller og diagrammer. Trekke konklusjoner.	6564 13299



Kommentarer til de åtte trinnene i tabellen:

Trinn 1

Betingelsene for et godt resultat av en sensorisk analyse ligger for en stor del på egnetheten av de sensoriske lokalene og de omgivelsene der prøveprepareringen skal foregå (se kapittel 3). Avhengig av hva som skal analyseres må det legges til rette for en objektiv analyse (unngå forstyrrelser, lyder, lukter, variasjon i lystilgang og lignende).

Trinn 2

For å kunne gjøre en grundig undersøkelse av de aktuelle prøvene er det viktig å kjenne til produkter med tilsvarende egenskaper som de som er med i undersøkelsen. Dette vil ofte føre til at vi raskere kan velge egenskaper som skal bedømmes i de aktuelle prøvene. Kjennskap til resepten kan antyde hvilke egenskaper som bør bedømmes. Men vær oppmerksom på at selv om resepten og instrumentelle undersøkelser dokumenterer innhold av et stoff som forventes å gi en bestemt sensorisk effekt, for eksempel brix-måling av sukkerinnhold i erter, er det ikke dermed sagt at et panel som ikke finner den forventede søtsmaken, har gjort en dårlig jobb. Andre smaker kan overdøve eller kamuflere søtsmaken.

Trinn 3

Se kapittel 3. Ved valg av dommere som er ansatt internt i bedriften er det viktig å ta i betraktning kunnskapen dommerne kan ha om prosjektet/produktene som skal analyseres. Forhåndskunnskap om spesiell behandling av prøvene vil lett forstyrre den objektive bedømmelsen og dommers detaljkunnskap må tas til følge når analyseresultatet skal vurderes.

Trinn 4

Dommerne får servert forskjellige prøver som er representative for oppdraget. I denne fasen er det ønskelig å få navngitt alle de egenskapene som er med på å gi prøvene den bestemte sensoriske profilen. Her er det aktuelt å la dommerne nevne alle ord som de finner i beskrivelse av produktet, gjerne kalt «hjernestorm» eller «brainstorming». Den





forhåpentligvis lange listen av egenskaper blir deretter bearbeidet til å inneholde de ønskede og viktigste sensoriske egenskapene. Det viser seg ofte at det er bedre å ha med litt for mange ord enn for få ord på den endelige listen av egenskaper, i alle fall når det er viktig å studerer detaljer ved produktet. På den andre side er det uhensiktsmessig å bruke tid på bedømmelse av egenskaper som ikke forventes å variere mellom produktene og heller ikke er av interesse for oppdragsgiveren.

Sensoriske egenskaper kan klassifiseres i forskjellige nivåer av detaljinformasjon, som eksempel kan vi tenke oss det mer overordnede ordet bærsmak i forhold til ordet jordbærsmak som er mer spesifikt. Det er nødvendig at alle dommerne er enige om forståelsen og definisjonen av hvert egenskapsord som skal bedømmes. Hvor detaljert (i dette eksemplet: bær eller jordbær) egenskapene skal beskrive produktet må avklares på forhånd i samråd med oppdragsgiver/produktansvarlig. Se Aromahjul (figur 4.4) for eksempel på gradert produktbeskrivelse

Gjenkjenning av egenskaper er avhengig av dommerens sensitivitet og tidligere erfaring med produkter med de aktuelle egenskapene. Trening i gjenkjenning av egenskaper i forskjellig intensitet gjøres ved først å bli presentert for det sterke og tydelige, for deretter å bli presentert for den samme egenskapen i svakere intensitet. Læring av gjenkjenning av en smak skjer ved en parsammenligning mellom den sterke smaken sammen med en stadig svakere smak.

Panelet skal være trent i de egenskapene som skal bedømmes. Panelet og panellederen blir enige om en definisjon av hver av de aktuelle egenskapene. Denne definisjonen følger også i rapporten, slik at oppdragsgiveren får samme tolking av egenskapene. Det advares mot å la et trent panel bedømme «totalinntrykk» av en prøve. Det er en hedonisk vurdering og skal følgelig utføres av forbrukere.

Trinn 5

Det er ikke likegyldig i hvilken rekkefølge egenskapene bedømmes. For matvarer er det oftest riktig å bedømme farge og utseende på prøven før lukt, smak og tekstur bedømmes. Selv om vi anstrenger oss til det ytterste for å oppnå en objektiv bedømmelse kan vi ikke





utelukke at dommere blir påvirket av tidligere erfaringer med produktene som skal bedømmes. For eksempel er det ofte en mindre intens jordbærsmak på et «brunt», gammelt jordbærsyltetøy enn et ferskt, rødt syltetøy. For bedømmelse av smak i dette produktet er det hensiktsmessig å presentere prøvene under rødt lys som kamouflerer den rette fargen, for at smaken skal bli mer objektivt bedømt (se kapittel 2). Smaksbedømmelsen i rødt lys bør følgelig gjennomføres før fargebedømmelsen i vanlig lys.

Trinn 6

Samtidig som det trenes på å kjenne igjen og vurdere intensitet av de valgte egenskapene, er det aktuelt å trene på bruk av intensitet på en gitt skala. Denne fasen med trening, ofte kalt for-forsøk eller treningsforsøk, er svært avgjørende for resultatet av hele analysen. Nøyaktigheten og presisjonen av bedømmelsen avhenger av paneldeltakernes felles oppfatning av sensoriske egenskaper, med definisjon og intensitet. Tid og penger er dessverre ofte begrensende faktorer for denne treningsperioden med de konsekvenser som følger for nøyaktigheten av analysen.

Ved for-forsøkene er det en fordel at alle prøvene kommer fra samme prøveuttak eller batch, slik at muligheten for enighet mellom dommerne blir størst mulig. Dommerne skal her «kalibreres» som et instrument. En eller gjerne flere prøver som er identisk fra dag til dag, og som er forskjellige fra hverandre med hensyn til en eller flere sensoriske egenskaper, vil være egnet som treningsobjekt. Disse prøvene forenkler kalibreringen av panelet mot skalaen.

Det er ikke nødvendig at alle dommerne gir den samme tallkarakteren for intensitet av prøven. Men det er svært viktig at den dommeren som bruker lave tall på skalaen alltid gjør det ved presentasjon av samme intensitet av en stimulus. Og tilsvarende for de som bruker de høye tallene på skalaen.

Det er gjennomsnittet av alle dommenes besvarelser som utgjør analyse-resultatet, ikke enkeltdommernes.

Bedømmelsesskjemaet, egenskapsforklaringen og framgangsmåten for bedømmelsen presenteres for dommerne. Alle dommerne serveres



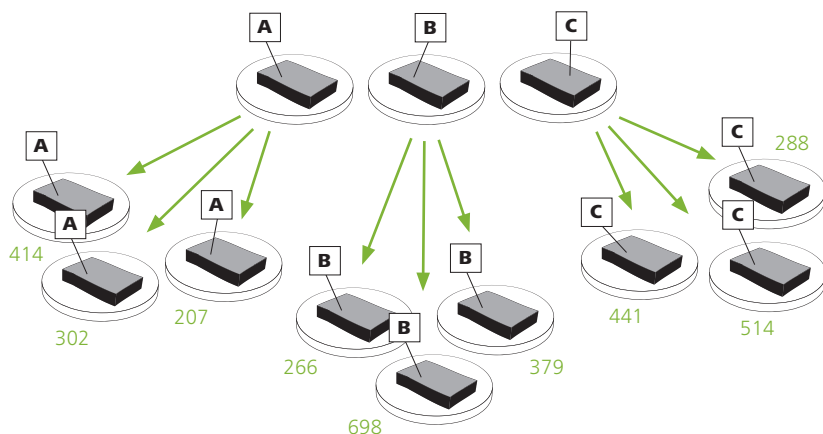


to eller flere prøver som er forskjellige, og testen gjennomføres. Ikke alltid er dommerne helt samkjørte ved første gangs testing av prøvene. Derfor bør panelet kalles sammen for en diskusjon for eventuelt justering av egne bedømmelser.

Følgende punkter kan ligge til grunn for diskusjonen: dommerne har ikke lik forståelse av egenskapene, er uenige om rekkefølgen på prøvene med tanke på intensitet av en egenskap, eller den valgte referanseprøven er uegnet.

Gjenta for-forsøket (treningen) til dommerne er samstemte ved bedømmelse. Under for-forsøkene bør en være oppmerksom på om noen av de planlagte egenskapene kan sløyfes fordi de ikke gir noen informasjon om variasjon mellom prøvene (alle er like søte), eller om egenskaper viser samvariasjon (enten tørr eller saftig).

En måte å øke sikkerheten i resultatet fra en sensorisk analyse på er å gjenta analysen en eller flere ganger. Det forventes at en dommer skal kunne gjenta sin karaktergivning likt for prøver som er like, for eksempel saft fra samme flaske servert i tre glass med forskjellige koder bør få samme karakterer (scoring) på innholdet i hvert glass.



Figur 4.4 Gjentak av prøver

Illustrasjon: Julie Gjermundsen





Ved bestemmelse av hvor mange repetisjoner som bør gjennomføres av en analyse, er det riktig å velge repetisjonen der variasjonen er størst. Forventes det større variasjon mellom to produkter (innen en produksjon, en behandling eller søsken i en flokk griser) enn variasjonen i et panels bedømmelse av ett og samme produkt, ja da er det større behov for repetisjon av prøvemeditet (flere griser, flere fisk, o.l.) enn av antall sensoriske gjentak av bedømmelser. Dette gjelder særlig for beskrivende tester, men til dels også for kvalitetskontrolltester.

Trinn 7

Se kapittel 3, praktisk gjennomføring. Panelansvarlig må ha kontroll med prepareringen av prøvene, serveringsrekkefølgen og kodingen av prøvene. Randomisering av prøveservering er viktig og det er kun i spesielle tilfeller det kan anbefales å servere prøver gruppevis med hensyn på intensitet av egenskaper. Panellederen bør være bevisst på effekten av å servere prøver med lav intensitet av egenskaper før de sterkere prøvene serveres. En slik serveringsrekkefølge vil gi dommerne en forventning om at intensiteten skal øke og dermed ikke nødvendigvis bedømme slik de sanser eller opplever prøven. Eksempler kan være pølser med lite krydder servert før pølser med mye krydder, eller drikke med lite/mye alkohol, eller lyder med høy/lav tone. I alle tilfeller er det viktig med nok av tid mellom hver prøve slik at smak og lukt fra den ene prøven er borte før den neste prøven skal bedømmes.

Trinn 8 Data fra den sensoriske analysen skal behandles med egnede statistiske metoder og resultatene rapporteres lett forståelig og med grundig henvisning til rådata. Rapporten bør beskrive prøvematerialet, dato for bedømmelse, valg av sensorisk metode, paneltreningen, liste med definisjoner av alle bedømte egenskapene, serveringsrekkefølgen av prøvene i tillegg til resultatet med konklusjon.

Språkutvikling, sensorisk leksikon og aromahjul

Wenche Emblem Larssen, Møreforskning AS

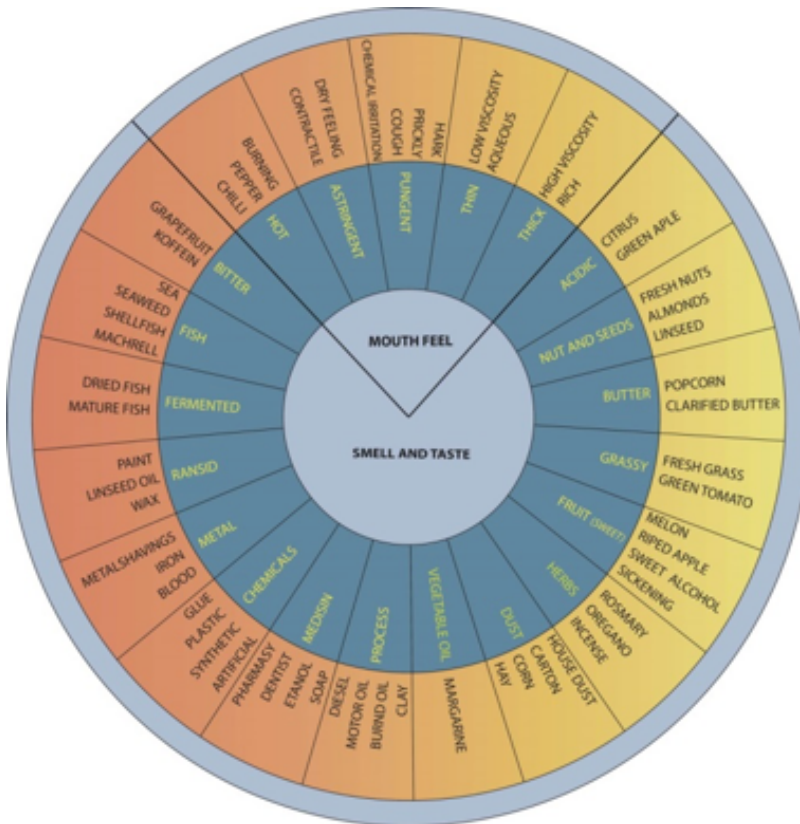
Før en kan gjennomføre en sensorisk vurdering av en produktgruppe må en samle inn ord og uttrykk som kan benyttes til en systematisk beskrivelse av de sensoriske egenskapene. Dette kan gjøres



Praktisk eksempel: Marine oljer (Larsen og Hersleth 2014)

Assosiasjonstest gjennomført på 44 marine oljer av ulikt råstoff og ulik kvaliteter.

Gjennomført av et profesjonelt sensorisk panel og av 6 bedriftspanel. Testen genererte 184 ulike ord og uttrykk der 100 var unike. Sortering og gruppering reduserte antallet til 60 ulike nøkkelord gruppert i 21 ulike kategorier. Kategoriene representeres av de nøkkelordene som gikk mest igjen i assosiasjonstesten. Grafisk fremstilt som et aromahjul.



Figur 4.5 Aromahjul for marine oljer

Illustrasjon: Larsen og Hersleth, 2014



ved å gjennomføre en assosiasjonstest. En assosiasjonstest ledes av en pannelleder og deltakerne kan være sensoriske dommere eller utvalgte fokusgrupper (konsumenter, næringsaktører ol.) Dommerne blir presentert for et utvalg av produktene som de skal vurdere for å bli kjent med de sensoriske variasjonene som finnes. Det kan være nyttig å gjennomføre flere runder med innsamling av ord og uttrykk, på et vidt spekter av produkt, med store individuelle forskjeller og ved bruk av ulike dommerpanel for å få en bredes mulig tilnærming (Lyon og Watson 1994, Monteleone og Langstaff 2014)

Den endelige listen med egenskapene (ord og uttrykk) generert gjennom assosiasjonstest blir deretter normalt redusert slik at den kun inneholder egenskaper som gir en klar og nøyaktig beskrivelse. Reduksjon kan gjøres ved at en først sorterer egenskapene alfabetisk og fjerner de som overlapper hverandre, før en deretter grupperer de uttrykkene som hører sammen (Lyon og Watson 1994, Mojet og de Jong 1994, Gawel et al. 2000, Koch et al. 2012). Det er også mulig å vekte de egenskapene som går mest igjen mellom dommerne og eventuelt mellom panelene som de mest viktige. Egenskapslisten kan presenteres som et sensorisk leksikon der de ulike grupperingene defineres eller grafisk som et aromahjul.

Andre beskrivende metoder

Marit Rødbotten og Mats Carlehög, Nofima

Andre metoder som også hører til i gruppen av beskrivende metoder er grundig beskrevet i Lawless og Heymann, 2010. Noen av metodene er kort beskrevet her.

Smaksbeskrivelse. (Flavor Profile[®], FP)

Metoden beskriver smak i produkter i fem kategorier; 1) egenskaper, 2) intensitet av egenskapene, 3) egenskapenes rekkefølge med hensyn på når de oppdages i munnen, 4) ettersmak og 5) totalinntrykk. Til å begynne med ble egenskapene målt etter følgende fem kategorier: Ikke merkbar, så vidt merkbar, litt, moderat, sterk. Denne skalaen har etter hvert blitt utvidet til en 17-punkt skala med bruk av halve poeng og pluss og minus for de fem kategoriene. Et godt trenet panel på 4-5





personer bedømmer prøvene og deretter diskuterer de seg frem til et svar de alle er enige om. Panellederen er aktivt med i bedømmelsen. Altså ikke et gjennomsnitt av dommernes bedømmelse, men en felles enighet om svaret.

Kvantitativ beskrivende analyse.

(Quantitative Descriptive Analysis®, QDA)

Denne metoden ble utviklet for å korrigere punkter som ble sett på som et stort problem for objektiviteten ved FP-metoden. QDA godtar ikke at resultatet fra analysen kommer som følge av konsensusdiskusjon mellom dommerne, men det skal regnes gjennomsnitt av alle dommernes karakterer for hver egenskap. I tillegg skal ikke panellederen være aktivt med i bedømmelsen og det skal brukes en ustrukturert linjeskala for hver egenskap.

Antall dommere bør være 10-12 og de velges med bakgrunn i gode resultater fra forskjellstester og evne til muntlig å karakterisere produkter. Egenskaplisten blir bestemt etter at dommerne er blitt presentert for mange prøver/produkter av lignende karakter som det som skal bedømmes, og de egenskapene som er mest egnet for prosjektets problemstilling blir valgt for hovedprosjektet.

Det legges vekt på at dommene bedømmer prøvene individuelt i bedømmelsesbåser. Det er ikke påkrevd at dommerne gir samme karakter på prøvene som sine meddommere, men det er viktig at hver dommer kan reprodusere sine egne bedømmelser ved å gi samme karakterer til to prøver som er like, men er kodet forskjellig.

Data kan behandles med statistiske modeller og presenteres med grafiske bilder.

Teksturbeskrivelse. (Texture Profile®, TP)

FP metoden ble raskt brukt i modifiserte varianter, som etter hvert endte opp som en spesiell metode egnet til bedømmelse av tekstur av matvarer, nemlig Textur Profile metoden. Metoden gir mulighet til å studere produkter i alle stadier av en tyggeprosess. Produkters geometriske- og mekaniske egenskaper blir målt i tillegg til saftighet, fethet, og det kan måles i hvilken rekkefølge de forskjellige





egenskapene er mest tydelige. Intensiteten av egenskapene blir angitt på en 15-punkt skala (med halve tall i tillegg) og hvert punkt på skalaen har en referanse i et kjent produkt. For eksempel har egenskapen hardhet følgende referansepunkter:

2,5=kokt eggehvite i fem minutter, 9,5=peanøtt og 14,5=hardt sukkertøy. Det er helt avgjørende for riktig bruk av TP at alle dommerne får den samme og grundig trening i bruk av skalaen med de samme referanseproduktene.

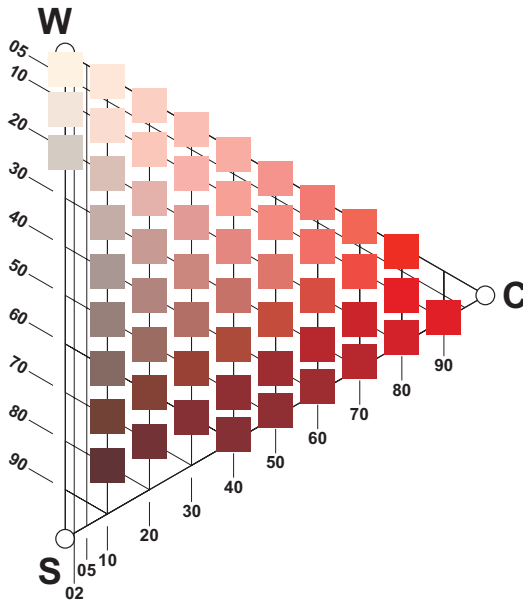
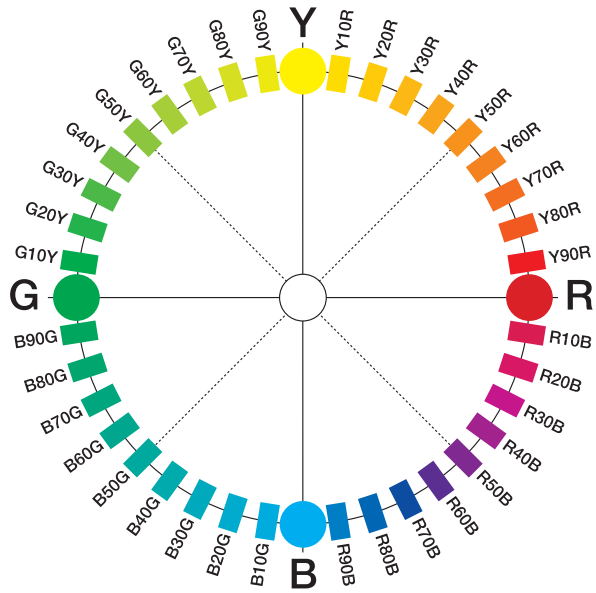
Spectrum metode. (Sensory Spectrum ®)

Denne metoden bygger på en kombinasjon av FP- og TP-metodene. Her kan alle sensoriske sanser brukes for bedømmelse av utseende, lukt, smak, tekstur og følelse. Ikke bare mat og drikke kan bedømmes med Spektrummetoden, men også hygieneartikler som kremer og papir. Skala brukes fra 0-15 med referanseprodukter for nesten alle punkter på skalaen for «alle» sensoriske egenskaper, inklusive grunnsmaker. Eksempel på salt smak: 5 poeng=«salted soda crackers, Premium», 9,5 poeng = «salted potato chips, Frito-lay». Alle egenskapsordene finnes i et leksikon som brukes for de forskjellige produktene og dommerne er ikke delaktige i utvikling av egenskapslisten som skal brukes i prosjektene. Dommerne trenger lang trening for å gjenkjenne de forskjellige egenskapsintensiteten i de forskjellige referansene.

4.3.2. Kvalitetskontrolltest

Serieproduserte varer som forventes å være like bør kontrolleres nettopp med tanke på likhet. Den sensoriske kvaliteten måles opp mot en standard eller referanse og metoden som oftest blir brukt er en poengbedømmelse, også kalt numerisk gradering eller «scoring». Den eller de egenskapene som skal bedømmes, blir kvantifisert ved hjelp av en tallskala (poengskala). Det vanligste er at den ene enden av skalaen angir riktig kvalitet, og den andre enden angir avvikende eller dårlig kvalitet. Metoden er illustrerende og anvendelig, men den stiller store krav til dommerne og panellederen. Grundig trening i bedømmelse av produktet og bruk av skalaen er viktig. Til denne treningen er det nødvendig å arbeide med produktteksempler med riktig kvalitet og





Figur 4.6 og 4.7 Fargesirkel og fargetriangel

Illustrasjon: NCS

110 Sensorikk – Måling med menneskelige sanser





eksempler med avvik fra riktig kvalitet. Et syltetøy kan ha riktig farge, lukt og konsistens, men ha feil intensitet av søtsmak.

I produktutviklingsfasen blir det bestemt hvordan produktet skal være (se kapittel 6) og senere kvalitetskontroll skal dokumentere grad av likhet eller ulikhet med det valgte referanseproduktet (se kapittel 8).

Den produktansvarlige bør være delaktig i prosessen når det sensoriske panelet skal trenes i gradering av avvik i den avvikende prøven, han/hun bør være delaktig i hvor «strengt» de forskjellige avvikene skal markeres.

Som oftest vil aksepttester fra forbrukere gi antydning om hvor grensen bør gå for salgsvare eller ikke, og en eller flere egenskaper kan være avgjørende her. Det kan for eksempel være viktig å angi avvik i fargen på prøven, for eksempel om jordbærsyltetøyet har fått en avvikende farge enten på grunn av lagring eller spesielle ingredienser. Å huske grad av farge er ikke alltid like lett. Som et hjelpemiddel kan en for eksempel benytte fargeprøver fra NCS Natural Colour System®©. NCS er et fargesystem som beskriver farger slik vi ser dem. I 1984 ble NCS en Norsk Standard innen fargebeskrivelse og systemet er brukt i hele Europa, noe som gjør det egnet som referansesystem for sensoriske analyser (www.ncscolour.no). Ved hjelp av NCS fargesirkel (figur 4.6) og NCS fargetriangel (figur nr 4.7) kan de fargene som representerer produkter med riktig og med avvikende farge lett illustreres som hjelpemiddel for dommerne til å gi korrekt angivelse av fargen på prøvene. For trening og kanskje også til bruk under selve bedømmelsen kan enkeltfarger i fargetriangler settes sammen etter økende grad av avvikende farge for produktet, som en rangering av fargene. Hver farge angir et markeringsnivå på bedømmelseskalaen. Det er mulig å finne farger som gjenspeiler alle typer produkter i NCS-systemet.

Det viser seg at dommerne ofte ikke benytter seg av hele skalaen. Poenggivingen er vanskelig å reproducere og skalaen har lett for å «gli» etter en tid. For å motvirke at skalaen glir, bør analysen også inneholde referanseprøver eller standarder som er kodet på lik linje med de ukjente testprøvene. Et veltrent panel bør kunne gi den kodete referanseprøven samme karakter gjennom hele prosjektet. Det kan i noen tilfeller være aktuelt å presentere referanseprøven merket





som «referanse» sammen med de kodede prøvene slik at dommerne kan oppfriske sin hukommelse av de forskjellige egenskapene og intensiteten av disse.

Det bør foreligge en skriftlig sensorisk spesifikasjon for hvert produkt, og denne spesifikasjonen skal i hovedsak beskrive det riktige produktet. En spesifikasjon må beskrive produktet med objektive termer. En beskrivelse av et produkt med for eksempel begrepene «passe salt» eller «riktig farge», er uegnet i denne sammenheng. For visuelle egenskaper som for eksempel mørke eller lyse flekker i produktet kan det være aktuelt å angi flekkens størrelse ved hjelp av standardiserte mål som for eksempel mm.

Ofte kan det være lettere å beskrive hvordan produktet ikke skal være. En slik negativ beskrivelse kan forekomme i produktspesifikasjonen. Da vil det være naturlig å beskrive kvalitetsfeil som har vist seg oftest i kontroll av produktet. Som eksempel kan vi tenke oss antall kvistmerker i et panelgulv eller mengde hull i en ost. En nomenklatur er en liste med aktuelle sensoriske egenskaper for produktet.

I kvalitetskontrollbedømmelsen må en på forhånd ha klart for seg hvilke krav en stiller til presisjonen av bedømmelsen. Hvor mange dommere som skal være med i panelet vil bli en avveining mellom behov for nøyaktighet i analysen, tid og økonomi. Det viktigste er at hver dommer er testet for sensitivitet for aktuelle produktegenskaper.

Vær oppmerksom på at et stort antall dommere ikke nødvendigvis betyr at presisjonen av bedømmelsen er stor. Dommernes sensoriske terskelverdi og deres trening i produktbedømmelsen er av avgjørende betydning for resultatet.

Inndelingen av skala til bruk i kvalitetskontroll er ofte enten 3-, 5- eller 9 poeng. Nordisk Metodikkomite for næringsmidler (NMKL) viser i prosedyre 16 (2005), Sensorisk kvalitetskontroll en sammenstilling av skalalengder.



Tabell 4.3 Sammenligning, 3-, 5- og 9-poengsskala

Kvalitetskategori	Skalalengde		
	3-poeng	5-poeng	9-poeng
Samsvar med spesifikasjon	3	5	9
			8
			7
Avvik fra spesifikasjon, men handeksvare	2	4	6
			5
		3	4
Vesentlig avvik, ikke handelsvare	1	2	3
			2
		1	1

4.3.3. Kvalitetsindeksmetode, QIM

Ferskhet på fisk er et viktig kriterium ved kjøp og salg av fisk og det er ikke alltid like lett å avgjøre fiskens ferskhet. En metode hvor forskjellige deler av rå hel, sløyet fisk blir vurdert i grad av ferskhet, kalles QIM-metoden, (Quality Index Method).

Metoden er basert på at man systematisk har fulgt utviklingen under tid av hel rå fisk lagret på is og dokumenterer viktige endringer av fiskens ferskhet. Viktige kvalitetsparametere som vurderes er skinn, øyne, gjeller, fiskekjøtt og blod. Det gis poeng fra 0 opptil 3 for hver parameter som vurderes. 0 angir mest fersk og 3 angir minst fersk (dårlig). Maksimumsummen (kvalitetsindeksen) av alle parameterne (24 for oppdrettslaks/*Salmo Salar*) viser en ubrukelig fisk for konsum, mens 0 poeng gis til en helt fersk fisk. Denne metoden er utviklet til bruk for de mest benyttede fiskearter og også annen sjømat. Metoden stiller store krav til dommers kunnskap til produktet og grundig trening i bedømmelse av produktet og bruk av skalaen er viktig. Vi viser til nærmere beskrivelse av metoden i www.qim-eurofish.com.



4.3.4. Tidsavhengige målinger

De mest brukte beskrivende metodene er metoder som måler maksimum intensitet av en egenskap. Det er imidlertid også av interesse å undersøke for eksempel når en smaksopplevelse oppstår og smakens endring i munnen med tanke på varighet og styrke.

Time-Intensity

Med Time-Intensity-metoden (TI) måles intensiteten av en egenskap over lengre tid. Bittersmaken i en sjokolade kan måles i et tidsintervall på for eksempel ett til tre minutter etter at prøven er puttet i munnen eller etter spyting/svelging. Det kan også være viktig å undersøke varigheten av smeltetiden av sjokolade i munnen, eller hvor lenge sjokoladens søt smak beholdes. De sensoriske dommerne skal her sanse intensiteten av en sensorisk stimulus, og samtidig angi intensiteten over et gitt tidsrom.

En utfordring ved TI-metoden er at den er best egnet ved analyse av bare en egenskap om gangen.

Temporal Dominance of Sensation

Temporal Dominance of Sensation (TDS) ble utviklet for å hjelpe på problemene som fulgte med TI-metoden, nemlig at TI bare kunne undersøke en egenskap om gangen. TDS er fortsatt under utvikling og det er ennå ikke etablert en bestemt standard prosedyre for gjennomføring av en analyse med bruk av denne metoden. Både trente og utrente dommeren kan bruke denne metoden. I Kapittel 5, forbrukeranalyser, beskrives denne metoden nærmere.





Referanser

- Amerine M.A., Pangborne R.M., Roessler, (1995). *Principals of Evaluation of Food*, Acad. Press, New York
- Ennis, Daniel M., (1993). The Power of Sensory Discrimination Methods, *Journal of Sensory Studies* 353-370
- ISO 10399, Sensory analysis – Methodology – Duo-trio-test
- ISO 13299, Sensory analysis – General guidance for establishing a sensory profile
- ISO 4120, Sensory analysis - Triangular test
- ISO 5495, Sensory analysis – Methodology – Paired comparison
- ISO 6564, Sensory analysis – Methodology – Flavour profile methods
- ISO 6658, Sensory analysis – Methodology – General guidance
- ISO 8587, Sensory analysis – Methodology – Ranking
- Kroll B.J., (1990). Evaluating rating scales for sensory testing with children, *Food Technology*. Vol. 11, pp. 78-86
- Land D.G., (1983). What is sensory quality? I: Sensory quality in foods and beverages: definition, measurement and control, Editor: Williams & Atkin, Chichester: Ellis Horwood
- Lawless H., Heymann H. (2010). *Sensory Evaluation of Food. Principles and practices*. Springer. ISBN 978-1-4419-6487-8
- Lawless H., Klein B.B., (1991). *Sensory Science Theory and Practice*, Ellis Horwood, Chichester, England
- Lawless H., Malone G., (1986). The discriminative efficiency of common scaling methods, *J. Sensory Studies*, 1, pp.85-98
- Lyon D., Francombe M.A., Hasdell T.A., Lawson K. (1992). *Guidelines for Sensory Analysis in Food Product Development and Quality Control*, Chapman & Hall, London
- Meilgaard M.C., Civile B., Carr T. (1991). *Sensory Evaluation Techniques*, (2 ed.) Boca Raton CRC Press, Inc.
- Munoz A.M., Civile G.V., Carr B.T. (1992). *Sensory Evaluation in Quality Control*, VanNostrandReinhold, New York
- Fortsetter neste side





Stone H., Sidel J.L. (1993). Sensory Evaluation Practices, Academic Press, London

www. qim-eurofish@com.

Referanser (Aromahjulet)

Gawel, R., A. Oberholster og I. L. Francis (2000). "A 'Mouth feel Wheel': terminology for communicating the mouth feel characteristics of red wine." Australian Journal of Grape and Wine Research 6(3): 203-207.

Koch, I. S., M. Muller, E. Joubert, M. van der Rijst og T. Næs (2012). "Sensory characterization of rooibos tea and the development of a rooibos sensory wheel and lexicon." Food Research International 46(1): 217-228.

Larssen, W. E. og M. Hersleth (2014). "Sensorisk kvalitet på marine oljer." Møreforskning. MA 14/16 51

Lyon, D. H. og M. P. Watson (1994). "Sensory profiling: a method for describing the sensory characteristics of virgin olive oil." Grasas y aceites 45(1-2): 20-25.

Mojet, J. og S. de Jong (1994). "The sensory wheel of virgin olive oil." Grasas y aceites 45(1-2): 42-47.

Monteleone, E. og S. A. Langstaff (2014). Olive Oil Sensory Science. First edition. Wiley Blackwell.





METODER FOR FORBRUKER- TESTER



Foto: Mari Svenningsen





METODER FOR FORBRUKERTESTER

Margrethe Hersleth og Valerie Lengard Almli, Nofima

5.1 Introduksjon

Innen forskning, produktvikling og kvalitetskontroll er det forbrukere som er den ultimate målgruppen for måling av sensorisk persepsjon. I vårt markedsorienterte samfunn har dette blitt mer og mer tydelig ved at forskningsprosjekter inkluderer hele verdikjeden (kunnskap om forbrukere og marked) og ved at industrien arbeider etter prinsipper for markedsdrevet produktutvikling og innovasjon. Det vil derfor være viktig både å benytte metoder for sensoriske målinger som inkluderer informasjon om forbrukeres preferanser og grad av aksept for produkter, samt deres matvaner og holdninger til mat. I dette kapitlet vil vi gjennomgå de mest vanlige tester som benyttes til slike formål.

5.2 Kvalitative studier

Kvalitative tester skiller seg fra preferansetesting og aksepttesting ved at man ikke innhenter data som egner seg for dataanalyse og statistiske tester, men at man innhenter kunnskap om spørsmål hvor det er nødvendig med samtaler og/eller diskusjoner med relevante forbrukere. Den mest vanlige form for kvalitative tester er dybdeintervju med en gruppe mennesker, noe som kalles en fokusgruppe (Lawless & Heymann, 2010). Fokusgrupper brukes innen markedsanalyse for å teste ut produktkonsepter og inkluderer ofte presentasjon av produkter, produkt informasjon og reklame. Men produktutviklere vil også ha stor nytte av metoden for å få direkte respons på produktets sensoriske egenskaper og produktets funksjonelle egenskaper sett fra forbrukerens synsvinkel.

118 Sensorikk – Måling med menneskelige sanser





I en fokusgruppe kan man diskutere forbrukeres motiver for kjøp og bruk av et produkt, deres vaner og atferd med produktkategorien, samt deres holdninger og verdier, og deres generelle følelser i forhold til et produkt eller en produktgruppe. Følgende trinn/prinsipper kan anbefales ved gjennomføring av en fokusgruppe:

- Definere problemstilling
- Spesifisere målgruppe og prinsipper for rekruttering
- Velge moderator
- Utarbeide og utteste intervjuguide
- Rekruttere deltakere
- Gjennomføre studien
- Analysere og tolke data
- Skrive rapport

Arbeidet med en fokusgruppe bør starte med å formulere en tydelig problemstilling slik at både oppdragsgiver og de som er involvert i gjennomføring av fokusgruppen har et tydelig bilde av hvilken kunnskap man ønsker å oppnå. Denne problemstillingen vil gi grunnlag for valg av målgruppe og rekruttering.

Valget av moderator er viktig. Det bør være en person som er trent i intervjuteknikk, som kommuniserer godt med målgruppen og som har evnen til å lytte og å balansere (legge til rette) en diskusjon mellom ulike mennesketyper. Det er viktig at moderator er nøytral i forhold til problemstillingen. I tillegg vil en viktig del av moderators rolle være å passe på at alle deltakere får anledning til å uttale seg om viktige spørsmål. Antall deltakere i en fokusgruppe er vanligvis mellom 8 og 12. En fokusgruppe bestående av 8 personer vil være enklere å håndtere for en moderator enn hvis antallet er større. En fokusgruppe bestående av 12 personer vil gi flere meninger og argumenter til analysen, men det vil sannsynligvis være mer utfordrende å styre diskusjonen. Det er en fordel om deltakerne i en fokusgruppe ikke kjenner hverandre på forhånd for det er oftest lettere å snakke helt fritt om sine meninger og vaner med personer en ikke skal «stå til rette overfor» privat eller på jobb.





Det er viktig å utarbeide en strukturert og tydelig intervjuguide som diskuteres av berørte parter og gjerne utprøves på forhånd. Dette vil være et viktig verktøy for styring av fokusgruppen og moderator bør bruke denne aktivt for å holde diskusjonen innenfor de definerte rammer. Det vil dessuten være viktig at en person er utpekt som observatør av fokus-gruppen og er tilstede sammen med moderator. Det vil være naturlig at denne observatøren er ansvarlig for skriving av rapport. Ved tolkning av resultater og skriving av en rapport fra en fokusgruppe er det viktig å behandle dataene kvalitativt. Man må ikke falle for fristelsen til eksempelvis å skrive at «80% av forbrukerne i gruppen prefererte produkt A framfor produkt B», men man bør forsøke å besvare definert problemstilling på en helhetlig måte. Sitater fra deltakere kan benyttes til å underbygge diskusjon og konklusjon.

Et alternativ til fokusgrupper kan være én-til-én intervju. Noen ganger gir forbrukere mer fullstendig informasjon når de kan uttale seg alene fremfor i en gruppe, og noen temaer kan det være vanskelig å snakke om i en gruppe, eksempelvis feilernæring og fedme-problematikk. Enkelte forbrukere kan dessuten lett påvirkes av andres meninger i en gruppesamtale, dette kan gjelde for eksempel barn og tenåringer. En ulempe ved én-til-én intervjuer framfor en gruppesamtale er imidlertid at man mister muligheten for gruppedynamikk. Man kan selvfølgelig henviser til utsagn fra tidligere intervjuer for å fremme videre diskusjoner, men man vil allikevel miste muligheten til direkte kommunikasjon mellom respondentene.

En tredje type kvalitativ studie er direkte observasjon av individer, dvs man studerer forbrukeres reaksjoner på og bruk av et produkt. Dette kan gjøres ved hjelp av video eller ved å delta i aktiviteter sammen med individene. Slike metoder vil være velegnet hvis man ønsker å studere eksempelvis tilberedelse av matretter, bruk av ingredienser ved tillaging av mat, åpning av pakninger og bruk av restråstoff (matsvinn). Dataene vil ha høy validitet siden forbrukere utsettes for direkte observasjon, men innsamling av slike data kan være tid- og ressurskrevende.

Det vil være hensiktsmessig å kombinere informasjon fra kvalitative og kvantitative tester i forbrukerstudier. Ved å innhente kvalitativ





informasjon fra forbrukere i målgruppen før man gjennomfører en kvantitativ spørreundersøkelse vil man få en større sikkerhet for at spørsmålene dekker alle relevante vinkler for problemstillingen, samt som stilles i et språk som forbrukere forstår. På den annen side, hvis man velger å gjennomføre en kvantitativ spørreundersøkelse med forbrukere i målgruppen i forkant av en kvalitativ oppfølgingsstudie, vil man ha muligheten til å utdype resultater fra spørreundersøkelsen gjennom fokusgrupper. Man kan da presentere resultater fra spørreundersøkelsen for fokusgruppedeltakerne for å kunne utdype mulige årsaksforklaringer.

5.3 Repertory grid og free-choice profilering

Repertory grid og free-choice profilering er to metoder som kan brukes både med trente dommere og med forbrukere. Felles for begge metoder, er at de lar dommeren/forbrukeren utvikle et eget beskrivende vokabular, noe som er nyttig til videre markedsføring og kommunikasjon på markedet. I repertory grid presenterer man triader (tre prøver samtidig) av produkter og oppgaven blir å beskrive på hvilken måte prøvene er like eller skiller seg ut fra de andre. Free-choice profilering ligner beskrivende analyse, med den forskjell at hver dommer evaluerer produktene basert på en egen-definert attributtlister. Ved hjelp av tilpasset multivariat analyse, kan man deretter håndtere synonymer og motsetninger som felles produktattributter på tvers av dommere, og oppnå en felles sensorisk profilering. Begge metodene krever instruksjoner uten tung trening. Begge metodene egner seg best med mindre forbrukergrupper, opp til ca. 30 deltakere.

5.4 Preferanse-testing

For å innhente kunnskap om hvilke produkter forbrukere liker, eventuelt ikke liker, er det vanlig å bruke preferansetesting og/eller aksept-testing. Dette avsnittet beskriver preferansetesting, mens avsnitt 5.5 vil beskrive prinsipper for aksepttesting.



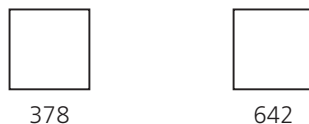


I en preferansetest presenterer man to produkter til forbrukere og spør: «hvilken prøve liker du best?». Preferansetesting innebærer at deltakeren foretar et valg mellom to produkter med hensyn på preferanse, men man får ikke informasjon om hvor godt produktene likes. Begge produkter kan likes meget godt, eller begge kan likes meget dårlig. Fordelen med denne type test er at den er enkel å forstå ettersom prinsippene kan sammenliknes med valg vi forbrukere gjør hver dag, enten det er i butikken eller foran kjøleskapet hjemme på kjøkkenet. Denne type test vil være egnet for små barn og for personer som av en eller annen grunn har problemer med å forholde seg til en skala. En preferansetest er i prinsippet en partest. I partester presentert for et trenet panel er spørsmålet hvilken prøve som har sterkeste intensitet av en egenskap, mens forbrukere i en preferansetest får spørsmålet hvilken prøve som likes best.

I en preferansetest får deltakerne utdelt 2 kodede prøver samtidig (A og B), og blir bedt om å identifisere den av de to som foretrekkes. En preferansetest har to mulige serveringsrekkefølger (AB, BA). Rekkefølgen bør balanseres slik at like mange får prøve A først og prøve B først (se partest, 4.2.2) Eksempel på skjema finnes i figur 5.1 og kapittel 16, vedlegg a. Vanligvis må respondenten foreta et valg, det er altså ikke mulig å svare «ingen preferanse». Årsaken er at mange forbrukere raskt vil kunne krysse av for «ingen preferanse» dersom dette er mulig, særlig hvis produktene er relativt like. Man ønsker å oppfordre respondentene til å velge. Dette gir et mer diskriminerende datagrunnlag, samtidig som analyse og tolkning av resultater blir enklere. Beskrivelse av hvordan håndtere data som inkluderer «ingen preferanse» finnes i Lawless & Heymann, 2010.

PREFERANSETEST

På bordet foran deg står det to prøver, smak på begge prøvene og begynn med prøven som står til venstre. Kryss av for den prøven du liker best.



Figur 5.1 Eksempel på skjema til en preferansetest





Preferansetesting med mer enn 2 produkter vil gjennomføres som en rangeringstest (se 4.2.6). Man serverer da produktene til forbrukere og ber de rangere disse etter hvor godt de likes. Hva som er maksimum antall produkter i en rangeringstest finnes det ingen regler for, dette vil være avhengig av type produkt og hvordan testen gjennomføres. Ved rangering av prøver med eplejuice, vil det sannsynligvis være mulig å inkludere 6-8 prøver ettersom dette er en relativ mild drikk. Men ved rangering av ølprøver, vil muligens 4-6 prøver være maksimum.

For preferansetest med to prøver vil sannsynligheten for å tilfeldig velge en prøve være en av to. Null hypotesen vil derfor være at når det ikke finnes noen preferanse for det ene produktet framfor det andre vil forbrukerne som gruppe velge hvert at produktene like mange ganger. Binomial fordelingen gir mulighet til å avgjøre om resultatet var tilfeldig eller om det finnes en preferanse. Rangeringsdata analyseres vanligvis med Friedman's test (se kapittel 11).

5.5 Aksept-testing

Ved gjennomføring av en aksept test, serverer man forbrukere ett og ett produkt og spør om grad av liking eller misliking. På sensorisk fagspråk kalles denne form for måling av aksept «hedonisk liking». På bakgrunn av svar fra forbrukere i en aksepttest er det mulig å konkludere hvilken prøve som prefereres av forbrukere, men det er ikke mulig å utlede grad av aksept basert på svar fra forbrukere i en preferansetest.

Den mest vanlige skala i en aksepttest er 9-punkts hedonisk skala (Peryam og Pilgrim, 1957). Den opprinnelige versjonen av skalaen ble utviklet med beskrivende engelske ord for hver av de 9 enhetene på skalaen og uttesting ble foretatt for å sikre at avstanden mellom hver enhet oppfattes som lik, og derved danner et grunnlag for normalfordelte data. Det er imidlertid vanskelig å finne tilsvarende beskrivende ord i det norske språket og en mye benyttet variant av 9-punkts akseptskala er vist i figur 5.2. Figuren viser at skalaen forankres med «liker ikke i det hele tatt» og «liker meget godt», og noen ganger inkluderer man et midtpunkt med «verken liker eller ikke liker».



Kode:

Hvor godt liker du spekeskinken?



Figur 5.2: 9-punkts akseptskala

Skalaen for aksepttesting kan også være basert på 5 punkter eller 7 punkter. Det anbefales imidlertid å benytte 7- eller 9-punkts skala for å oppnå en best mulig differensiering mellom produkter. Årsaken er at mange forbrukere har en tendens til å unngå å bruke ytterpunktene av skalaen, noe som medfører at det vil være lite «spillerom» ved bruk av 5 punkts skala. Som nevnt ovenfor (figur 5.2) er det også vanlig å inkludere et midtpunkt på skalaen som kan beskrives som «verken liker eller ikke liker», men det finnes argumenter for at et nøytralt midtpunkt gjør skalaen mindre diskriminerende. Hedoniske skalaer kan være kontinuerlig (uten angivelse av tall eller bokser) eller kategorisk (med tall eller bokser) (se 4.1).

Prosedyren for preferansetesting og aksepttesting er lik, bortsett fra preferanse-/aksept-spørsmålet og inkluderer følgende trinn:

- Bestemme hensikten med test, prøver, tidspunkt, kriterier for rekruttering av forbrukere etc.
- Bestemme testbetingelser (prøvestørrelse, temperatur osv)
- Lage spørreskjema
- Rekruttere forbrukere iht definerte kriterier (bruk av produktet, alder, osv).
- Sette opp balansert serveringsrekkefølge
- Definere koder og merke alle prøver
- Gjennomføre testen
- Analysere resultater
- Rapportere resultater



Selv om man oppnår kunnskap i form av en tallverdi om grad av hedonisk liking fra en aksepttest er det viktig å være oppmerksom på at bedømmelsen som gis er relativ. Bedømmelsen som gis av forbrukere vil være avhengig av det utvalg av prøver som inngår i testen. For eksempel hvis man serverer 4 prøver i en aksept test hvor alle prøvene er middels likt av forbrukere vil scorene sannsynligvis spre seg rundt midten på skalaen. Hvis man inkluderer en 5. prøve som er veldig dårlig likt kan dette trekke bedømmelsen av de middels prøvene i positiv retning. Hvis man derimot inkluderer en 5. prøve som er veldig godt likt, vil dette kunne trekke bedømmelsen av middels prøver i negativ retning.

Rekkefølgen for servering av prøver vil også ha betydning. Det er alltid viktig å servere prøvene i balansert rekkefølge (Mac Fie et al, 1989) og det anbefales å inkludere en «dummy-prøve» dersom dette er mulig. En balansert rekkefølge sikrer at alle prøvene i testen er servert like mange ganger som nr. 1 som som nr. 2 osv, og at alle prøver presenteres like mange ganger før som etter hver av de andre prøvene. På den måten blir smakspåvirkningen av en prøve på den neste annullert over datamaterialet fra samtlige deltakere. En dummy-prøve bør representere et sensorisk midtpunkt av alle påfølgende prøver. Når denne serveres først vil den kunne hjelpe forbrukere til å orientere seg på skalaen og skape passende sensoriske forventninger til de kommende prøver. Scoren for dummyprøven benyttes ikke i videre analyser.

Ved aksept testing av produkter for barn er det vanlig å bruke såkalte «smiley» skalaer. Men det er også vist at skaler med verbal beskrivelse kan fungere godt for barn helt ned til 5 år.

Analysert i sammenheng med beskrivende analysedata, er akseptdata brukt i preference mapping for å kartlegge hvilke produkttegenskaper påvirker liking og i hvilken retning.

5.6 Nyere sensoriske metoder

Det har de siste årene blitt økt etterspørsel etter kvantitative metoder som gir beskrivende sensorisk informasjon om produkter direkte fra

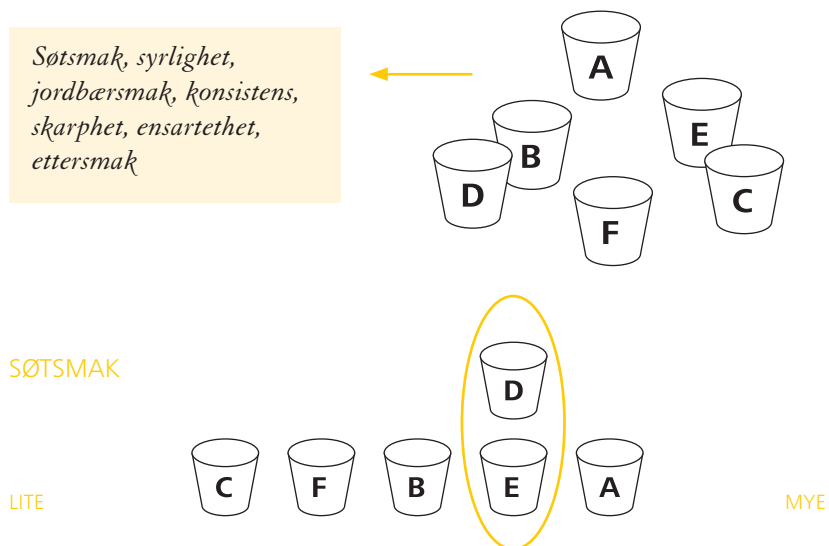




forbrukere. Motivasjonen for slike metoder er at ved å bruke et større antall utrente forbrukere som allerede er rekruttert til aksepttesting sparer man tid og penger i forhold til å trene et sensorisk panel. I dette avsnittet vil vi kort gå igjennom fire ulike metoder som alle gir sensorisk informasjon om produkter og som er relativt brukervennlige. Metodene som beskrives er «flash profiling», «sorting», «projective mapping» og “check all that apply” (CATA). Navnene angis på engelsk ettersom dette er vanlig i fagspråket.

Flash profiling

Flash profiling er en metode for sensorisk beskrivelse som er en variant av free choice profiling (Dairou and Sieffermann, 2002). Metoden gjennomføres i to sesjoner, eller i en sesjon med to trinn. I første trinn vil forbrukere bedømme alle prøvene og angi alle egenskaper som de mener er egnet for å diskriminere mellom produktene. Dette betyr at hver respondent kan ha ulikt antall egenskaper. I neste trinn blir de bedt om



Figur 5.3: Flash profiling av jordbæryoghurt

Illustrasjon: Julie Gjermundsen



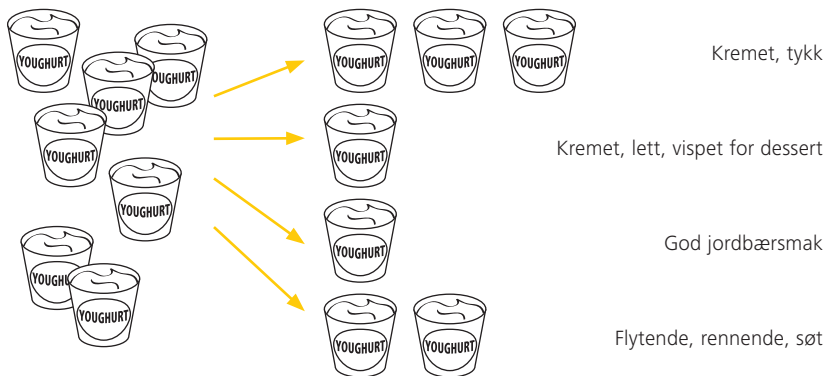


å rangere alle prøvene fra lav til høy for hver av de valgte egenskaper. Flash profilering kan gjennomføres med både trente dommere og forbrukere og prinsippene er vist i figur 5.3. Beskrivelse av data-analyse og applikasjoner for Flash profilering finnes i Varela & Ares, 2012.

Sorting

Sorting (sortering) av produkter har tidligere blitt brukt for å frembringe informasjon om sensoriske beskrivelse av produkter (Lawless et al, 1995). Metoden baseres på klassifikasjon eller en gruppering av produkter i kategorier basert på egendefinerte kriterier, og dette er en metode som benyttes bla innen psykologi og sosiologi.


Ved sensorisk sortering blir respondenter bedt om å måle grad av likhet og ulikhet mellom prøver ved bruk av deres egne personlige kriterier. Dernest skal de prøver som oppfattes som like plasseres i den samme gruppe, og de prøver som oppfattes som forskjellige skal plasseres i forskjellige grupper. Det er vanligvis et krav under sorteringen å etablere minst to grupper, og det er ikke tillatt å ha bare en prøve i hver gruppe. Respondentene blir deretter bedt om å oppgi beskrivende ord for hver gruppe av produkter. Sortering av produkter kan gjennomføres med både trente dommere og forbrukere og prinsippene er vist



Figur 5.4: Eksempel på sortering av jordbæryoghurt

Illustrasjon: Julie Gjermundsen





i figur 5.4. Beskrivelse av data-analyse og applikasjoner for sortering finnes i Varela & Ares, 2012.

Projective mapping

Projective mapping (napping) stammer fra psykologi, og er tidligere gjennomført i sensorisk forskning med forbrukere for å etablere kunnskap om sensoriske beskrivelse av produkter (Risvik et al, 1994). I projective mapping presenteres prøvene samtidig og hver respondent blir bedt om å plassere de i et to-dimensjonalt rom, vanligvis et A2 ark (60 x 40 cm). Respondentene blir bedt om å vurdere prøvene (se, lukte, smake, føle på) og å plassere de i henhold til forskjeller og likheter på en slik måte at jo mer like de er, jo nærmere skal de plasseres på arket og jo mer forskjellig de er, jo lengre fra hverandre. Kriteria for plassering velges individuelt, noe som gjør projective mapping til en fleksibel og spontan metode. Etter å ha plassert prøvene blir deltakerne ofte spurt om å skrive ned kommentarer for å beskrive prøver eller grupper av prøver, dvs en begrunnelse for prøvenes plassering (ultra flash profiling).

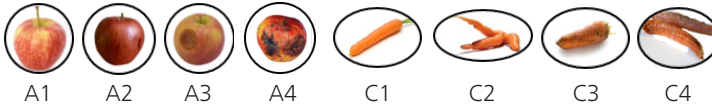
Projective mapping kan gjennomføres både med kodete prøver (blindt) og med merkede prøver og respondenter kan både være trente dommere eller forbrukere. Når det gjennomføres blindt kan det være nyttig å legge til et prøve-replikat for å sjekke om respondentene er konsistente. Data fra projective mapping analyseres med multivariat analyse og et eksempel på konsensusplot gjennomført med prøver med epler og gulrøtter er vist i figur 5.5. Mer informasjon om analyser, samt eksempler på bruk finnes i Varela & Ares, 2012.

Check all that apply

Check all that apply (CATA) er betegnelsen på multiple valg spørsmål og er mye benyttet i markedsanalyser for å gjøre oppgaven enkel for respondenter. Spørsmålene består av en liste med ord eller påstander som respondenten kan velge om passer til å beskrive et tema eller en prøve. I sensoriske tester benyttes CATA som en enkel metode for å innhente informasjon om forbrukeres sensoriske opplevelse av produkter. Produktene presenteres enkeltvis til forbrukerne ifølge en balansert

PROJETIVE MAPPING EKSEMPEL

1. Deltakere mottar 8 prøver og et A2 ark.

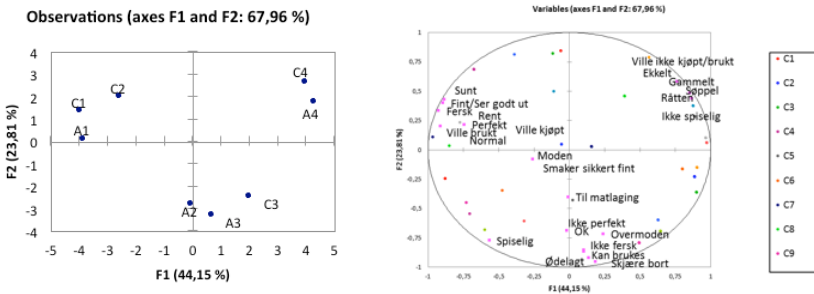


2. Hver deltaker lager et kart ved å plassere prøvene på arket i forhold til likheter og ulikheter. Deltakerne skriver en kort beskrivelse for hver prøve eller hver gruppe.



Kart fra en deltaker

3. Koordinater for prøvene og attributtene samles i en datatabell. Dataene analyseres med Multiple Factor Analysis (MFA)



Figur. 5.5: Konsensus produkt-kart (venstre) og attributt-kart (høyre) basert på multivariat analyse (MVA) av individuelle kart.



rekkefølge. Respondentene blir så spurt om å vurdere produktet og å krysse av for de egenskaper som de synes passer til å beskrive prøven. Det er vanligvis ikke noen begrensning for hvor mange egenskaper som kan velges av respondentene. Listen som presenteres kan inneholde kun sensoriske produktegenskaper, eller den kan også inkludere ikke-sensoriske egenskaper som for eksempel bruksområder, merkevare etc.

Kode 274:

Vennligst spis opp den siste biten samtidig som du vurderer hvilke egenskaper du synes beskriver brødet.

Kryss av for alle egenskapene du mener gjelder

- | | | |
|---|--|---|
| <input type="checkbox"/> Sur smak | <input type="checkbox"/> Tungt | <input type="checkbox"/> God smak |
| <input type="checkbox"/> Myk | <input type="checkbox"/> Søt smak | <input type="checkbox"/> Veldig god |
| <input type="checkbox"/> Smak av surdeig | <input type="checkbox"/> Smak av korn | <input type="checkbox"/> Dårlig smak |
| <input type="checkbox"/> Saftig | <input type="checkbox"/> Klebrig | <input type="checkbox"/> Mellomgrovt |
| <input type="checkbox"/> Kompakt | <input type="checkbox"/> Luftig | <input type="checkbox"/> Bitter smak |
| <input type="checkbox"/> Porøst | <input type="checkbox"/> Tørr | <input type="checkbox"/> Lite grovt |
| <input type="checkbox"/> Mye tyggemotstand | <input type="checkbox"/> Gjærsmak | <input type="checkbox"/> Hardt |
| <input type="checkbox"/> Deigete | <input type="checkbox"/> Smulete | |
| <input type="checkbox"/> Passer til frokost | <input type="checkbox"/> Dette ville jeg kjøpt | <input type="checkbox"/> Brød til helgen |
| <input type="checkbox"/> Fiberrikt | <input type="checkbox"/> Usunt | <input type="checkbox"/> Tiltalende |
| <input type="checkbox"/> Hverdagsbrød | <input type="checkbox"/> Passer til matpakke | <input type="checkbox"/> Mettende |
| <input type="checkbox"/> Dette ville jeg ikke kjøpt | <input type="checkbox"/> Passer til kveldsmat | <input type="checkbox"/> Passer til middag (f.eks. suppe) |
| <input type="checkbox"/> Ikke tiltalende | <input type="checkbox"/> Passer til lunsj | <input type="checkbox"/> Sunt/mye næring |

Figur 5.6: Eksempel på CATA spørreskjema.

130 Sensorikk – Måling med menneskelige sanser





Valg av egenskaper i CATA er en hovedutfordring med metoden og krever et godt forarbeid. Det er vanlig å benytte ord fra trente sensoriske dommere, ord fra produktutviklere eller ord fra fokusgrupper. CATA er mest vanlig å gjennomføre med forbrukere. Dataanalysen for CATA gjennomføres ved hjelp av multivariat analyse. For ytterligere informasjon om data-analyse, samt eksempler på applikasjoner henvises til Varela & Ares, 2012. Eksempel på et CATA spørreskjema finnes i figur 5.6. Metoden refereres på norsk som KASPER (Kryss Alt Som PassER)(se kapittel 6).

5.7 Valg av prøver ved conjoint design

I forbrukertester er man ofte interessert i å finne ut hvilke sensoriske og ikke-sensoriske egenskaper som styrer forbrukernes vurdering av prøvene. En effektiv tilnærming til problemstillingen er å bruke conjoint analyse. Conjoint metoden består i å teste ut prøver som varierer med kontrollerte kriterier og nivåer i henhold til et statistisk forsøksdesign. Prøvene kan enten variere i eksterne parametere og produktinformasjon alene (f. eks. pris, emballasjedesign, nøkkelhull...) eller i kombinasjon av både sensoriske og eksterne faktorer (f. eks. saltsmak, tekstur, utseende, og informasjon om saltinnhold, opprinnelse eller prosess). Typisk ber man deltakerne vurdere prøver på en aksept skala, rangere prøver, eller velge ut én av flere prøver i en rekke valgsett, men tilnærmingen til prøveutvalget kan også kombineres med sortering, projective mapping og andre nyere responsmetoder.

5.8 Betydning av testsituasjon (kontekst)

Spise-situasjonen eller konteksten inkluderer tid, sted, på hvilken måte og sammen med hvem et måltid inntas og alle disse faktorene kan ha stor betydning for matopplevelsen. Meiselman (1996) diskuterer følgende kontekstuelle faktorer: Maten, situasjonen og individet. Maten i seg selv er en kontekstuell faktor ettersom den kan serveres alene, i en matrett med flere matvarer eller som en del av et helt måltid. Situasjonen er en kontekstuell faktor som både består av fysiske og sosiale omgivelser. Et individ eller en forbruker, bringer med seg mange





erfaringer, holdninger og forventninger til en spise-situasjon, noe som også kan påvirke spiseopplevelsen betraktelig.

I forbindelse med preferansetesting og aksepttesting er det viktig å tenke igjennom mulig påvirkning av kontekstuelle faktorer og utforme testen slik at resultatene blir mest mulig gyldige i forhold til den definerte problemstilling i testen og i forhold til fremtidige spisesituasjoner for det aktuelle produkt. Produktet bør tilberedes mest mulig likt slik som det vanligvis spises. Dette vil si at ved gjennomføring av for eksempel en forbrukertest på leverpostei bør det vurderes å servere brød til, mens ved en forbrukertest av ketchup bør det vurderes å servere denne som en del av en egnet rett. Andre faktorer som bør vurderes er temperatur på prøven og mengde prøve som passer til en naturlig spiseopplevelse.

Når det gjelder testsituasjon er det mest vanlig å innkalle forbrukere til en test i et laboratorium, i en kantine, til et kjøpesenter eller å sende produkter med forbrukere hjem. I forhold til disse nevnte testsituasjoner er graden av kontroll med testbetingelser avtakene og grad av naturlighet for spisesituasjonen økende. De sosiale omgivelser vil også være forskjellige for de nevnte testsituasjonene. I en sensorisk testbås sitter man vanligvis alene i fred og ro. I en kantine eller et kjøpesenter vil det sannsynligvis være muligheter for sosial interaksjon med andre personer og mer støy. Ved en hjemme-test vil man potensielt være sammen med venner og familie. Man vil derfor sannsynligvis være mer sensitiv for forskjeller mellom prøver i en sensorisk testbås enn hjemme ved kjøkkenbordet, og optimal testsituasjon vil derfor avhenge av formål ved testen.

Variasjon i kontekstuelle faktorer knyttet til individet (forbrukerne) kan delvis reduseres gjennom kriterier for rekruttering og deltakelse. Hvis man etablerer kriterier for deltakelse knyttet til kjønn, alder, bruksfrekvens, matvaner o.a. vil man sikre en mindre variasjon hos forbrukere, og produkter som velges på bakgrunn av resultater fra testen vil være mer skreddersydd til en slik målgruppe. Men man må også da være oppmerksom på at de muligens er mindre skreddersydd til målgrupper som avviker fra disse kriteriene. Det finnes flere studier som påviser effekter av kontekstuelle faktorer i forbrukertester. Resultater





fra studiene viser at slike effekter kan være produktavhengige. For eksempel vises det i Hersleth et al (2003) at aksept for Chardonnay-vin målt i en forbrukertest ble påvirket både av test-situasjon og tilbehør, mens en tilsvarende studie av norske oster viste en stabil aksept over ulike test-situasjoner (Hersleth et al, 2005).

Selv om det i de fleste tilfeller ikke vil være mulig å gjenskape en virkelig spisesituasjon i en forbrukertest, kan mye legges til rette ved å gi gode instruksjoner til forbrukere under testen. Det kan være viktig å be forbrukere tenke seg et spesielt måltid eller en spesiell spisesituasjon, det kan være mulig å be de vektlegge spesielle egenskaper ved maten og det kan være mulig å vise forbrukere bilder av måltider. Et eksempel på en spesifikk kontekst er to ulike spisesituasjoner for spekeskinke (tradisjonelt spekematmåltid og tapas-måltid) som ble introdusert til forbrukere v.h.a bilder og tekst. Viktigheten av å tenke igjennom og å teste ut type informasjon som gis til forbrukere og eventuelt type informasjon som ikke gis til forbrukere i en forbrukertest vil være en god investering for å få gyldige resultater.

5.9 Nettesting

Det er i det siste tiåret blitt veldig vanlig å gjennomføre forbrukertester og forbrukerundersøkelser via internett, i takt med utviklingen av tilsvarende tekniske løsninger. En fordel med nettesting i et stort land som Norge er muligheten til å inkludere deltakere fra forskjellige geografiske områder, noe som ofte ikke er mulig i lab-tester. En ulempe er at en eventuell produktsmaking er utfordrende å organisere og må kombineres med utsendelser av produkter. Nettesting brukes derfor oftest i studier fokusert på holdninger og vaner, samt i produkttester uten produktsmaking der deltakerne vurderer produkter på basis av produktbilder og/eller angitt produktinformasjon. Conjoint er en velegnet metode i nettesting (Øvrum et al, 2012). Også sortering, projective mapping og andre nyere metoder er tilgjengelige i nettbaserte løsninger. Nettester er forøvrig en fleksibel løsning for deltakerne som kan velge hvor og når de ønsker å besvare testen, dette være seg på jobben i lunsjpausen, på bussen om ettermiddagen eller fra sofaen i





helgen. Dagens teknologi tillater å samle besvarelser via både datamaskiner, smarttelefoner og nettbrett. Det kan imidlertid være stor variasjon i hvorvidt deltakeren er fokusert på produktvurderingene under besvarelsen. Det anbefales å inkludere noen «lurespørsmål» underveis i testen, utformet for å sjekke om deltakeren leser spørsmålene ordentlig, slik at en kan se bort ifra upålitelige besvarelser.

Referanser

Dairou, V. and J.-M. Sieffermann. 2002. A comparison of 14 jams characterized by conventional profile and a quick original method, the Flash profile. *Journal of Food Science* 67 (2), 826-834

Hersleth, M., Monteleone, E., Segtnan, A. and Næs, T. 2015. Effects of evoked context on consumers' responses to intrinsic and extrinsic product attributes in dry cured ham. *Food Quality and Preference*, 40, 191-198.

Hersleth, M., Illseng, M.A., Martens, M., Næs, T. 2005. Perception of cheese - a comparison of quality ratings by experts, descriptive profiling by a trained panel and consumer response. *Journal of Food Quality*, Vol 28, pp 333-349.

Hersleth, M., Mevik, B.-H., Næs, T., Guinard, J.-X. 2003. Effect of contextual factors on liking for wine - use of robust design methodology. *Food Quality and Preference*, Vol 14, 7, pp 615-622.

Lawless, H.T., Sheng, N., and Knoops, S.S.C.P. 1995. Multidimensional scaling of sorting data applied to cheese perception. *Food Quality and Preference* 6: 91-98.

Lawless, H., & Heymann, H. (2010). *Sensory evaluation of food, Principles and Practices*, New York: Springer.

MacFie, H. J., Bratchell, N., Greenhoff, K., & Vallis, L. V. (1989). Designs to balance the effect of order of presentation and first-order carry-over effects in hall tests. *Journal of Sensory Studies*, 4, 129-148.





Meiselman, H. L. (1996). The contextual basis for food acceptance, food choice and food intake: the food, the situation and the individual. In H. L. Meiselman & H. J. H. MacFie (Eds.), *Food choice acceptance and consumption* (pp. 139-263). Glasgow: Blackie Academic and Professional

Relevant litteratur

Almli, V. L., Næs, T., Sulmont-Rossé, C., Enderli, G., Issanchou, S., & Hersleth, M. (2011). Consumers' acceptance of innovations in traditional cheese. A comparative study in France and Norway. *Appetite*, 57, 110-120.

Almli, V.L. & Hersleth, M. (2012). Salt replacement and injection salting in smoked salmon evaluated from descriptive and hedonic sensory perspectives. *Aquaculture International*, (DOI) 10.1007/s10499-012-9615-4.

Helgesen, H., Solheim, R., Næs, T (1997). Consumer preference mapping of dry fermented lam sausages. *Food Quality and Preference* 8 (2), 97-109.

Kroll, B. J. 1990. Evaluation rating scales for sensory testing with children. *Food Technology*, 44 (11), 78-80, 82, 84, 86.

Peryam, D:R.; Pilgrim, F:J: (1957). Hedonic scale method of measuring food preferences. *Food Technology* 11, 9-14





Foto: Mari Svenningsen





PRODUKTUTVIKLING

Urd Bente Andersen, Tone Eikrem Nyvold og Yvonne Mainusch,
Orkla Foods Norge

6.1 Produktutvikling

Produktutvikling er viktig for å skape og vedlikeholde en merkevare. En produsent ønsker å utvikle nye produkter som er innovative og/eller bedre enn konkurrentenes. Som produktutvikler må man kunne utvikle produkter med egenskaper (smak, utseende, osv.) som forbrukeren liker. I tillegg må produktutvikleren være i stand til å måle disse egenskapene under utviklingsprosessen. Dette kan gjøres ved hjelp av sensoriske analyser som beskrives nærmere i dette kapittelet.

Før utviklingsarbeidet starter, blir det gjennomført markedsundersøkelser for å kartlegge hva forbrukeren vil ha, samt for å undersøke om det er et reelt behov for et produkt i markedet. I større bedrifter er det markedsavdelingen som gjennomfører markedsundersøkelser. I mindre bedrifter må ofte ledelsen eller produktutviklere selv undersøke markedet for å finne ut hva forbrukerne vil ha. Mer om markedsundersøkelser i kapittel 5.

Når markedsbehovet er identifisert, kan produktutviklingsarbeidet starte. For en produktutvikler er det først og fremst viktig å forstå markedsidéen/-konseptet. Det er også viktig å vite hvem målgruppen er. Skal produktet for eksempel være til hele familien eller til barna?

Videre gjelder det å definere noen funksjonelle/tekniske og sensoriske egenskaper som produktet må ha. Skal det være et lavkalori-produkt? Skal produktet ha spesielt lang holdbarhet? Skal det oppbevares romtemperert eller kjølt? Hvilke egenskaper ved utseende, konsistens og lukt/smak er viktige i produktet man ønsker å lansere?





Visste du at?

Aroma er en vanlig ingrediens i mange næringsmidler og tilsettes for å få frem smak og lukt. Aromaer blir produsert av bedrifter som er spesialisert innen feltet, såkalte aromahus. Aromahusene bruker mye sensorikk i utviklingen av aromaer. Dette fordi aromaer består av veldig mange ulike aromastoffer, som er komplekse å analysere. For eksempel består en standard jordbæraroma av omtrent 100 ulike aromastoffer, og en kaffearoma kan gjerne bestå av mer enn 1000 ulike aromastoffer.

For å kunne utvikle aromaer benyttes derfor spesielt trente personer, kalt flavourister. Flavouristene jobber med sitt sensoriske apparat, spesielt luktesansen, for å gjenkjenne aromastoffer.

I tillegg beskriver flavouristene aromaer med egne begreper. En bringebæraroma kan for eksempel beskrives som «bærsmak», «grønn», «parfymert», «frøsmak», «såpe», osv.

En flavourist kan jobbe med 5000 forskjellige råmaterialer (aromastoffer og løsemidler) og har kunnskap om forholdet mellom disse i typiske aromaer, som for eksempel jordbær, bringebær og kaffe.

I den første fasen av aromautvikling spiller orthonasal lukt en viktig rolle. Orthonasal lukt betyr at aromastoffene flyter gjennom nesen direkte til luktesansen (se figur 2.7 s. 40). Ved å lukte på ulike aromastoffer setter flavouristen sammen en aromaresept. Når man har kommet frem til en aromaresept, er det viktig å undersøke aromaens funksjonalitet i sluttproduktet. Her lukter man retronasalt, som betyr lukt via munnen. Aromastoffene settes fri ved å tygge på produktet og så flyter stoffene opp til luktesansen hvor de interagerer med luktreseptorer.

En aroma kan gi ulike smaksopplevelser i forskjellige matvarer. Dette kan skje fordi produktsammensetningen påvirker aromaen. I melk, for eksempel, maskerer fettmolekylene ofte aromastoffene slik at profilen blir en annen enn når aromaen tilsettes et matprodukt uten fett. I tillegg kan aromaer forandres under produksjonsprosessen av matvaren. Når det lages brus med jordbærsmak, er det viktig at jordbæraromaen ikke blir ødelagt av varmen under pasteurisering.

Aromafunksjonaliteten testes derfor i sluttproduktet, altså når produktet har gjennomgått hele produksjonsprosessen. Dette gjøres av flavourister i samarbeid med sensoriske paneller. Det er vanlig å bruke testpaneller som er spesialisert innenfor ett aromaområde, for eksempel fruktige aromaer, citrusaromaer eller kremaktige aromaer. Avhengig av hva man ønsker undersøkt, brukes forskjellige tester som for eksempel triangeltest og partest, eller beskrivende tester som for eksempel profileringstest.

– Yvonne Mainusch





Gjennom utviklingsprosessen jobber man for å oppnå de sensoriske egenskapene ved produktet som ble fastsatt i idéfasen. Ved hjelp av sensorikk er det mulig å analysere viktige sensoriske egenskaper under hele utviklingsprosessen. Ofte brukes begrepet «smak» som en fellesbetegnelse på disse egenskapene. Dersom smaken blir endret underveis i prosessen, vil sensoriske analyser kunne dokumentere og gi et mål på hvor stor endringen er. Sensorikk kan derfor brukes til å avdekke om smaken er riktig i forhold til det som opprinnelig ble bestemt. I tillegg til de sensoriske egenskapene vil ytre egenskaper, som har med emballasje og kjøpsituasjonen å gjøre, spille inn på forbrukerens valg. Faglitteraturen bruker betegnelsen *intrinsic factors* (indre faktorer) om de sensoriske egenskapene knyttet til selve produktet. Alle andre ytre faktorer blir betegnet som *extrinsic factors*.

6.2 Prosjektarbeid og prosjektgruppen

For å lykkes med et nytt produkt er det viktig å avsette tid til arbeidet, ha en finansiell plan for prosjektet og et godt samarbeid mellom ulike prosjektmedlemmer. Det må opprettes en prosjektgruppe med en prosjektleder som styrer arbeidet gjennom hele prosessen. Prosjektgruppen bør bestå av teknisk personell (for eksempel ingeniører og kokker) som kjenner produktene og prosessene, samt markedsførere som har innsikt i markedssituasjonen.

Prosjektmedlemmene må forstå hverandre og ha et felles «språk». For eksempel kan en sensorisk beskrivelse av et produkt brukes til å kommunisere hva man vil frem til. De som jobber med marked og emballasje (produksjefer, designere, emballasjeingeniører) må forstå de sensoriske begrepene og hva de innebærer. På den annen side må produktutviklerne og sensorikerne skjønne betydningen av design, pris og emballasjestørrelse.

Det er viktig at de interne sensoriske analysene ses i lys av eksterne tester både av egne og andre produkter i markedet. Prosjektgruppen må derfor følge med på eksterne undersøkelser og trender i markedet og relatere disse til produktene man jobber med.





6.3 Produktutviklingsprosessen (PU-prosessen)

Noen av de større næringsmiddelbedriftene som har en stor produktutviklingsavdeling, bruker Cooper's Stage-Gate modell for utviklingsprosessen (se figur 6.1). Robert G. Cooper forklarer denne modellen i boken «Winning at New Products».

Modellen går ut på å dele PU-prosessen inn i faser. Etter hver fase tas det en beslutning om prosessen er klar for å fortsette, om det er nødvendig å gjøre forbedringer eller utvikling av produktet skal stoppes. Figur 6.2 viser en oversikt ulike faser, med spesielt fokus på de sensoriske aspektene. Omfanget av analyser vil variere avhengig av type produkt og hvor i prosessen man er.

Fase 1: Definisjon av produktegenskaper

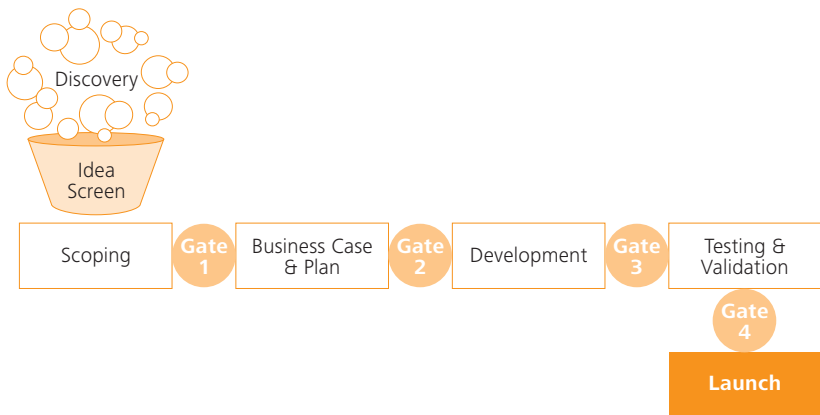
I fase 1 av produktutviklingsprosessen handler det om å skape idéer til et nytt produkt.

I større bedrifter er det ofte en egen avdeling (marketing) som har hovedansvaret for å studere markedssituasjonen og trender. Markedssituasjonen analyseres ved hjelp av markedsdatabaser som for eksempel Mintel og Nielsen. Ved å intervjuere forbrukere eller gjennomføre online forbrukerundersøkelser får man et bilde av hva forbrukeren ønsker. Når behovet er definert, jobbes det med å skape idéer.

Markedsansvarlige og produktutviklere jobber da tett sammen for å definere produkt-egenskapene og kritiske suksessfaktorer.

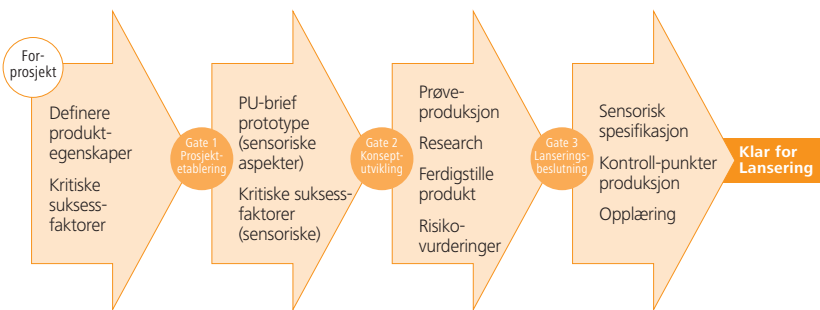
For å få inspirasjon reiser man gjerne til store byer i utlandet for å besøke supermarkeder eller messer. Vel hjemme sammenlignes aktuelle produkter fra utlandet med egne produkter og/eller tilsvarende produkter fra hjemmemarkedet. Emballasje, design, merking og deklarerings studeres og man smaker seg gjennom produktene. Som regel er både markedsførerne og produktutviklerne til stede under denne typen smakinger. I tillegg stiller ofte ledelsen i bedriften opp. Det er en fin måte for ledelsen å få en gjennomgang og oversikt over markedet og se muligheter for egne produktkategorier. Under en slik seanse er også personer med sensorisk kompetanse med for å strukturere smakingen. Det er viktig å få kartlagt hvilke sensoriske parametre som karakteriserer de aktuelle produktene. Disse danner grunnlag for hvilke egenskaper man vil at det nye produktet skal ha.





Figur 6.1 Skisse av Cooper's Stage-Gate modell

Illustrasjon: Julie Gjermundsen



Figur 6.2 Eksempel på ulike faser og beslutningspunkter (gates) i en produktutviklingsprosess

Illustrasjon: Julie Gjermundsen





Fase 2: Utvikling av prototypeprodukt.

Basert på informasjonen fra fase 1 begynner kokker og produktutviklere arbeidet med å lage en prototype. I denne fasen er det mye prøving og feiling for å finne gode prototyper. Produktutviklere i matvareindustrien vil ofte starte arbeidet med å kokkelere og lage rettene fra bunnen av med de beste råvarene. De ulike variantene av f.eks. ferdig pizza eller tørre supper og sauser starter med utgangspunkt i tradisjonelle oppskrifter og ferske råvarer. Etter hvert må produktet tilpasses industrien og prosessene man har til rådighet.

Produktutviklerne etterligner utseende, smak og konsistens i det ferske produktet ved å bruke råvarer som egner seg for et industrielt fremstilt produkt. Dette er ofte råvarer som bidrar til den ønskede holdbarheten til produktet, som er innenfor kostnadsbegrensninger og som er stabile i fremstillingsprosessen på fabrikken.

Under hele fasen spiller sensorikk en viktig rolle for å bekrefte at smaksbildet til prototypen er i henhold til produkttegenskapene man har definert i den første fasen.

I denne fasen må man også bli enige om hvilke sensoriske egenskapene som er viktige for at produktet skal bli en suksess – de kritiske suksessfaktorene. Dette kan beskrives i en såkalt produktutviklingsbrief (PU-brief) som beskriver hva ønsker man at PU skal utvikle.

Fase 3: Prøveproduksjoner og ferdigstillelse av produkt

I fase 3 handler det om å oppskalere prototypen. Produktet som frem til nå har blitt laget med laboratorie- eller kjøkkenutstyr, må nå lages under reelle produksjonsbetingelser.

Storskala produksjonsutstyr kan ha en påvirkning på produktenes egenskaper.

For eksempel kan sterk varmebehandling, som er nødvendig for å sikre mikrobiologisk stabilitet, endre smaken til et produkt betraktelig. I denne fasen er det derfor viktig å kontrollere de sensoriske egenskapene til produkt fra storskalaproduksjon opp mot produkt fra laboratorieskalaproduksjon.

Når én eller flere prototyper er ferdigutviklet, gjennomføres det gjerne forbrukertester for å se om produktet treffer med tanke på smak, altså de sensoriske egenskapene. Dette gjøres ved hjelp av liking- og aksepttester (se kapittel 5).

Testene gir svar på om et produkt er klart for lansering eller om det må jobbes videre med justeringer, før det kan lanseres. I verste fall kan dette ende med at produktet ikke kvalifiserer til lansering. Uansett vil produktidéen kunne hentes opp igjen ved en senere anledning.





Fase 4: Sensorisk kvalitetskontroll

I den siste fasen av utviklingsprosessen er det viktig å dokumentere de kritiske sensoriske suksessfaktorene i det ferdige produktet.

Disse beskrives i form av en ferdigvarespesifikasjon som er en del av kvalitetssystemet på fabrikken. I ferdigvarespesifikasjonen står det blant annet hvilke sensoriske egenskaper produktet skal ha og hvordan disse skal kontrolleres. Operatørene på fabrikken får opplæring i hvordan de skal kontrollere sensoriske egenskaper på det ferdige produktet. De må vite hvordan produktet skal være etter produksjonsprosessen og kunne ta ut stikkprøver som tilberedes til sensoriske analyser. Ofte brukes standarder/referanseprøver, som for eksempel tilsvarende produkt fra en tidligere produksjon eller en prototype fra produktutviklingsprosessen (lab-standard), til å sammenligne sensoriske egenskaper.

Først når stikkprøven er sensorisk godkjent, kan produktet sendes ut på markedet. Se også kapittel 8.

6.4 Sensorikk i produktutvikling

Produktutvikling er et nitidig arbeid der mye skal klaffe og sensorikken er et viktig redskap for å måle om man er på riktig vei i forhold til det som man bestemte seg for i utgangspunktet. Å sette opp en riktig sensorisk analyse krever at man vet hva formålet med analysen er og hva man ønsker å få svar på. Følgende tre problemstillinger er aktuelle i forbindelse med produktutvikling:

- Benchmark – å måle et produkts kvalitet i forhold til tilsvarende konkurrerende produkter i markedet.
- Like bra eller bedre enn dagens produkt (standard)
- Hva er godt nok – hvordan skal det nye produktet smake

Ad. pkt 1. Benchmark vil si produktet man måler mot (man kan også utføre en benchmark). Det kan enten være et produkt av samme type som det man skal undersøke (f.eks. Knorr suppe mot Toro suppe, Lerum syltetøy mot Nora syltetøy) – eller hvis ikke det finnes et tilsvarende produkt i handelen, kan man måle mot en annen type produkt i markedet som produktet kan erstatte, konkurrere mot eller supplere.





Ad. pkt 2. Ofte utfører produktutvikleren sensoriske analyser for å undersøke om nytt produkt er like bra eller bedre enn dagens standard, dvs. om det har en like god eller bedre liking hos forbruker. Det gir først og fremst svar på om sensoriske parametere som utseende, konsistens og smak er blitt bedre i det nye produktet. Men i tillegg til smaken kan man også spørre om andre faktorer som kan påvirke om produktet blir kjøpt eller ikke (eks. emballasje, form, mengde produkt eller lignende).

Ad. pkt 3. Hvordan skal det nye produktet smake – her har man fokus på objektiv måling av de sensoriske egenskapene. Når det ikke er noe produkt å måle mot, kan prosjektgruppen selv ha definert hvilket produkt de ønsker å utvikle og hvilken kvalitet dette skal ha. Det kan være ut fra en prototype eller en benchmark mot et produkt i en helt annen kategori som de ønsker å ta markedsandeler fra. Eksempelvis kan et søtpålegg sammenlignes mot en smak i et godteri eller en ny smak i et drikkeprodukt kan sammenlignes mot en naturlig juice av frukt eller bær.

6.5 Metoder og panel

I produktutvikling benyttes både objektive/analytiske og subjektive/affektive sensoriske metoder. Bedriftene har som regel testet og trent opp egne ansatte å utføre objektive bedømmelser og mange har også ansatte med spesielt ansvar for sensorikk.

Alle spørsmål knyttet til liking (subjektive bedømmelser) må forbrukeren svare på i forbrukerundersøkelser (se kapittel 5). Som regel er det er en oppgave for markedsavdelingen, men mye av de subjektive bedømmelsene foregår også internt i bedriften. Det kan være hjemme-tester der de ansatte får med seg produkter hjem, eller at ansatte blir invitert til smakinger enten i kantinen eller til et sensorisk testrom. Når man utfører denne type interne tester må man være klar over at utvalget av personer som deltar er begrenset og ikke representative for befolkningen som helhet. Disse testene er derfor bare veiledende og kan evt. bare bekrefte prosjektgruppens egne vurderinger.





Figur 6.4 viser hvilke metoder som egner seg i de ulike fasene av produktutviklingsprosessen. Det er også noen forklaringer på hva metodene gir svar på i de ulike fasene.

Se kapittel 4 og 5 for nærmere beskrivelse av metodene.

Fase 1: Definisjon av produktegenskaper

Fokusgrupper: Behov og ønsker fra forbruker i målgruppen

Napping, sorting: Gode metoder for å få frem egenskaper – karakterisering av produkter i kategorien.
Panel er forbrukere eller prosjektgruppen.

CATA: Kartlegger ulike produkters sensoriske forskjeller og likheter. Egner seg som enkel metode som kan brukes gjennom hele PU-prosessen; «har vi nådd målet?», har produktet de egenskapene som vi tidlig bestemte var de riktige osv.

KASPER (Kryss-Alt-Som-PassER): Er en forenkelt utgave av metoden CATA. Man bruker et mindre panel gjerne bestående av personer som har spesiell kunnskap om produktet.

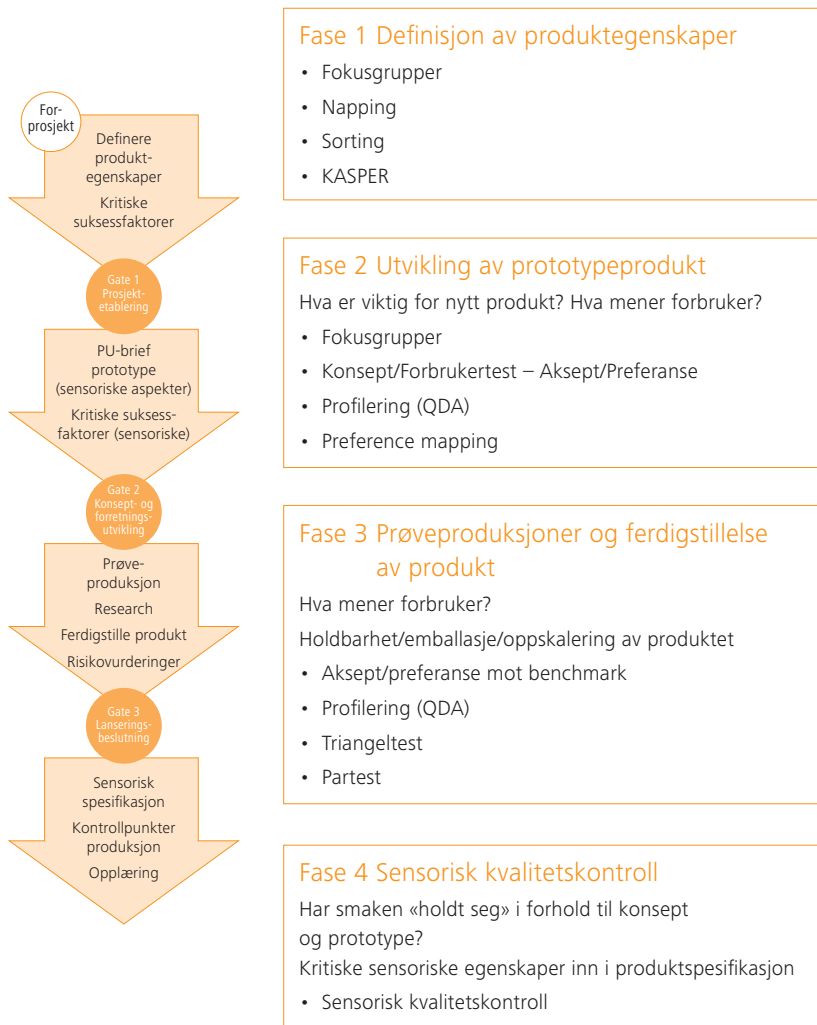
KASPER kan brukes i prosjektgruppen til måle hvor langt unna en prototype eller ulike produkter som allerede finnes i markedet er fra det ideelle produktet. Metoden egner seg også til «å følge og passe på» et produkt gjennom PU-prosessen slik at det ikke endrer seg for mye fra det som ble bestemt i den første fasen.

Fase 2: Utvikling av prototypeprodukt

Fokusgrupper og konsept/forbrukertester + preference mapping: Informasjon om forbrukernes brukssituasjon, behov, oppfattelse av prototype e.lign. Det vil si man får mer kunnskap på et tidspunkt hvor man selv har jobbet litt med produktet. Man har selv fått innsikt og erfaring og kanskje også fått svar på ubesvarte spørsmål.

Profilerings: Kunnskap om egne eller konkurrerende produkter. Dette vil også være nødvendig for å kunne utføre en preference mapping.





Figur 6.3 Aktuelle metoder i ulike steg av PU-prosessen

Illustrasjon: Julie Gjermundsen





Fase 3: Prøveproduksjoner og ferdigstillelse av produkt

Aksept/preferanse: Er produktet godt nok?

Profilering og KASPER: Har produktet de opprinnelige sensoriske egenskapene?

Triangel- og partest: Holdbarhetstester (emballasje, tid, temperatur) og test av produktets robusthet.

Fase 4: Sensorisk kvalitetskontroll

Metoder for sensorisk kvalitetskontroll og utarbeidelse av en ferdigvarespesifikasjon med kritiske sensoriske egenskaper.

Referanser

Robert G. Cooper (1993) *Winning at New Products. Accelerating the Process from Idea to Launch*. Second Edition. Addison-Wesley Publishing Company, Inc, 95-120.





Foto: Mari Svenningsen





FORSKNING

Margrethe Hersleth, Nofima

En videreutvikling av sensorisk analyse og sensorisk vitenskap er viktig for å bygge ny kunnskap. I dette kapittelet vil vi belyse to sider av sensorisk forskning. Først vil vi diskutere forskning og utvikling *innen sensoriske metoder*, dernest vil vi belyse og eksemplifisere forskning hvor sensorisk kunnskap *benyttes til andre formål*.

7.1 Forskning innen sensoriske metoder

Beskrivende analyse, slik som QDA (4.3.1) som gjennomføres med trente dommere har i mange tiår representert en viktig metode for å oppnå en fullstendig beskrivelse av de sensoriske egenskaper i et utvalg av produkter. Beskrivende analyse er imidlertid relativt ressurskrevende å gjennomføre og de siste årene har det derfor blitt utviklet en del alternative metoder for beskrivelse av produkter. Metodene benevnes som *flash-profiling*, *sorting*, *projective-mapping* og *check-all-that-apply* og omtales alle i kapittel 5.6. Felles for disse metodene er at de krever mindre eller ingen trening, dvs de kan utføres av trente dommere, semi-trente dommere og utrente dommere (forbrukere). Sensorisk litteratur viser at metodene kan være velegnet når man har behov for rask informasjon om sensoriske egenskaper. Gjennomført med forbrukere kan man også oppnå en bedre forståelse for deres opplevelse av produkter, samt forbrukeres vokabular for å beskrive matvarer. Metodene åpner opp for nye muligheter for bedrifter som ikke har mulighet til å trene og å opprettholde et trent sensorisk panel. Det er imidlertid viktig å være klar over metodenes styrker og svakheter, og sammenliknende studier har derfor vært et viktig forskingsfelt de siste årene. En annen sensorisk metode som nylig har fått mye oppmerksomhet er Temporal Dominance of Sensation (TDS). Metoden er beskrevet i 4.3.4.





TDS er en metode som egner seg til relativt komplekse produkter slik som vin eller sjokolade, men har også være benyttet til finjustering av resepter for lav-salt produkter. Ved Nofima har metoden også blitt testet ut ved sensorisk bedømmelse av måltider (Paulsen, 2013). I dette prosjektet ble det gjennomført sensorisk bedømmelse av sensorisk samspill mellom laks og sauser med ulike grunnsmaker og målsetningen var å se hvordan sausene påvirket smaken av laks.

Forskning for å oppnå gode metoder for å innhente kunnskap om forbrukeres opplevelse av mat har pågått kontinuerlig i mange år og har oppnådd mer og mer oppmerksomhet ettersom fagområdet har utviklet seg. Det har vært arbeidet med mange ulike temaer og vi kan nevne følgende: Sensometri og statistisk analyse, validitet i sensoriske responser, sensorisk sensitivitet, betydning av genetikk, tverrkulturelle studier, effekter av kontekstuelle faktorer og betydning av følelser og assosiasjoner ved matopplevelse.

7.2 Forskning basert på sensorisk kunnskap

Sensorisk vitenskap og sensorisk kunnskap er en viktig brikke innen mange andre fagområder. Innenfor ernæringsvitenskap er sensorisk analyse en nødvendighet. Dette kan presiseres gjennom følgende utsagn:

«Maten er ikke sunn før den er spist».

Hvis man skal gjennomføre kostholdstudier er det viktig at dietten settes sammen på en sensorisk optimal måte slik at respondentene klarer å gjennomføre intervensjonen uten at de går trette og faller fra. Når man skal utvikle sunne produkter med for eksempel redusert fett, salt og/eller sukker er sensorisk analyse en nødvendighet. Reseptoptimalisering kombinert med målrettet produktinformasjon vil kunne bidra til at forbrukere velger og spiser sunne produkter.

Kulinarisk vitenskap er et annet område som inkluderer sensorisk kunnskap. Begrepet «Molekylær gastronomi» har flere definisjoner. Vi nevner her en fra Hervé This, (2008): «Molecular gastronomi is a branch of science that studies the physic-chemical transformations of





edible materials during cooking and the sensory phenomena associated with their consumption». I en doktorgrad innen molekylær gastronomi utført ved Universitet i København med tittel: «Emulsions from a culinary perspective – the case of Hollandaise sauce and its derivatives» (Rognså, 2014), ble sensorisk analyse benyttet som målemetode for å studere endringer i tekstur og smak i hollandes saus som følge av ulike tilberedningsmetoder og ingredienser (vin og vinreduksjoner).

Forbrukergruppene barn og eldre har de siste årene fått mye oppmerksomhet knyttet til betydning av sensorisk kunnskap for riktig ernæring og for å motvirke utvikling av fedme. Kunnskap om mekanismer som påvirker barns smaksoppfattelse og deres kosthold er viktig ettersom grunnlaget for en god helse legges i barndommen.

Barn spiser det de liker og liker det de kjenner, og utvikling av matvaner er derfor en kritisk faktor i et barns oppvekst.

Forskning innen sensorisk sensitivitet og preferanser for ulike matvarer er viktig og nødvendig og må kombineres med kunnskap om faktorer som neofobi, eksponering, sosiale faktorer og rollemodeller.

Behovet for sensorisk forskning relatert til kosthold hos eldre er motivert av andre faktorer. Først og fremst blir gruppen eldre en stadig større del av befolkningen ettersom levealderen øker og antall barn i familiene går ned. Dernest skjer det en endring i sensorisk opplevelse hos eldre særlig gjennom svekking av luktesans og tyggemusklene (tenner). Det er nødvendig med mer kunnskap om disse sensoriske endringer og det er dernest viktig å utvikle matretter som er tilpasset eldres behov. I tillegg bør mattilbudet til eldre tilpasses både mhp porsjonsstørrelser, forenklet tilberedning og tilgjengelighet.

Sensorikk relatert til holdbarhet og/eller emballering av næringsmidler er et annet eksempel på et viktig forskningsmessig anvendelsesområde. En større studie på holdbarhet på laks med ulike bakteriekulturer og ulik lagringstid ble nylig gjennomført på Nofima på Ås. Et trent panel gjennomførte bedømmelse av laks basert på lukt-egenskaper og resultatene viste gode korrelasjoner mellom antall lagringsdager, utvikling i ulike bakterietall og utvikling i aromakomponenter.





Kunnskap innenfor sensorisk persepsjon er også viktig innen forbruker- og/eller markeds kunnskap. De siste årene har det fremkommet viktig kunnskap om betydning av produktinformasjon for forbrukeres forventninger til mat og samspillet mellom sensorisk opplevelse og forventningene. Forskningen viser at hva vi vet om maten påvirker våre forventninger til spiseopplevelsen. Hvis forbrukeres forventninger er positive trekkes opplevelsen som oftest i positiv retning, men hvis forbrukernes forventninger er negative trekkes opplevelsen ofte i negativ retning.

Referanser

Rognså, G.H., (2014), Emulsions from a culinary perspective – the case of Hollandaise sauce and its derivatives, Department of Food Science, Faculty of Science, University of Copenhagen, Denmark.
ISBN 978-87-996934-2-9

Paulsen, M.T. (2013). Sensory interactions in meals: influence of tasting techniques, dynamic aspects and culinary principles, Department of Chemistry, Biotechnology and Food Science, Norwegian university of Life Sciences, Norway.
Thesis number 2013:55. ISSN1503-1667, ISBN 978-82-575-1155-5

This, H. (2008). Molecular gastronomy, a scientific look at cooking. INRA/AgroParisTech.





KVALITETS- KONTROLL

Foto: TINE





KVALITETSKONTROLL

Hilde Kraggerud og Elin Simonstad Valle, TINE FoU

De fleste produsenter av næringsmidler har en detaljert beskrivelse av alle sine produkter. I en slik produktspesifikasjon inngår som oftest ingredienssammensetning, næringsinnhold og merking i tillegg til kjemiske, fysiske, mikrobiologiske og sensoriske normer. Disse normene bygger fortrinnsvis på forbrukeres synspunkter på hvordan produktet bør være. Ofte gjøres det forbrukertesting i produktutviklingen for å optimalisere produktet.

De sensoriske normene beskriver de sensoriske egenskapene ved produktene som utseende, lukt, smak og konsistens. For ulike produkter kan disse beskrivelsene være mer eller mindre detaljerte. Det er eksempelvis enklere å beskrive utseendet på en storhullet, fast ost enn å beskrive smaken. For noen produkter, der det er stor forskjell på hvordan produktet ser ut på overflaten og inni, snakker vi ofte om indre og ytre utseende.

Den sensoriske beskrivelsen i produktspesifikasjonen beskriver som oftest det ideelle produktet, og toleransegrenser for variasjon. Ved sensorisk analyse i kvalitetskontroll bedømmer man i hvilken grad produktet er i henhold til norm, og eventuelt hvordan det skiller seg fra normen (se kapittel 4. 3). Siden det er vanskelig å beskrive et produkt sensorisk slik at en person som er helt ukjent med produktet kan bedømme om det holder mål, er man avhengig av at de som skal bedømme produktet vet hvordan det skal være. De må trene på å bedømme produktene de skal være dommer for sammen med godt trent dommere, og opparbeide seg en «memory standard» for hvordan produktene skal være.





Sensorisk analyse i kvalitetskontroll kan utføres på ulike steg i produksjonsprosessen

- Råvarekontroll
- Prosesskontroll
- Ferdigvarekontroll

Råvarekontroll

Når man får inn nye råvarer i bedriften bør disse kvalitetskontrolleres for å sikre at man opprettholder en jevn og korrekt kvalitet på ferdigproduktene. Ofte gjøres dette ved at man kontrollerer mot en prøve fra en tidligere godkjent batch av råvaren. For noen råvarer kan man sammenligne de ulike batchene direkte mens for andre må man lage prøver av det ferdige produktet som deretter kvalitetskontrolleres.

Prosesskontroll

Under prosessen er det viktig at man hele tiden sikrer at produktet holder ønsket kvalitet, derfor legger mange produsenter inn sjekkpunkter i prosessen der de går inn og kontrollerer produktet. Man kan kontrollere flere steder underveis, avhengig av hvor kompleks produksjonen er, men som et minimumskrav bør man kontrollere første produkt som kommer ut av pakke/tappe-maskin for å være sikker på at det er rett produkt i rett emballasje og med riktig kvalitet. Produsentene bør trene opp operatørene til å gjøre slik prosesskontroll.

Ferdigvarekontroll

Ferdig produkt kan kontrolleres gjennom hele holdbarhetstiden. Man kontrollerer straksprøver for å sjekke at kvaliteten er i henhold til standard. Dersom man har mulighet kan det være lurt å fremprovosere eventuelle feil som kan oppstå under lagring for å oppdage prøver der man kan få kvalitetsproblemer i løpet av lagringstiden. En slik fremprovosering kan f.eks. være lagring av melk fra straks etter tapping i ett døgn ved 17 °C. Man bør også gjøre jevnlig kvalitetskontroll ved holdbarhetstidens utløp for å sjekke at man klarer å holde kvaliteten gjennom hele holdbarhets-perioden.





Man kan også som et ledd i kontrollen av ferdigvarer gå ut å kjøpe inn fra omsetningsleddet og deretter kvalitetsbedømme disse produktene. Da får man et bilde på hvordan produktene klarer seg gjennom transport, lagring hos grossist og under lagring og eksponering i butikk.



Figur 8.1 (kapitelforside) og 8.2 Kontroll av ost

Foto: TINE

Dommere til kvalitetskontroll

Det er ofte hensiktsmessig å benytte egne ansatte til kvalitetskontroll, dvs personer som er ansatt for hovedsakelig å utføre andre typer arbeider enn sensorisk kvalitetskontroll. Sensorisk kvalitetskontroll blir en spesifisert del av arbeidsoppgavene til den ansatte. Man må trene dem opp slik at de kjenner de aktuelle produktene og opparbeider en «memory standard». Hvis ønskelig kan man gjennomføre en sertifisering av dommerne slik at man vet at de har de kunnskapene de trenger. Dommerne bør følges opp ved at man av og til legger inn prøver med vanlige feil som kan forekomme i gjeldende produkt for å sjekke at dommeren klarer å gjenkjenne disse.

Bruk av skala

Det kan benyttes flere ulike typer skala ved kvalitetskontroll. Det enkleste er «godkjent/ikke godkjent». Det vanligste er en tall-skala som





kan ha varierende lengde. I /ISOIDF standarden for meieriprodukter er det for eksempel definert en skala fra 1-5 poeng. I Danmark benyttes en 15 punkts skala. Poeng benyttes for å beskrive produktkvaliteten, om produktet svarer til spesifikasjonen eller graden av avvik fra produktets gjeldende sensoriske spesifikasjon og gis i henhold til en klart definert skala. Bedriften bør definere klare regler for anvendelse av produkter med ulik grad/type av avvik, f.eks. ved å ha en poeng-grense for hva som skal selges som «normalt» produkt, og hva som må omarbeides, evt. anvendes alternativt eller i ytterste konsekvens destrueres.

Når man bedømmer en prøve til å ha avvik er det vanlig at man må angi hvilken feil man har trukket for. Ofte lages det en liste over de vanligste feil som kan oppstå i ulike produktgrupper (feilnomenklatur). Man kan gradere feilene etter hvor alvorlige de er og hvilke poeng man kan gi for ulike feil.

Standardiserte metoder

Det finnes metodestandarder innenfor sensorisk kvalitetskontroll, både fra NMKL, ISO og ASTM (se kapittel 14).

Referanser

Kilcast, D, editor (2010). *Sensory Analysis for Food and Beverage Quality Control, A Practical Guide*, Woodhead Publishing Ltd. (ISBN: 978-1-84569-476-0)

Hersleth, M., Ilseng, M. A., Martens, M., & Næs, T. (2005). Perception of cheese: A comparison of quality scoring, descriptive analysis and consumer responses. *Journal of Food Quality*, 28, 333–349.

Kraggerud, Solem & Abrahamsen (2012). Quality scoring – a tool for sensory evaluation of cheese? *Food Quality and Preference*, 26, 221-230.

International Organization for Standardization, & International Dairy Federation. (2009). *ISO 22935/IDF 99 milk and milk products – Sensory analysis – Part 1, 2, 3*

QIM metoden: <http://www.qim-eurofish.com/>





§





OFFENTLIG TILSYN

Egil Weie Berg, Mattilsynet

9.1 Innledning

Offentlig regelverk har lang historisk tradisjon for å ha krav til sensoriske egenskaper i næringsmidler.

Mattilsynet har tilsynet med planter, fisk, dyr og næringsmidler samt kosmetikk og legemidler solgt utenom apotek. Denne virksomheten omfatter objektiv bedømmelse av produkters egenskaper ved hjelp av menneskelige sanser.

Forbrukerne har sensorisk kvalitet som et viktig kriterium i vurdering av et næringsmiddel. Forbrukere kan henvende seg til Mattilsynet med spørsmål angående næringsmidler og ofte gjelder dette sensoriske egenskaper ved produktene eller mistanke om næringsmiddelrelatert sykdom.

Mattilsynet har ikke egne sensoriske laboratorier eller sensoriske paneler. Enklere analyser av sensoriske parametere skjer på laboratorier som brukes til flere formål eller andre rom som er egnet til slik bruk. Et eksempel på dette er ved soppkontroll hvor publikum kan få hjelp til sort/artsbestemmelse. Ved behov kan Mattilsynet kjøpe kvalitetssikrede sensoriske analyser fra private eller offentlige laboratorier.





9.2 Inspeksjon og tilsyn

Ved inspeksjoner vil tilsynet delvis basere seg på inspektørens faglige skjønnsmessige vurdering av sensoriske inntrykk.

Sensorisk vurdering er ikke nødvendigvis tilstrekkelig grunnlag i seg selv til å reagere med virkemiddelbruk som f.eks. forvaltningsvedtak. Kjemiske eller mikrobiologiske analyser er ofte nødvendig i tillegg til de sensoriske analysene. Dette kan for eksempel gjelde ved mistanke om ulovlig bruk av tilsetningsstoffer. Unntak er hvis virksomheten selv bekrefter avviket. På noen områder som for eksempel ved vurdering av lukt og konsistens av rå fisk kan sensoriske observasjoner av faktiske forhold være hovedgrunnlaget for pålegg og vedtak.

Sanseinntrykk fra lokalene i tilknytning til produksjon eller lagring av næringsmidler og/eller emballasje kan gi grunnlag for videre undersøkelser i en virksomhet. For eksempel kan maling eller desinfeksjonsmidler avgi lukt som kan endre næringsmidlenes sensoriske kvalitet.

9.3 Regelverk

Dagens lover og forskrifter stiller krav til sensoriske egenskaper ved omsetning av næringsmidler. Slik var det også i Christian Vs lov fra 1682 som hadde bestemmelser om straff for den som forfalsket mat eller kamouflerte bedrevet mat.

Regelverket som forvaltes av Mattilsynet er kontinuerlig i endring. Myndighetene har nå mindre fokus på sensorisk produktkvalitet og mer fokus på hygienisk og helsemessig kvalitet.

Norge har de senere år tilpasset sitt regelverk til EUs regelverk på en rekke områder. I tillegg til næringsmidler, gjelder dette også dyrevern, dyrehelse og plantehelse.

Sentralt for næringsmidler er *forskrift om næringsmiddelhygiene (næringsmiddelhygieneforskriften)* med forordning 852/2004 som stiller krav til fareanalyse og vurdering av kritiske styringspunkter (HACCP – Hazard Analysis & Critical Control Points). For hvert styringstiltak





som er tilknyttet et kritisk punkt, bør det spesifiseres kritiske grenser/grenseverdier for aksept (se kapittel 4.3, kvalitetskontroll). De blir fastsatt for parametere som kan observeres eller måles og dermed kunne gi dokumenterte bevis i forbindelse med en kontroll av prosessen. Eksempler hvor slike verdier fastsettes er sensoriske parametere som utseende, lukt eller tekstur.

Forskrift om hygieneregler for næringsmidler av animalsk opprinnelse har en rekke bestemmelser om visuelle undersøkelser av kjøtt, fisk og fiskevarer. Også forskrift om kvalitet på fisk og fiskevarer har slike bestemmelser.

Sensoriske analyser gir ofte opplysninger av stor verdi for vurdering av vannkilden. *Drikkevannsforskriften* har bestemmelser om at materialer i ledningsnett og i vannforsyningsystemet ikke skal forringe vannets sensoriske egenskaper. Drikkevannet skal være klart og uten framtreddende lukt, smak eller farge når det leveres til mottakeren. Forskriften har også en tabell med grenseverdier for de sensoriske parameterne lukt og smak. Vannverkene skal undersøke sensoriske parametere som en del av sin rutinekontroll. Mattilsynet fører tilsyn med at vannverkene foretar disse undersøkelsene.

Forskrift om mineralvann og kildevann, sier at «Naturlig mineralvann og kildevann skal ved omsetning ikke ha noen sensoriske feil».

Forskrift om tilsetningsstoffer til næringsmidler 2011-06-06-668 regulerer bruk av tilsetningsstoffer. En av de akseptable grunnene til å bruke tilsetningsstoffer kan være for å forbedre produktets sensoriske egenskaper. Regelverket forutsetter at man ikke endrer næringsmidlets art, beskaffenhet eller kvalitet på en måte som kan villedde forbrukeren. Man kan for eksempel ikke bruke fargestoffer for å gi inntrykk av at en matvare er friskere eller ferskere enn den i virkeligheten er.

En gruppe tilsetningsstoffer er smaksforsterkere. Dette er stoffer som forsterker et næringsmiddels naturlige smak.





Andre eksempler på tilsetningsstoffer som påvirker sensoriske egenskaper er:

- Syrer som øker et næringsmiddels surhetsgrad og/eller gir det en sur smak.
- Søtstoffer som gir et næringsmiddel søt smak. Næringsmidler med søtende egenskaper regnes ikke som søtstoffer i denne forskriften.
- Konsistensbevarere som gjør eller holder frukt eller grønnsaker faste eller sprø, eller som reagerer med geleringsmidler og danner eller styrker en gel.

Næringsmiddelhygieneforskriften omhandler også bestemmelser for lokaler, innredning og utstyr. Alle gjenstander og anlegg i næringsmiddelvirksomheter skal være utformet slik at det er minst mulig fare for forurensing av næringsmidlene. Dette gjelder også i forhold til sensorisk kvalitet. Innredning og luftkvalitet skal for ikke avgi stoffer som setter smak eller lukt på næringsmidlene.

Emballasjeforskriften har bestemmelser om materialer og gjenstander som kommer i kontakt med næringsmidler. Disse skal under normale eller forutsigelige bruksvilkår ikke avgi stoffer til næringsmidler i en slik mengde at de kan føre til en forringelse av næringsmidlenes sensoriske egenskaper.

Aktive materialer og gjenstander i kontakt med næringsmidler er definert som «materialer og gjenstander beregnet på å forlenge holdbarheten eller forbedre tilstanden til pakkeede næringsmidler». De er utformet for bevisst å omfatte bestanddeler som vil avgi stoffer til eller absorbere stoffer fra det pakkeede næringsmiddelet eller miljøet rundt det.

For eksempel skal ikke aktive materialer og gjenstander villedende forbrukeren, de skal ikke skjule vond lukt eller at næringsmidlene er bedrevet. Aktive materialer og gjenstander kan bare endre næringsmidlenes sensoriske egenskaper dersom det er i samsvar med regelverket. Regelverket på Mattilsynets fagområde blir oppdatert og utviklet. Nærmere informasjon finnes på www.mattilsynet.no og www.lovdatab.no.





9.4 Metoder og analyser

Internasjonale standarder, for eksempel de som utgis av Nordisk Metodikkomité for Næringsmidler (NMKL) og den internasjonale standardiseringsorganisasjonen (ISO) blir akseptert og anbefalt brukt i tilsynssammenheng. Det er kvalitetssikringskrav til sensoriske analyser når Mattilsynet skal benytte resultatene til forvaltningsvedtak.

Eksempler på områder hvor sensorikk brukes i praksis:

Kjøttkontroll

Mattilsynet utfører visuell kontroll av slakt på slakteriene. Slakteskrotter blir inspisert for synlige forandringer som kan skyldes sykdom eller skader. Tilsynspersonalet tar også prøver av kjøtt som kokes for så å bli luktet og smakt på. Dette kan avdekke for eksempel rånelukt og rånese smak eller andre avvik. Koke- og pH-prøve skal tas ut i de tilfeller en er i tvil om kjøttets sensoriske kvalitet. Ved forekomst av flyktige luktstoffer kan det være best å lukte på selve slakteskrotten.

Vegetabilier

Potetforskriften, som er en kvalitetsforskrift, har bestemmelser om at potetene skal være fri for fremmed lukt og smak og at de skal være faste. Potetkontrollen utføres blant annet ved å lukte, se på og føle på potetene. Potetene deles opp for å oppdage indre sensoriske avvik. Enkelte skader på potet kan gi bestemte lukter. Mattilsynet utfører kontroll av poteter.

Forskrift om vegetabiliske konserver sier at varene skal være pakket i emballasje som ikke påvirker lukt, smak eller farge. Forskriften har en rekke krav til sensoriske kvalitetsegenskaper for frukt og grønnsaker.

Fisk/sjømat

Bedervelsesbakterier vil raskt kunne gi fisk en vond lukt. Dette gjør at luktanalyse er et naturlig og godt utgangspunkt for bedømmelse av ferskhet.

Både tilsyn og forbruker kan gå ut i fra at hel sløyet fisk er fersk nok til å bli spist dersom den har frisk lukt, smaker godt og har fast elastisk muskulatur. Gammel fisk vil derimot ofte lukte vondt, være slimete, og





ha redusert glans og spenst. Dersom tilsynet finner fisk med slike sensoriske avvik kan det være grunnlag nok for å nekte omsetning av varen.

Sopp

Flere av de ansatte i Mattilsynet har avlagt eksamen i regi av Norges sopp- og nyttevekstforbund for å bli soppsakkyndige.

I Mattilsynets soppkontroll brukes sanseintrykk inkludert vurdering av lukt og smak. Dette gjøres blant annet ved soppkontroll som i soppsesongen utføres av kontrollører ved en rekke av Mattilsynets distriktskontorer.

Referanser

Forskrift 2008-12-22-1623 om næringsmiddelhygiene (næringsmiddelhygieneforskriften)

Kommisjonsforordning (EF) nr. 2073/2005 av 15. november 2005 om mikrobiologiske kriterier for næringsmidler

Forskrift 2008-12-22-1624 om særlige hygieneregler for næringsmidler av animalsk opprinnelse (animaliehygieneforskriften)

Forskrift 2001-12-04-1372 om vannforsyning og drikkevann (drikkevannsforskriften)

Forskrift 2004-10-04-1316 om naturlig mineralvann og kildevann

Forskrift 1993-12-21-1381 om materialer og gjenstander i kontakt med næringsmidler (matkontaktforskriften)

Forskrift 2011-06-06-668 om tilsetningsstoffer til næringsmidler

Forskrift 2013-06-28-844 om kvalitet på fisk og fiskevarer

www.mattilsynet.no

www.lovdatabasen.no





KVALITETS- SIKRING AV SENSORISKE ANALYSER



Foto: Mari Svenningsen





KVALITETSIKRING AV SENSORISKE ANALYSER

Urd Bente Andersen, Orkla Foods Norge

Akkreditering, Sertifisering og God laboratoriepraksis (GLP)

For noen laboratorier er det viktig å ha en godkjenning som gjør dem i stand til å dokumentere sine analyser. Det kan være analyser som krever ekstra nøyaktighet eller der analysene har stor økonomisk betydning. Offentlige institusjoner eller større bedrifter krever ofte slik dokumentasjon. Eksempler på slike ordninger er ISO sertifisering, sertifisering etter BRC standard, akkreditering eller GLP godkjenning.

En *akkreditering* og en *sertifisering* vil si en uavhengig bedømmelse av en virksomhets kompetanse og evne til å utføre en bestemt oppgave slik at den er i overensstemmelse med gitte spesifikasjoner. Bedømmelsen utføres av en upartisk og uavhengig tredjepart. Akkreditering er den strengeste ordningen og akkrediteringsmyndigheten blir utnevnt av statlige myndighetene. En sertifisering er mindre streng og utføres ofte av private stiftelser eller organisasjoner som Det Norske Veritas.

Norsk Akkreditering (NA) utfører teknisk akkreditering i Norge. NA er etablert som forvaltningsorgan under Nærings- og handelsdepartementet. I Danmark er Danske Akkrediterings- og Metrologifond (DANAK) nasjonalt akkrediteringsorgan, i Sverige Styrelsen för akkreditering och teknisk kontroll (SWEDAC) og i Finland er det The Finnish Accreditation Service (FINAS) som har denne oppgaven.





Krav til akkrediterte virksomheter er fastlagt i internasjonale avtaler. Slike avtaler er:

- European cooperation for Accreditation (EA)
- International Laboratory Accreditation Cooperation (ILAC)
- International Accreditation Forum (IAF)

Hvilke standarder som skal følges avhenger av område som skal akkrediteres. Internasjonale avtaler om gjensidig aksept av akkreditering gjør at en akkreditering i Norge vil være likeverdig med tilsvarende akkrediterte analyser fra et laboratorium i f.eks i Sverige. Et akkreditert laboratorium må følge kravene i standarden EN ISO/IEC 17025.

God laboratoriepraksis (GLP) er en spesiell godkjenning som brukes av offentlige myndigheter. Formålet er å få fram opplysninger om produkter/kjemikaliers egenskaper og deres potensielle skadelige effekter på helse og miljø. Alle laboratorier som skal rapportere innenfor områdene legemiddel, pesticidier (plantevernmidler), kosmetiske produkter, veterinærmedisin, tilsetningsstoffer til mat og fôr samt industrikjemikalier skal være GLP godkjente. Disse laboratoriene må følge GLP prinsippene fastsatt av OECD i The OECD Principles of Good Laboratory Practice. I Norge er det Norsk Akkreditering (NA) som gir teknisk akkreditering av GLP.

GLP slik den brukes av myndighetene har en spesiell hensikt og ikke ment for sensoriske analyser. Det bør imidlertid være mulig å bruke prinsippene også i sensoriske laboratorier – spesielt for de laboratorier som jobber med vitenskapelige undersøkelser.

Akkreditering vil si en uavhengig bedømmelse og formell anerkjennelse av et prøvningslaboratorium som gjør det i stand til å utføre bestemte prøvninger eller bestemte typer av prøvninger som f. eks sensoriske analyser.

Det er forskjellige måter å akkreditere laboratorier på avhengig av hva slags oppgaver laboratoriet har. Det skilles mellom prøvningslaboratorier og kalibreringslaboratorier. De følger samme standard, men kravene er forskjellige på noen punkter. NA akkrediterer også





sertifiseringsorgan og kontrollorgan som f. eks sertifiseringsorgan for kvalitetsstyringssystemer, for miljøstyringssystemer eller for personell samt GLP godkjenning som nevnt tidligere.

Prøvings- og kalibreringslaboratorier blir akkreditert etter EN ISO/IEC 17025 Generelle krav til prøvings og kalibreringslaboratoriers kompetanse. Krav til prøvetaking inngår også som en del av standarden. Sensoriske laboratorier har i tillegg nytte av å følge retningslinjene gitt i *EA-4/09 Accreditation for sensory testing laboratories* som gir spesielle råd om akkreditering av sensoriske laboratorier.

Akkrediterte virksomheter skal oppfylle generelle krav til virksomhetens organisasjon, kvalitetsstyring, uhildethet og kompetanse. Det stilles spesifikke krav til faglig kompetanse angående metode som brukes i prøvings- og kalibreringslaboratorier. Virksomheten skal kunne dokumentere kontroll vedrørende spesielle krav til prosesser, produkter eller personer. Det innebærer at de må kunne dokumentere overfor kunder, myndigheter og andre samarbeidspartnere at virksomhetens produkter eller serviceytelser lever opp til kravene i standarden. Det betyr bl.a at sertifiserings- og inspeksjonsvirksomheter må ha personale som er akkreditert til å utføre inspeksjoner eller blitt gitt myndighet til å gi en sertifisering.

ISO/IEC 17025:2005, Generelle krav til prøvings- og kalibreringslaboratoriers kompetanse dekker prøving og kalibrering som utføres i henhold til

- Standardiserte metoder
- Ikke-standardiserte metoder
- Metoder som er utviklet av laboratoriet

Når et laboratorium ikke utfører en eller flere av de aktiviteter som er dekket av standarden, slik som prøvetaking og utvikling av nye metoder, gjelder ikke kravene i de punktene. For øvrig må alle deler av standarden være oppfylt. Oversikt over hvilke laboratorier som er akkreditert i Norge finnes på akkreditert.no





Akkreditering av sensoriske analyser innebærer:

- En frivillig ordning
- Skal sikre kvaliteten i det arbeidet en virksomhet utfører
- Krav til ledelse
- Tekniske krav
- Akkrediteringsorganet gir akkreditering når de finner at virksomheten som søker om akkreditering tilfredsstillere alle akkrediteringskravene
- For å beholde akkrediteringen må bedriften bevise for akkrediteringsmyndigheten at de klarer å tilfredsstillere kravene

Akkreditering av sensoriske metoder innebærer at hver analyse skal være under kontroll. Laboratoriet skal ha definerte grenser som gir samme resultat, og når det er mulig gi sammenlignbare resultat med andre laboratorier. Det er laboratoriets ansvar å kunne demonstrere for akkrediteringsmyndighetenes bedømmere at ved å bruke riktige teknikker, så oppnår de kriteriene for akkreditering. Akkreditering kan omfatte fagspesifikke, nasjonale og internasjonale standardiserte metoder eller dokumenterte validerte egne (in house) metoder.

Dokumentet EA-4/09 tar for seg sensoriske analyser og hva som er viktig ved akkreditering av disse. Den er utarbeidet av European cooperation for Accreditation (EA), Laboratory Committee, Food Testing Task Force og er et supplement til EN ISO/IEC 17025. Dokumentet EA-4/09 kan betraktes som et «anvendt dokument» og er nyttig både for sensoriske dommere og sensoriske laboratorier som forbereder seg til akkreditering

Sikkerhet

Hensyn til dommerne er viktig. Dommernes sikkerhet skal alltid komme i første rekke. Man skal f.eks ikke servere produkter som kan påføre dommeren helseskade eller allergiske reaksjoner. Man må ta hensyn til sammensetningen av prøvene, lengden på prøveseriene, passe på å legge inn pauser osv. for å unngå tretthet hos dommeren.





Det finnes også andre standarder som er nyttige for sensoriske laboratorier som ønsker å akkreditere eller kvalitetssikre sine analyser. Nordisk Metodikkomite for næringsmidler (NMKL) har utgitt flere prosedyrer som sensoriske laboratorier kan ha nytte av.

EA-4/09 tar for seg akkreditering av objektive analyser som er dokumentert og validert. Det innebærer at hver analyse skal være under kontroll. Laboratoriet skal ha definerte grenser som gir samme resultat og når det er mulig kunne gi sammenlignbare resultat med andre laboratorier. Eksempler på sensoriske analyser som kan akkrediteres er

- Forskjellstester: triangeltest, par -test, rangering
- Beskrivende tester: kvantitative beskrivende tester som profilering

Det er laboratoriets ansvar å kunne demonstrere for akkrediteringsmyndighetenes bedømmere at ved å bruke riktige teknikker så oppnår de kriteriene for akkreditering

Et akkreditert sensorisk laboratorium må kunne dokumentere (EA-4/09):

- Validering av metode
- Dokumentasjon av metode
- Trening og autorisasjon av personale
- Å ha tilfredsstillende lokaler
- Planlegging, organisering og styring av lokaler
- Ivareta og kalibrering av utstyr
- Prosedyrer for utvelgelse og trening av dommere
- Kontinuerlig kvalitetskontroll
- Kontinuerlig overvåking av sensoriske dommere og sensorisk panel
- Benytte passende referansematerialer eller annet treningsmateriale
- Prosedyrer for å sjekke data
- Overvåke og dokumentere testforholdene





Krav til metode og validering av metode (ISO 17025 § 5.4)

Laboratoriet bør bruke oppdaterte og standardiserte metoder. Det vil ofte være nødvendig å beskrive metoden i en prosedyre. Egne metoder som skal akkrediteres må være dokumentert, validert og evaluert (prøvd). Metoden skal inneholde informasjon om (EA-4/09):

- Krav til trening av dommere
- Prøvepreparering og presentasjon
- Sammensetning av det sensoriske panelet
- Overvåking av dommer og dommerprestasjoner
- Spesielle krav til omgivelser og lokaler
- Statistiske analyser av resultater

For å sikre at samme sensoriske problem alltid gjennomføres på samme måte, må det lages en prosedyre for praktisk gjennomføring av metoden der hele prosessen må være godt dokumentert, og hvert steg og type personell som er ansvarlig er beskrevet.

Angivelse av usikkerheten for sensoriske analyser som nøyaktighet, deteksjonsgrense, sensitivitet, linearitet, repeterbarhet, reproduserbarhet, robusthet og kryss-sensitivitet er vanskelig og vil bare kunne angis på en forenklet måte. Som regel settes det krav til kontroll med dommernes prestasjoner ved repeterbarhet (gjentak og omprøving) og reproduserbarhet (i sammenlignende laboratorieprøvinger, SLP).





Referanser:

Generelle retningslinier for kvalitetssikring af sensoriske laboratorier (NMKL-prosedyre nr 6, 1998)

OECD's Principle of GLP, No 1: OECD Principles on Good Laboratory Practice (ENV/MC/CHEM(98)17)

EN ISO/IEC 17025 Generelle krav til prøvings- og kalibreringslaboratoriers kompetanse

EA-4/09 Accreditation For Sensory Testing Laboratories

NMKL-procedure nr 4 (2010) Vejledning i kvalitetssikring for kemiske levnedsmiddellaboratorier

NA Dok. 41 God laboratoriepraksis. Den norske GLP-ordningen.

www.akkreditert.no





STATISTIKK





STATISTISK BEHANDLING OG VURDERING AV DATA

Per Lea, Nofima

11.1 Innledning

Tidligere i boka har vi gjennomgått forskjellige metoder som nyttes ved sensorisk analyse. Videre har vi lært hvilke metoder og paneltyper som er mest hensiktsmessige i ulike bedømmelsessituasjoner. Når slike analyser er utført, vil en sitte igjen med et tallmateriale, men det er ikke uten videre gitt hvilke konklusjoner vi kan trekke på grunnlag av disse tallene.

Datasettet er alltid et utvalg – at vi endte opp med det datasettet vi virkelig gjorde, har et element av tilfeldighet i seg. Men vi ønsker vanligvis ikke å uttale oss om det konkrete utvalget vi har endt opp med; vi ønsker å generalisere resultatene. Hvis f.eks. et forsøk resulterer i at italiensk salat basert på råvarer fra leverandør A har mer gulrotsmak enn italiensk salat basert på råvarer fra leverandør B, så er det ikke noe interessant om dette gjelder bare for de 10 pakningene som inngikk i forsøket. Forsøksopplegget og den statistiske testen må være av en slik art at vi med en viss grad av sikkerhet kan si at leverandør A leverer råvarer som gir mer gulrotsmak på den italienske salaten over en lengre tid. Hvis vi ikke er villige til å generalisere ut fra det utvalget som inngår i selve forsøket, får vi en helt absurd situasjon: da må vi gjøre omfattende sammenlikninger hver eneste dag, og kanskje bytte leverandør flere ganger i uka. Mesteparten av det som produseres vil trolig gå med til testing, og dermed bli spist opp, slik at det nesten ikke blir noen ting igjen å selge!

Kan vi ikke generalisere resultatene av den sensoriske analysen vil det heller ikke være noen grunn til å utføre den!





11.2 Gangen i en statistisk test

En hypotese (nullhypotese, ofte symbolisert H_0 og uttalt «H-null») framsettes. Denne kan f.eks. være at italiensk salat produsert etter resept A, B og C alle har like mye gulrotsmak. Etter en sensorisk bedømmelse kan kanskje resultatet være at panelet har gitt salat A verdien 6.2 for gulrotsmak, B har fått 5,9, og C har fått 3,9. Når vi skal avgjøre om en slik hypotese kan forkastes eller ikke, må vi se på hvor sannsynlig det resultatet vi observerer i forsøket, faktisk er. Å beregne en slik sannsynlighet er ikke uten videre enkelt, og for å komme videre må vi gjøre noen antakelser. Vi antar rett og slett at nullhypotesen er sann, dvs. at de tre reseptene faktisk gir like mye gulrotsmak «i det lange løp». Har vi først antatt dette, betyr det at de forskjellene i middelveidier vi har observert, bare skyldes tilfeldigheter.

I disse middelveidene inngår at vi har «midlet bort» dommerne, og også eventuelle gjentak og andre faktorer som inngår i forsøket. Siden vi nå har antatt at de tre reseptene gir like mye gulrotsmak, så kan vi også beregne sannsynligheten for at resultatene skal avvike så mye fra hverandre – eller mer - enn det vi har observert. At resept C har endt opp med en så lav verdi som 3,9, minst 2 sensoriske enheter mindre enn de to andre, kan synes mistenkelig. Men det er bare det statistiske forsøksopplegget som kan gi oss svar på om denne forskjellen er «stor nok». Eller sagt på en annen måte: hvor sannsynlig er det at vi får så store forskjeller som 2,3 og 2 enheter mellom de to høyeste og den laveste verdien? Hvis denne sannsynligheten er «liten», er det grunn til å tro at ett eller annet er «galt».

Og hva kan så være galt? – vi forutsetter selvsagt at alle beregninger og praktiske ting rundt forsøket har vært utført på en korrekt måte. Da gjenstår bare den antakelsen vi startet med, altså nullhypotesen. Hvis antakelsen om at de tre reseptene gir like mye gulrotsmak er lite sannsynlig, betyr det at antakelsen ikke er sann, det vil si at vi forkaster nullhypotesen. Det er det samme som at nullhypotesen ikke er sann, som i sin tur betyr at alternativet er sant. Alternativet i dette eksemplet er at minst to av reseptene er forskjellige.





Som «lite sannsynlig» er det vanlig å bruke grenseverdien 0,05 (5%), som vi kaller nivået for testen, eller signifikansnivået og ofte angir med p . Av historiske grunner bruker vi oftest $p = 0,01$, $p = 0,05$ eller $p = 0,10$. Det som avgjør om en nullhypotese kan forkastes er altså ikke bare et spørsmål om hvor store forskjeller vi har observert, men også om hvilket nivå vi ønsker å teste på.

Når vi først velger oss et nivå for en test, så betyr det at dette er den største sannsynligheten vi er villig til å akseptere for at vi forkaster en hypotese som er sann. Dette kalles ofte for «feil av type I», eller i engelsk litteratur: «Type I Error». Vi kan aldri være 100% sikker på at den konklusjonen vi trekker er sann. Det er selvsagt ønskelig at sannsynligheten for å trekke en feil konklusjon er så liten som mulig, men gjør vi nivået mindre, så betyr det også at vi må finne større forskjeller i middelverdiene før vi kan forkaste H_0 .

For å vurdere om de forskjellene vi har observert ovenfor virkelig gir grunnlag for å forkaste H_0 , må vi gjøre ytterligere endel antakelser. Slik vi har formulert situasjonen kan det kanskje være naturlig å bruke en Variansanalyse. Da forutsetter vi bl.a. at data er på en intervall- eller ratioskala (se kapittel 4.1). Uansett hvilken statistisk modell som velges, så kan vi skjematisk sette opp gangen i en hypoteseprøving slik:

1. Nullhypotesen og alternativet formuleres.
2. Signifikansnivået bestemmes.
3. Utsagnet i nullhypotesen antas å være sant.
4. Forsøket utføres.
5. Resultatene noteres, og de aktuelle beregningene gjøres.
6. Sannsynligheten for å få minst et så avvikende resultat som det vi fikk, beregnes.
7. Hvis sannsynligheten i punkt 6 er mindre enn det signifikansnivået vi bestemte i punkt 2, er det grunn til å betvile påstanden i punkt 3. Følgelig forkaster vi påstanden i punkt 1 - vi forkaster H_0 - og påstår at alternativet er sant





11.3 Forskjellstester

Forskjellstester kan deles inn i to grupper, generelle forskjellstester og spesifikke forskjellstester (Se kapittel 4.2).

Forskjellstester basert på den binomiske fordeling har den store fordel at det er (relativt) enkle regneoperasjoner og enkel statistisk teori som ligger bak. Men alt har sin pris: disse enkle testene kan bare gi svar på enkle problemstillinger, slik som:

- Er det forskjell mellom prøve A og prøve B?
- Er prøve C søtere enn prøve D?
- Hvilken prøve foretrekker du: E eller F?

Siden et sensorisk panel ikke benyttes til å uttale seg om preferanser, er den siste problemstillinga bare av interesse i forbindelse med forbrukertester.

The Lady tasting Tea

Et eksempel av statistisk-historisk interesse er «The Lady tasting Tea», første gang referert til av R A Fisher i hans bok «The Design of Experiments» som utkom i 1935. En dame påstår hun kan kjenne på smaken hvorvidt teen eller melka ble helt opp i koppen først. Fisher diskuterer så hvordan en slik påstand skal kunne testes. Det kan jo tenkes at damen bløffer, så et eksperiment er absolutt på sin plass. I boka gir hele episoden inntrykk av å være oppkonstruert, men i en biografi publisert mye seinere, påstås det at eksemplet har sitt utspring i en virkelig hendelse. Fisher - gentleman som han var - ville være høflig og tilbød i en tepause en kvinnelig medarbeider en kopp te. «Men kjære Dem, Hr. Fisher, denne kan jeg jo ikke drikke, De har jo hatt i melken først!», svarte damen forferdet. Fisher ristet på hodet og påsto at dette kunne jo ikke damen kjenne forskjell på. Myten vil så ha det til at han gikk inn i sitt lønnekammer og skrev ned innledningskapitlet til en av de store klassikerne innen statistisk litteratur.

Damen med teen lever videre – i 2001 ga David Salsburg ut boka «The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century» – en bok om statistikk og personene som la grunnlaget for statistisk teori. Mange år seinere ble for øvrig damen som ble utsatt for Fisher's lett tilslørte sjekketriks navngitt som Muriel Bristol, forlovet, og seinere gift, med en annen deltaker rundt te-bordet: William Roach. Sjekketriksset var mislykket, men den moderne statistikk var vinneren. Og frøken Roach besto testen: hun kunne virkelig kjenne forskjell!





Et vanlig sensorisk panel består som regel av så få personer at en forskjellstest blir ytterst ustabil: at bare en dommer endrer oppfatning kan snu om på hele konklusjonen. Forskjellstester brukes derfor mer i forbindelse med forbrukertester. Vanligvis vil det delta 100–200 personer i slike tester, og den n som vi seinere skal se inngår i de generelle formlene er lik antall personer som deltar i testen.

Nå kunne det vært fristende å oppnå en høyere verdi av n ved å la panelet gjøre bedømmelsen flere ganger, f.eks. la et panel på 12 personer gjøre en binomisk test 5 ganger for å oppnå $n = 60$. Denne framgangsmåten frarådes, da det får konsekvenser for den teoretiske vurderingen av testens godhet. Derfor unngår man en del teoretiske komplikasjoner ved å holde seg til prinsippet om at binomiske gjøres enten på et panel ved at hver dommer gjør en bedømmelse ($n = \text{antall dommere}$), eller ved at én person gjør n bedømmelser. I det siste tilfellet er det personen som sådan som vi tester.

Et alternativ kan være at vi sier en dommer har rett hvis vedkommende har identifisert rett prøve i for eksempel minst 3 av 4 gjentak. Da vil n i alle formlene fremdeles være lik antall dommere, men sannsynligheten for å gjette riktig vil avhenge av hvordan vi definerer å ha «rett». For eksempel vil en triangeltest hvor vi forlanger minst 3 av 4 rette gi oss en binomisk test med $p = 1/9 = 0,1111$. Hvis vi i stedet vil forlange minst 4 av 5 rette, blir $p = 11/243 = 0,0453$.

11.3.1 Duo-trio-test

Denne testen brukes for å gi svar på spørsmålet: – Er det forskjell mellom prøve A og prøve B? (Se kapittel 4.2. Forskjellstester)

Hvis prøvene virkelig er forskjellige, vil vi forvente at mange dommere vil gjenkjenne den prøven som er ulik referansen. Hvis derimot prøvene er omtrent like, vil vi forvente at bare omtrent halvparten gjør det. Er det ingen forskjell mellom prøvene, må dommerne gjette seg til hvilken prøve som er ulik referansen, og sannsynligheten for å gjette riktig er $p = 1/2$.





Nullhypotesen er at A og B er like, alternativet er at de er forskjellige, det vil si at det er flere dommere som identifiserer riktig prøve enn det man skulle forvente ved ren gjetning. Følgelig blir:

$$H_0: p = \frac{1}{2} \text{ mot } p > \frac{1}{2}$$

Hvis 7 av 12 personer har identifisert korrekt, må vi beregne sannsynligheten for at vi får et slikt – eller mer ekstremt – resultat (punkt 6 i den skjematiske framstillinga av gangen i en hypoteseprøving). Vi må altså beregne sannsynligheten for at minst 7 av 12 personer skal velge A, under forutsetning av at A og B er like. Denne sannsynligheten kan vi finne ved hjelp av den binomiske fordelinga beskrevet i forrige kapittel:

$$\Pr(X \geq 7) = \sum_{i=7}^{12} \binom{12}{i} \left(\frac{1}{2}\right)^i \left(1 - \frac{1}{2}\right)^{12-i} = \sum_{i=7}^{12} \binom{12}{i} \left(\frac{1}{2}\right)^{12} = 0,3872$$

I praksis vil vi aldri beregne en slik sannsynlighet for hånd, men enten bruke et regneark eller et statistikkprogram. I Excel vil vi få fram denne sannsynligheten ved å skrive

=1-BINOM.FORDELING.N(6;12;0,5;SANN) eller:

=1-BINOM.FORDELING(6;12;0,5;SANN)

Da har vi benyttet oss av at $\Pr(\bar{A}) = 1 - \Pr(A)$ for en vilkårlig begivenhet A. Eller med andre ord: sannsynligheten for en begivenhet (her 7 eller flere) er lik 1 minus sannsynligheten for den omvendte sannsynligheten (her: færre enn 7). Når begivenheten A betyr $X \geq 7$, blir dette: $\Pr(X \geq 7) = 1 - \Pr(X < 7) = 1 - \Pr(X \leq 6)$. Legg merke til at det motsatte av $X \geq 7$ altså er $X < 7$, som igjen er lik $X \leq 6$ fordi X nødvendigvis er et helt tall.

I statistikkprogrammet R (som er gratis), får vi den samme sannsynligheten ved hjelp av:

sum(dbinom(7:12,12,0.5))

En sannsynlighet på 0,3872 kan definitivt ikke betraktes som «liten»





her. Det betyr at selv om 7 av 12 (altså et simpelt flertall, eller 58%) korrekt identifiserer A til å være lik referansen, så kan ikke det kalles usannsynlig. Det er derfor ingen grunn til å forkaste H_0 , og vi har altså ingen grunn til å påstå at det er noen forskjell mellom de to prøvene.

Hvis vi hadde benyttet langt flere enn 12 personer, for eksempel 150, så ville det faktisk at minst 58% (dvs minst 87 personer) av disse hadde valgt prøve A, hatt en sannsynlighet på bare 0,030 hvis hypotesen om at de egentlig var like, var sann.

Et alternativ til å gjøre disse beregningene er å slå opp i statistiske tabeller for å finne ut hvor mange korrekte identifikasjoner man må ha for å kunne forkaste H_0 – etter først å ha bestemt seg for nivået for testen. Slike tabeller pleier – selv om statistikkprogrammene stort sett har overtatt alt regnearbeidet – å være tilgjengelige i elementære lærebøker i statistikk.

Visste du at...?

I sannsynlighetsregning er det vanlig å snakke om begivenheter, suksesser og gunstige utfall. For andre enn statistikere kan det virke rart at man for eksempel i forbindelse med dødsfall snakker om «suksesser» for å omtale personer som døde i et medisinsk forsøk. Vi får bare akseptere at «sånn er det».....

11.3.2 Triangeltest

I likhet med duo-trio-testen har vi også her to prøver som skal sammenlignes, og hver av forsøkspersonene får tre testprøver. To av testprøvene er like, mens den tredje er forskjellig fra de to første. Siden det er 3 testprøver som blir presentert, er det nå en sannsynlighet på $1/3$ for å gjette riktig.

Nullhypotesen blir:

$$H_0: p = \frac{1}{2} \text{ mot } p > \frac{1}{2}$$

I likhet med under duo-trio-testen, vet altså forsøkslederen hva som er riktig: fasiten er kjent. For å avgjøre om det er signifikant forskjell





mellom de to prøvene må vi beregne sannsynligheten for at det resultat vi observerte kan oppnås ved tilfeldigheter. Sagt på en annen måte: hvis f. eks. 23 av 45 personer korrekt identifiserte den testprøven som var forskjellig fra de to andre – hvor stor sannsynlighet er det for et slikt resultat hvis de to prøvene i virkeligheten er like? Denne gang blir $p=1/3$ i formelen, og vi får:

$$\Pr(X \geq 23) = \sum_{i=23}^{45} \binom{45}{i} \left(\frac{1}{3}\right)^i \left(\frac{2}{3}\right)^{45-i}$$

som vi ved hjelp av Excel, R eller et annet regneark- eller statistikkprogram finner er lik 0,0103.

Følgelig er det en liten sannsynlighet for at 23 av 45 personer skal utpeke den ene prøven som forskjellig fra referansen hvis det i virkeligheten ikke er noen forskjell. Siden den sannsynligheten vi fant er liten, er det naturlig å forkaste hypotesen om at prøvene er like.

Samme konklusjon ville vi ha fått ved å konsultere tabellen i vedlegg 16b. Der framgår det at med 45 personer er det tilstrekkelig med 21 korrekte svar for å forkaste hypotesen om at prøvene er like hvis vi ønsker nivå 0,05 for testen.

11.3.3 Partest

Ved en partest får dommeren to prøver som skal sammenlignes, for eksempel hvilken av de to som er søttest, har mest gulrotssmak, eller hvilken av de to prøvene synes du smaker best? I det etterfølgende holder vi oss til eksempel med søtssmak. En naturlig nullhypotese vil være at de to prøvene er like søte.

Partestene kan ha ensidige eller tosidige alternativ. Hvis vi bare er interessert i å vite om det er en forskjell, men ikke har noen forutinntatt mening om hvilken vei forskjellen kan gå, har vi en tosidig test: konklusjonen kan bli at A er mest søt, eller at B er mest søt.





Da sier vi at testen er tosidig, med symboler skriver vi nullhypotesen og alternativet slik:

$$H_0: A=B \text{ mot } A \neq B$$

Men i noen situasjoner kan det tenkes at bare det ene alternativet er av interesse. En slik situasjon kan være at vi vurderer å innføre en ny, billigere ingrediens i et produkt. Denne ingrediensen kan tenkes å ha innflytelse på søtheten. Hvis søt smak er en ønsket egenskap, ønsker vi å avsløre det hvis den nye resepten resulterer i mindre søt smak enn den gamle. I så fall vil vi ikke introdusere den billigere ingrediensen. Om den nye ingrediensen gir enda søtere smak enn den gamle er egentlig ikke så interessant, det blir nærmest bare en bonus. Det er sannsynlighetsteoretiske grunner til at vi i en slik situasjon velger en ensidig test. Det er nemlig slik at det nivået vi velger for testen, er en form for forsikring mot å gjøre feil. Ved en tosidig test kan vi komme til å påstå at A er søtere enn B når faktisk H_0 er sann, og vi kan komme til å påstå at B er søtere enn A når faktisk H_0 er sann. Siden vi ikke er interessert i alternativet B er søtere enn A, er det heller ingen grunn til å «bruke opp» noe av nivået for å gardere oss mot en slik «feil». Altså er testen ensidig, med symboler skriver vi nullhypotesen og alternativet slik:

$$H_0: A=B \text{ mot } A > B$$

Alternativet er altså at A er søtere enn B. Hvis prøvene virkelig er like, vil dommernes valg være bestemt av rene tilfeldigheter, og vi vil forvente at omtrent like mange peker ut den ene som den andre prøven. Selve den statistiske testen består i å notere hvor mange som har valgt prøve A og hvor mange som har valgt prøve B, og sjekke hvor sannsynlig dette er *under forutsetning av at prøvene faktisk er like*.

Sannsynligheten for å velge den ene eller den andre prøven, forutsatt at nullhypotesen er sann (dvs. at prøvene er like), er lik $\frac{1}{2}$. Hvis 7 av 12 personer har sagt at italiensk salat produsert etter resept A er best, mens 5 har foretrukket resept B, må vi beregne sannsynligheten for at vi får et slikt – eller mer ekstremt – resultat (punkt 6 i den skjematiske framstillinga av gangen i en hypoteseprøving). Vi må altså beregne sannsynligheten for at minst 7 av 12 skal velge A, under forutsetning





av at A og B er like. Denne sannsynligheten finner vi – akkurat på samme måte som for duo-trio-testen ved hjelp av:

$$\Pr(X \geq 7) = \sum_{i=7}^{12} \binom{12}{i} \left(\frac{1}{2}\right)^i \left(1 - \frac{1}{2}\right)^{12-i} = \sum_{i=7}^{12} \binom{12}{i} \left(\frac{1}{2}\right)^{12} = 0,3872$$

Konklusjonen blir også den samme: det er ingen grunn til å påstå at A er søtere enn B.

Et eksempel på en par-test hvor det er aktuelt med et tosidig alternativ, er følgende: vi har valget mellom å tilsette 2 forskjellige ingredienser og er interessert i hvilken som gir det søtteste produktet. Men på forhånd vet vi ingenting om hvilken av de 2 som vil gi det søtteste produktet. Det betyr at vi forkaster H_0 om at de er like, hvis *mange* dommere sier at A er søttest, eller at mange sier at B er søttest. Eller sagt på en annen måte: Vi forkaster H_0 hvis mange svarer A, eller få svarer A. Dette følger av at dommerne må svare noe, slik at antall som svarer A pluss antall som svarer B er lik antall dommere.

Anta at en partest med 15 dommere, så har 12 påstått at A er søttest, men 3 har påstått at B er søttest. Hvor sannsynlig er det at et slikt resultat skal oppstå av rene tilfeldigheter, det vil si hvis H_0 faktisk er sann? Før vi regner ut svaret, kan vi selvfølgelig sjekke med tabellen i vedlegg 16b og konkludere med at sannsynligheten for dette er mindre enn 0,05. Vil vi sjekke det, blir formelen:

$$\begin{aligned} \Pr(X \geq 12) + \Pr(X \leq 2) &= \sum_{i=12}^{15} \binom{15}{i} \left(\frac{1}{2}\right)^i \left(1 - \frac{1}{2}\right)^{15-i} + \sum_{i=0}^2 \binom{15}{i} \left(\frac{1}{2}\right)^i \left(1 - \frac{1}{2}\right)^{15-i} \\ &= \sum_{i=12}^{15} \binom{15}{i} \left(\frac{1}{2}\right)^{15} + \sum_{i=0}^2 \binom{15}{i} \left(\frac{1}{2}\right)^{15} = 0,0213 \end{aligned}$$

I tillegg til at dette bekrefter konklusjonen vi fikk ved å sjekke tabellen, får vi den tilleggsinformasjonen at vi ligger relativt godt under nivået 0,05 – konklusjonen burde derfor være klar.





11.3.4 To-av-5-test

Denne testen er en variant av triangeltesten og har fått en viss utbredelse, trolig fordi sannsynligheten for å gjette riktig hvis det ikke er noen forskjell mellom prøvene, er så liten som 0,1. Igjen er vi interessert i å sammenlikne 2 prøver, men nå får dommerne 2 testprøver av den ene prøven og 3 av den andre. Oppgaven er å finne hvilke 3 som hører sammen.

I denne testen vil vi derfor igjen bruke den binomiske fordelinga, nå med $p=0,1$ der vi brukte $p = 1/3$ i triangeltesten. Med 16 dommere og korrekt identifikasjon fra 5 av dem, får vi at

$$\Pr(X \geq 5) = \sum_{i=5}^{16} \binom{16}{i} 0,1^i (1-0,1)^{16-i} = 0,0170$$

som må regnes som en «liten» sannsynlighet, altså vil vi forkaste H_0 : $A = B$ og påstå at de er forskjellige.

I prinsippet kunne man konstruere massevis av slike n-av-m-tester. Sannsynligheten p for å gjette riktig vil da være gitt ved 1 dividert med den binomiske koeffisienten:

$$p = \frac{1}{\binom{m}{n}} = \frac{1}{\frac{m!}{(m-n)!n!}}$$

En 2-av-4-test ($n = 2, m = 4$) vil ha $p = 1/6 = 0,1667$, en 4-av-9-test ($n = 4, m = 9$) vil ha $p = 1/126 = 0,0079$.

11.4. Beskrivende tester

I en beskrivende test bedømmer dommerne produktene etter en eller flere egenskaper ved å gi dem karakterer på en skala. Vanlige skalaer er 1-7, 1-9, 0-100, ofte med muligheten til å bruke desimaler.





11.4.1 Enveis variansanalyse – urealistisk modell

En statistisk test som ofte brukes til å analysere data fra beskrivende test er slike data, er variansanalyse, ofte omtalt som ANOVA etter den engelske betegnelsen *Analysis of Variance*. En variansanalyse kan være alt fra en helt enkel og oversiktlig situasjon til en komplisert modell som kan være vanskelig å fortolke og som i praksis ville vært nær umulig å beregne manuelt.

En helt enkel situasjonen kan eksemplifiseres ved det følgende urealistiske datasettet, hvor 16 dommere har vært i aksjon og bedømt søtthet: 4 av dem har bedømt sort A, 4 har bedømt sort B, 4 har bedømt sort C og 4 har bedømt sort D. Her kan det ikke understrekes tydelig nok at dette ikke er en realistisk situasjon i sensorikk: der lar vi – nesten uten unntak – alle dommerne bedømme alle sortene. *Men denne modellen lar oss demonstrere gangen i den aller enkleste formen for variansanalyse.*

Lesere som kjenner til begrepet T-test kan her bli fristet til å benytte den flere ganger: først sammenlikne A med B, så A med C, A med D, B med C, B med D og C med D, totalt 6 T-tester. Dette er en framgangsmåte som det sterkt advares mot, fordi vi da mister kontrollen over nivået for hele testsituasjonen. Selv om nivået for hver av testene er for eksempel 0,05, så gjelder dette ikke hvis vi ser alle testene under ett, det vil si å teste $H_0: A = B = C = D$.

Grunntanken i en variansanalyse er – ikke overraskende – å sammenlikne varianser. Det som derimot er overraskende, er at det vi egentlig uttaler oss om, er middelerverdier. Idéen er ikke umiddelbart opplagt og krever en viss matematisk skoleing for å kunne forstås helt ut. La oss anta at resultatene ble som i tabell 11.1:

Tabell 11.1: Middelerverdier for 4 sorter

SORT A	SORT B	SORT C	SORT D
3	4	5	8
6	1	5	5
4	2	4	8
7	5	6	7





Middelverdiene for de 4 sortene blir 5, 3, 5 og 7. Variansen i datamaterialet kan beregnes på to måter, forutsatt at det ikke er noen systematisk forskjell mellom sortene. En måte å regne på, er å beregne variansen innen hver av de 4 sortene, og så ta gjennomsnittet av de 4 variansene. Avrundet til 2 desimaler blir de 4 variansene 3,33 – 3,33 – 0,67 og 2,00, og gjennomsnittet av dem blir 2,33.

Men hvis det ikke er noen forskjell mellom sortene, så vil variansen for de 4 gjennomsnittene også være et fornuftig uttrykk for variansen i datasettet. Denne verdien blir 2,67. Når variansen for et sett observasjoner er σ^2 , så er variansen for middelverdien over n slike observasjoner lik σ^2/n . For at uttrykket ovenfor (2,67) skal være sammenlignbart med den gjennomsnittlige variansen vi fant (2,33), må vi derfor multiplisere den førstnevnte med antall observasjoner som ligger bak hvert middeltall, nemlig 4, og vi får 10,67. (Her har vi brukt de ikke-avrundete delresultatene i beregningene). Siden vi har to tall som begge er et uttrykk for variansen, forutsatt at det ikke er noen forskjell mellom sortene, er det naturlig å sammenlikne dem, for eksempel ved å dividere dem med hverandre:

$$F = \frac{10,67}{2,33} = 4,57$$

Under forutsetning av at H_0 var sann, det vil si at det bare er tilfeldige variasjoner som gjør at ikke alle 4 middelverdiene blir identiske, så ville vi forventet at F var nær 1. Når den er større enn 1, tyder det på at det estimatet for totalvariansen i datasettet som vi beregnet, inneholder noe mer enn bare den tilfeldige variasjonen mellom middelverdiene. Eller med andre ord: de 4 middelverdiene kan ikke betraktes som å være like allikevel. Telleren i F består altså av noe mer enn variasjonen i datasettet og dette «noe» kan direkte tilskrives forskjellene i middelverdier.

Hvor sannsynlig er det at det skal bli så store forskjeller som det har blitt, forutsatt at H_0 er sann? Det kan vi få ut ved hjelp av et av de mange statistikkprogrammene som er tilgjengelige (vi har igjen benyttet Statistix 9):



**One-Way AOV for X by Sort**

Source	DF	SS	MS	F	P
Sort	3	32.0000	10.6667	4.57	0.0234
Error	12	28.0000	2.3333		
Total	15	60.0000			

Her kjenner vi igjen verdiene 10,67 og 2,33 fra den manuelle beregningen av variansene, samt F-verdien 4,57. Programmet har riktignok skrevet ut variansene med 4 desimaler i stedet for våre 2. Sannsynligheten finner vi under overskriften P: den er 0,0234 og er ikke noe vi kan regne ut selv. For å teste H_0 på nivå 0,05 uten tilgang til program av denne typen måtte vi ha funnet ut at 0,95-fraktilen i F-fordelinga med 3 og 12 frihetsgrader var 3,4903. Siden den F vi beregnet er større enn den verdien vi fant i tabellen, må vi altså forkaste H_0 .

Formlene vi har benyttet, skrevet ut i en mer formell drakt, ser slik ut (lesere med formelfobi gis herved tillatelse til å hoppe over dette avsnittet):

Variansen innen gruppe nr. i:

$$\frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_{i\cdot})^2$$

Middelverdien over disse variansene:

$$Q_0 = \frac{1}{m} \sum_{i=1}^m \frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_{i\cdot})^2$$

Variansen for gruppemiddeltallene blir:

$$\frac{1}{m-1} \sum_{i=1}^m (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2$$





Etter multiplikasjon med n :

$$Q_1 = \frac{n}{m-1} \sum_{i=1}^m (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2$$

$$F = \frac{Q_1}{Q_0}$$

Beviset for at F er Fisher-fordelt med $m-1$ og $m(n-1)$ frihetsgrader når H_0 er sann, forbigår vi denne sammenhengen, men henviser til lærebøker i generell statistikk.

Selv om en variansanalyse kan ses på som en komplisert prosedyre rent beregnings-teknisk sett, eksisterer det altså en helt mekanisk løsning av problemet:

1. Mat dataene inn i programmet.
2. Trykk på de rette knappene(!).
3. Sjekk om p -verdien er mindre enn det nivået man vil teste på.

11.4.2 Toveis variansanalyse uten gjentak

Vi vil nå se på en mer realistisk sensorisk situasjon, hvor vi lar et panel på 9 dommere bedømme bitterhet i 3 forskjellige kaffesorter, for enkelthets skyld kalt A, B og C. Bedømmelsene ga følgende resultat:

Tabell 11.2: Dommerbedømmelser og middelveidier for 3 sorter

DOMMER	SORT A	SORT B	SORT C
1	4,6	3,8	8,2
2	4,2	7,1	8,1
3	5,9	5,2	8,4
4	6,0	3,2	7,6
5	5,5	3,1	5,1
6	4,7	3,6	7,3
7	5,0	3,9	7,7
8	4,8	4,1	8,0
9	4,3	6,5	7,1
Middel:	5,0	4,5	7,5





Som andre målinger (kjemiske, fysiske, ...) er også disse sensoriske målingene forbundet med en usikkerhet. Selve «måleinstrumentet», nemlig dommerne i panelet, varierer; og det kan de gjøre både på en systematisk og en tilfeldig måte. Videre kan det være variasjon i prøve-materialet: selv om vi tilbereder kaffen på samme måte hver gang, kan det være råvareforskjeller også innen en enkelt sort. Vi kan derfor ikke uten videre konkludere med at kaffesort C er mer bitter enn de to andre. Spørsmålet vi må svare på, er om de forskjellene vi observerer, er signifikante. Sagt på en annen måte: kan hypotesen $H_0: A = B = C$ forkastes på et gitt nivå, f.eks. 0,05?

Selv om tabellen over i sin struktur likner tabellen i det forrige delkapitlet til forveksling, er det en viktig forskjell: i kaffe-eksemplet er det de samme 9 dommerne som har bedømt alle sortene, slik at vi har ikke bare en sorteffekt vi må ta hensyn til i modellen, men også en dommereffekt.

Kortversjonen av hvordan vi skal analysere disse kaffedataene, er som følger:

1. Mat dataene inn i programmet.
2. Trykk på de rette knappene(!).
3. Sjekk om p-verdien er mindre enn det nivået man vil teste på.

Statistix 9 gir oss:

Analysis of Variance Table for SensScore

Source	DF	SS	MS	F	P
Sort	2	46.5000	23.2500	19.14	0.0001
Dommer	8	8.6800	1.0850	0.89	0.5441
Error	16	19.4400	1.2150		
Total	26				



Middelverdiene og standardavvik kan vi også få:

Descriptive Statistics for Sort = A

Variable	N	Mean	SD
SensScore	9	5.0000	0.6595

Descriptive Statistics for Sort = B

Variable	N	Mean	SD
SensScore	9	4.5000	1.4457

Descriptive Statistics for Sort = C

Variable	N	Mean	SD
SensScore	9	7.5000	0.9950

Et hvilket som helst statistikkprogram vil gi oss den samme informasjonen, men den kan være organisert på en annen måte. Et gratis-program som R, her i versjon 3.1.1 og med modulene («package» i R-terminologi) *mixlm* og *describe* gir følgende utskrift, lettere redigert av plasshensyn:

Analysis of variance (unrestricted model)

Response: SensScore

	Mean Sq	Sum Sq	Df	F value	Pr(>F)
Sort	23.25	46.50	2	19.14	0.0001
Dommer	1.09	8.68	8	0.89	0.5441
Residuals	1.21	19.44	16		

Group A

	n	mean	sd
SensScore	9	5.00	0.66

Group B

	n	mean	sd
SensScore	9	4.50	1.45

Group C

	n	mean	sd
SensScore	9	7.50	0.99



I praksis har altså data fra en beskrivende test minst en toveis-analyse, og hvis de forskjellige prøvene kan deles inn i flere undergrupper eller forsøksledd (for eksempel: sorter, lagringsforhold, dyrkingssted ...) kan det være snakk om tre-, fire-, og flerveisanalyser. Selv om dette kompliserer formlene og gjør en manuell utregning nesten uoverkommelig, er prinsippet grovt sett det samme uansett hvor komplisert modellen er: totalvariansen deles inn i delvarianser som kan tilskrives de forskjellige forsøksleddene som inngår i modellen, samt at det alltid vil være igjen en rest – i utskrifter fra dataprogrammene gjerne betegnet Error. En annen betegnelse som benyttes på denne restvariansen, er MSE (Mean Square Error) – et begrep som vil dukke opp seinere under beskrivelsen av programvaren PanelCheck.

11.4.3 Rangeringer: Friedman's test

I stedet for å bruke bedømmelsene fra dommerne direkte, kan vi gjøre dem om til rangeringer. Da er det naturlig å analysere data ved hjelp av Friedman's test. Den tar utgangspunkt i rangeringer, for eksempel ser vi av tabellen i kapittel 11.2 at dommer 5 har rangert prøvene i rekkefølgen A – C – B, hvor A har fått høyest score. Da vil resultatene for dommer 5 for sort A kodes om til 3, sort C til 2 og sort B til 1. Tilsvarende blir gjort for de andre dommerne. Tabellen omgjort til rangeringer blir da:

Tabell 11.3: Rangering av 3 sorter, rangsum og gjennomsnittlig rang

DOMMER	SORT A	SORT B	SORT C
1	2	1	3
2	1	2	3
3	2	1	3
4	2	1	3
5	3	1	2
6	2	1	3
7	2	1	3
8	2	1	3
9	1	2	3
Rangsum	17	11	26
Gj.snittlig rang	1,89	1,22	2,89





Hvis sortene er like, vil vi forvente at rangeringene fordeler seg slik at A, B og C får omtrent samme rangsum (= summen av rangeringene). Spørsmålet i eksemplet blir nå om rangsummene 17, 11, 26 er så forskjellige at vi må forkaste hypotesen om at de er like, eller at forskjellene i rangsummer like gjerne skyldes rene tilfeldigheter. Å regne ut sannsynligheten for å få så store (eller større) forskjeller er ingen enkel oppgave. Vi støtter oss her til statistisk teori, nærmere bestemt Friedman's formel, som sier at T definert under er tilnærmet kji -kvadratfordelt:

$$T = \frac{12}{ds(s+1)} \sum_{i=1}^s R_i^2 - 3d(s+1)$$

der d = antall dommere (= 9 i eksemplet over), s = antall sorter (= 3), og R_i ($i=1,2, \dots, s$) er rangsummene (17, 11 og 26 i tabellen over). Da blir:

$$T = \frac{12}{9 \times 3 \times (3+1)} (17^2 + 11^2 + 26^2) - 3 \times 9 \times (3+1) = 12,67$$

Siden T under H_0 er tilnærmet kji -kvadratfordelt med $s - 1$ frihetsgrader (= 2 i dette eksemplet), må vi sammenlikne $T = 12,67$ med 0,95-fraktilen i denne fordelinga. Av statistiske tabeller (vedlegg 16 b, tabell 3) ser vi at den verdien er lik 5,9915. Siden $T > 5,9915$, må derfor H_0 forkastes. I praksis vil vi vanligvis la et statistikkprogram stå for selve utregningene, og da kan resultatet bli seende slik ut (her har vi benyttet Statistix 9)

Friedman Two-Way Nonparametric AOV

Variable	Mean Rank	Sample Size
Sort A	1.89	9
Sort B	1.22	9
Sort C	2.89	9

Friedman Statistic	12.667
P-value, Chi-Squared Approximation	0.0018
Degrees of Freedom	2





I stedet for å sammenlikne T (*Friedman Statistic* i utskriften fra programmet) med verdien fra tabellen, sammenlikner vi p -verdien (0,0018) i utskriften med nivået vi tester på (0,05). Siden den beregnede p -verdien i utskriften er mye mindre enn nivået, forkaster vi H_0 . Konklusjonen blir altså den samme, men bruk av statistikkprogrammet vil i større grad kunne si oss hvor signifikant forskjellene er: det er forskjell på $p = 0,0003$ og $p = 0,0499$ selv om begge sier at det er forskjell på nivå 0,05.

Det programmet som er benyttet her, skriver ut de gjennomsnittlige rangeringene i stedet for rangsummene 17, 11 og 26; dette vil variere fra program til program.

Hvis en dommer har bedømt 2 eller flere sorter likt, deles rangeringene: hvis to prøver har fått 6,0 og en prøve 4,7 så får de to første hver rang 2,5 (= gjennomsnittet av 3 og 2, som de ville fått hvis de i stedet hadde hatt for eksempel verdiene 6,1 og 6,0) og den tredje får rang 1.

Friedman's test kan benyttes enten dommerne har bedømt egenskapene langs en skala, eller de har gjort rangeringene direkte. Testen er et interessant alternativ hvis man ikke er overbevist om at data tilfredsstiller de kravene som en variansanalyse stiller. Hvis dommerne har lite trening i å bruke en skala, for eksempel hvis det dreier seg om forbrukere, kan det være en ide å bruke Friedman's test til analysene selv om data opprinnelig er på en 1-9 skala.

11.4.4 Multiple sammenlikninger

Å forkaste H_0 er vanligvis ikke en fullstendig konklusjon. Hvis vi har flere enn 2 grupper, vil vi gjerne vite hvilke grupper som er forskjellige. I eksemplet fra kapitlet over kan vi stille spørsmål som:

Er alle 3 sortene forskjellige fra hverandre?

Er sort 1 og sort 2 forskjellige fra sort 3, mens vi ikke kan skille sort 1 fra sort 2?

Er sort 2 forskjellig fra sort 3, mens vi ikke kan skille sort 1 fra sort 2 eller sort 1 fra sort 3?





Her er det nærliggende å benytte T-tester for hver av sammenlikningene:

Sort 1 mot sort 2

Sort 1 mot sort 3

Sort 2 mot sort 3

Den teknikken må det advares sterk mot av så vel sannsynlighetsteoretiske som praktiske grunner. Den korrekte måten å besvare slike spørsmål på, er å benytte metoder som går under betegnelsen multiple sammenlikninger. Av dem finnes det flere, og den vanligste er sannsynligvis Tukey's test. Også den vil ligge inne i alle statistikkprogrammer med respekt for seg selv. En grei måte å presentere slike resultater på, er denne, hentet fra Statistix 9:

Tukey HSD All-Pairwise Comparisons Test of SensScore for Sort

Sort	Mean	Homogeneous Groups
Sort C	7.5000	A
Sort A	5.0000	B
Sort B	4.5000	B

Alpha	0.05	Standard Error for Comparison	0.5196
Critical Q Value	3.651	Critical Value for Comparison	1.3413

Error term used: Sort*Dommer, 16 DF

There are 2 groups (A and B) in which the means are not significantly different from one another.

Hva kan vi så lese ut av dette? Prinsippet er at sorter som ikke er signifikant forskjellige fra hverandre, markeres med samme bokstav i kolonnen Homogeneous Groups. Hovedkonklusjonen er at det er signifikant forskjell mellom A og C og mellom B og C, men ikke mellom A og B. Alpha er nivået for testen, og Critical Value for Comparison er den verdien som middelverdiene sammenliknes med. Forskjellen mellom A og B er 0,5 som er mindre enn 1,3413: altså er det ikke forskjell mellom A og B. Men forskjellene mellom A og C (2,5) og mellom B og





C (3,0) er begge større enn 1,3413 – følgelig er både A og B forskjellige fra C.

I større forsøk kan resultatet av multiple sammenlikninger bli ganske komplisert, som i dette forsøket med 8 prøver:

	Mean	G1	G2	G3	G4
Prøve JK407	4.104762	A			
Prøve JO171	3.861905	A			
Prøve BG891	3.566667	A	B		
Prøve BY100	3.366667	A	B		
Prøve CX301	3.161905	A	B	C	
Prøve EE291	2.119048		B	C	D
Referanse	1.685714			C	D
Prøve LK197	1.461905				D

Skal vi beskrive dette med ord, blir det:

JK407 og JO171 er forskjellige fra EE291, Referanse og LK197
 BG891 og BY100 er forskjellige fra Referanse og LK197
 CX301 er forskjellig fra LK197

Mellom de øvrige prøvene er det ingen forskjeller. Merk at her er forskjellig fra benyttet i betydningen signifikant forskjellig på nivå 0,05.

I en type situasjoner kan det være aktuelt med en annen test, nemlig hvis vi ikke er interessert i å sammenlikne et visst antall sorter som i utgangspunktet er likeverdige, men ønsker å sammenlikne disse sortene mot en standard eller kontroll. Med 4 sorter A, B, C, D og en kontroll K, er vi altså ikke interessert i å sammenlikne alle 5 sortene innbyrdes, men bare de 4 første mot den siste. I en viss forstand slipper vi å «kaste bort» noe av nivået for testen ved å gardere oss mot feilaktig å påstå at det er forskjeller innen A, B, C og D. Denne testen kalles Dunnett's test og er tilgjengelig i alle standard statistikkprogram. I det ovenstående eksemplet med 8 prøver kunne Dunnett's test vært et alternativ, siden vi har en prøve som kan betraktes som en kontroll eller en referanse.





Også i Friedman's test er det aktuelt å gjøre multiple sammenlikninger hvis vi har forkastet H_0 . I motsetning til Tukey's test er dette ikke en standardanalyse, og den må derfor gjennomføres for hånd. Første punkt er å beregne den kritiske verdien som alle parene skal sammenliknes med. En navnløs test beskrevet i Hochberg og Tamhane (1987) gir en kritisk verdi c gitt ved:

$$c = \frac{n}{\sqrt{2}} q_{k,\infty;\alpha} \sqrt{\frac{k(k+1)}{6n}}$$

Her er n lik antall dommere, k = antall grupper, og $q_{k,\infty;\alpha}$ er den kritiske verdien for den Studentiserte variasjonsbredde (Studentized range) for k grupper, uendelig mange frihetsgrader og nivå α for testen. Siden den tabellen ikke er vanlig å gjengi i statistiske lærebøker, er en tabell for $k = 3, 4, \dots, 20$ og $\alpha = 0,01, 0,05$ og $0,10$ gjengitt i vedlegget. De gruppene som har rangsummer som avviker med mer enn c , er signifikant forskjellige. Med $q_{k,\infty;\alpha} = 3,315$ fra tabellen i vedlegg 16 får vi:

$$c = \frac{9}{\sqrt{2}} \times 3,315 \times \sqrt{\frac{3 \times 4}{6 \times 9}} = 9,945$$

Siden forskjellen mellom A og C og mellom A og B er mindre enn denne kritiske verdien, mens forskjellen mellom B og C er større enn den, kan konklusjonen framstilles slik:

C	26 ^a
A	17 ^{ab}
B	11 ^b

Altså: det er ikke signifikant forskjell mellom A og C, og heller ikke mellom A og B, mens det er signifikant forskjell mellom B og C. Legg merke til at denne konklusjonen er noe svakere enn den vi kunne trekke i variansanalyseeksemplet: der var både A og B signifikant forskjellige fra C.

11.4.5 Toveis variansanalyse med gjentak

Dette er kanskje den vanligste modellen i sensorisk analyse: vi har et visst antall prøver: A, B, C, osv., og alle dommerne i et sensorisk panel har bedømt alle prøvene flere ganger (som oftest 2 eller 3). Den underliggende metoden for variansanalyse som vi har sett på tidligere, benyttes også her. Et konkret eksempel:



Tabell 11.4: Data fra 3 sorter bedømt av 10 dommere i 2 gjentak

	Sort A	Sort A	Sort B	Sort B	Sort C	Sort C
Dommer	Gjentak 1	Gjentak 2	Gjentak 1	Gjentak 2	Gjentak 1	Gjentak 2
1	3,6	3,9	1,9	2,7	4,0	3,1
2	2,1	1,6	1,0	1,0	2,0	2,8
3	3,9	4,5	1,0	3,7	4,3	4,0
4	1,0	3,0	1,0	1,0	5,0	4,8
5	1,5	6,6	3,1	2,5	6,6	5,1
6	2,5	2,7	2,3	3,7	2,7	3,4
7	3,3	5,6	1,0	1,0	3,6	2,8
8	2,9	2,6	3,6	3,3	2,6	4,1
9	3,8	1,3	1,5	1,1	4,0	1,2
10	3,9	6,1	2,6	1,9	5,0	7,7

I denne modellen oppstår tre nye begreper som vi må ta hensyn til. Det første er relativt greit å forholde seg til: samspillet mellom to faktorer, i dette tilfellet: samspillet mellom sort og dommer. At det er samspill mellom faktorene sort og dommer, betyr i korthet at ikke alle dommerne har bedømt alle sortene likt: dommer 1 kan ha gitt høyere score til prøve A enn til prøve B, mens dommer 2 har gitt høyere score til prøve B enn til prøve A. Aller helst hadde vi jo ønsket at dommerne utviste større enighet enn det. Det er ikke så farlig om de legger seg på forskjellige deler av skalaen, så lenge de er konsekvente. Med konsekvente menes her at de bedømmer prøvene i omtrent samme rekkefølge.

Det andre begrepet er at effekter kan deles inn i de som er tilfeldige (engelsk: random) eller faste (engelsk: fixed). Om en effekt er fast eller tilfeldig, får konsekvenser for beregningen av F-verdien. I de to enkle modellene vi har sett på hittil, får det ingen praktiske konsekvenser om effektene er faste eller tilfeldige: formlene – og dermed alle utregningene – blir akkurat de samme. Men med en gang vi får tilfeldige effekter inn i modellen, endrer det til dels dramatisk noe av det teoretiske grunnlaget for testene.



En fast effekt er en effekt hvor vi er interessert i de nivåene, eller verdiene, som effekten har. I et eksempel med 3 sorter A, B og C er det naturlig at disse 3 sortene er valgt ut fordi det er nettopp disse sortene vi er interessert i. En annen situasjon ville vi fått hvis det var snakk om eple sorter, og vi i stedet for å bestemme oss på forhånd hvilke sorter vi ville undersøke, så gikk vi i en butikk og kjøpte de sortene som tilfeldigvis var til salgs den dagen. En tilfeldig effekt kan bringe tanken hen på at det er nettopp noe tilsvarende vi gjør når en effekt defineres som tilfeldig. Det er imidlertid ikke tilfelle når det gjelder de tilfeldige effekten i herværende modell, nemlig dommereffekten. Når vi definerer den som tilfeldig, betyr det selvfølgelig ikke at det sensoriske panelet består av personer trukket tilfeldig ut fra en befolkningen over 18 år. De aller fleste sensoriske panel er hentet fra personer som er trent til oppgaver innen sensoriske analyser. Men det er likevel tilfeldig at det er nettopp disse personene som utgjør panelet.

Vi holder oss til den tradisjonen som definerer en fast effekt hvor vi er interessert i de konkrete verdiene (sortene) vi analyserer, og alt annet defineres som tilfeldige effekter.

Det tredje av de nye begrepene er skillet mellom kryssete (engelsk: crossed) og nøstet (engelsk: nested), eller hierarkiske effekter. I den modellen vi ser på her, er dommer og sort krysset: alle dommerne har bedømt alle sortene. Hvis vi hadde benyttet totalt 30 dommere, og 10 dommere hadde bedømt sort A, 10 andre dommere hadde bedømt sort B og de 10 siste dommerne hadde bedømt sort C, så hadde dommereffekten vært hierarkisk under sorteffekten. I andre, og mer kompliserte modeller, hvor vi har 3 slaktemetoder og 3 dyr fra hver metode, ville det være naturlig å modellere dyreeffekten som hierarkisk under slaktemetoden. Ett enkelt dyr kan av opplagte grunner bare bli slaktet etter en metode. I en slik situasjon ville ikke dyr 1 fra metode A hatt noe som helst felles med dyr 1 fra metode B og C, noe som vi må ta hensyn til i utregningene. Fôringforsøk er en annen type forsøk hvor det er naturlig å benytte en modell hvor dyr er hierarkisk under et fôringsregime.

Et annet eksempel på hvor dette er aktuelt, er følgende situasjon: vi skal sammenlikne sauser som leveres i poser med pulver, 3 poser fra





hver sort. Da kan vi lage 3 sauser av hver sort og la hver pose være et gjentak, eller blande alle 3 posene og lage en stor porsjon som vi serverer 3 gjentak fra til hver dommer. I det første tilfellet vil vi ha en pose-effekt hierarkisk under sorteffekten: pose 1 fra sort D vil ikke ha noe felles (bortsett fra den tilfeldige nummereringen) med pose 1 fra sort C. En slik modell vil gi oss informasjon om variasjonen mellom enkeltposer, mens den andre lager en form for middelvei ved å blande de 3 posene og dermed miste all informasjon om variasjon mellom enkeltposene.

Gjentakene er vanligvis både en tilfeldig og hierarkisk effekt, i hvert fall i slike situasjoner hvor gjentak 1 og gjentak 2 bare er tilfeldige nummereringer. Noe annet ville det vært hvis gjentak 1 ble foretatt på en viss dato, og gjentak 2 en uke seinere. Da ville gjentakseffekten vært krysset med både sort og dommer, og vi ville havnet utenfor rammene til dette underkapitlet.

En variansanalyse av data fra en modell som beskrevet: en fast sortseffekt krysset med en tilfeldig dommereffekt, samspill mellom sort og dommer, og en hierarkisk og tilfeldig gjentakseffekt, vil gi oss svar som dette, lettere redigert for lesbarhetens skyld (her er benyttet SAS versjon 9.4):

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Sort	2	37.340333	18.670167	11.19	0.0007
Dmr	9	39.661500	4.406833	2.64	0.0380
Error	18	30.033000	1.668500		

Error: MS(Sort*Dmr)

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Sort*Dmr	18	30.033000	1.668500	1.24	0.2913
Error: MS(Error)	30	40.275000	1.342500		

Her ser vi at det er god grunn til å forkaste H_0 , siden p-verdien er så liten som 0,0007. At dommereffekten også er signifikant, er av mindre betydning, det betyr bare at de har benyttet forskjellige deler av





skalaen. At samspillet mellom sort og dommer ikke er signifikant, er også interessant – det betyr at dommerne stort sett har vært enige. Hvis for eksempel en dommer bedømmer søtheten i prøvene A – B – C til henholdsvis 2,3 – 3,1 – 4,0 og en annen dommer bedømmer de samme prøvene til 5,2 – 5,5 – 6,1 så er disse enige: C er søtere enn B som igjen er søtere enn A. Sagt på en annen måte: de to dommerne har rangert prøvene på samme måte med hensyn på søthet.

11.4.6 Flerveis variansanalyse

I prinsippet kan variansanalysemodellen utvides nærmest i det uendelige ved å legge til flere effekter, som både kan være krysset med eller være hierarkisk under en eller flere andre effekter, være faste eller tilfeldige. Med tilgang til statistisk programvare vil det være få problemer med å gjøre selve utregningene som skal til, men det kan fort oppstå problemer med tolkningene. Hvis det for eksempel er signifikante samspill, får dette betydning for tolkningen av hovedeffektene.

En annen type komplikasjon inntreffer når datasettet ikke er balansert. Et ubalansert datasett har vi når for eksempel en eller flere av dommerne ikke har bedømt alle prøvene fordi de har vært fraværende under deler av forsøket. Ubalanse kan også oppstå hvis en produksjon har gått galt, og det ikke går an å kjøre produksjonen på nytt. Sammenlikningene kan bli urettferdige og gale.

I noen situasjoner utelukker man enkelte kombinasjoner av effekter rett og slett fordi et fullstendig design med alle effektene kombinert med alle de andre effektene ville gi et uoverkommelig stort datasett. Såkalte «2-i-n'te» forsøk (2^n -forsøk) faller i denne kategorien. Slike forsøk består i å analysere et spesifisert antall effekter (n), hver på 2 nivåer. Da kan antall analyseprøver fort bli u håndterlig stort: hvis 7 slike effekter inngår, blir det $2^7=128$ prøver som skal bedømmes – og da har vi sett helt bort fra eventuelle gjentak. Her er det viktig at man følger et definert forsøksopplegg – i litteraturen finner man detaljene under Design of Experiments eller Experimental Designs.





11.4.7 Manglende verdier

I forsøk som går over flere dager, hender det at dommere er forhindret fra å være tilstede under hele bedømmelsen. Slike «hull» i datamatrisa gjør at teorien bak variansanalysen blir mer komplisert, og noen statistikkprogrammer kan også få problemer med å utføre beregningene hvis antall manglende verdier blir for stort. Hvis det mangler en eller flere verdier i datamatrisa, sier vi at modellen er ubalansert. Fortolkningsmessig er ubalanserte modeller mer kompliserte enn balanserte. Det er derfor en stor fordel å ha balanserte modeller. Det finnes en opplagt måte å gjøre en ubalansert modell balansert på, og det finnes en som er mindre opplagt. Hvis vi i utgangspunktet har relativt mange dommere (mer enn ca. 10), og en av disse bare har vært tilstede på noen av bedømmelsene, kan vi velge å utelate denne dommeren helt i beregningene. Dermed er modellen balansert. Et annet alternativ er å erstatte de manglende verdiene med ett eller annet gjennomsnitt. Denne framgangsmåten kan være akseptabel hvis programmet selv beregner hva disse manglende verdiene skal erstattes med, men er ikke å anbefale hvis de må beregnes mer eller mindre for hånd. Å erstatte manglende verdier med middelverdier innebærer også at man må justere frihetsgradene for restkvadratsummen.

Har man en ubalansert modell er altså valget om man vil kaste bort noen av dataene og dermed få en enkel, balansert og lett tolkbar modell, eller man vil analysere data med en ubalansert modell med de fortolkningsproblemer det medfører.

En annen årsak enn dommere som uteblir til at man får manglende verdier, ser vi ofte i dommerpanel hvor data registreres for hånd. I slike situasjoner er det fort gjort å glemme å fylle inn en eller flere karakterer.

11.5 Flere egenskaper samtidig: PCA

Alle metodene vi har sett på så langt i dette kapitlet har vært såkalte univariate – vi ser på hver egenskap for seg. Noen ganger er det også av interesse å se hvordan egenskapene varierer i sammen, og til det





trenger vi multivariable metoder. En mye brukt metode er prinsipal komponentanalyse, forkortet PCA etter den engelske betegnelsen Principal Component Analysis. En vanlig datamatrikse i et sensorisk forsøk vil ikke bare se slik ut som de tabellene vi hittil har presentert, men de vil ha en rad som tilsvare bedømmelsene for en dommer av en sort i ett gjentak, og så en kolonne for hver egenskap. Et eksempel med 8 dommere, sorter A-H, 2 gjentak og totalt 26 sensoriske egenskaper (lukteegenskaper L1-L9, smaksegenskaper S1-S12 og teksturegenskaper T1-T5) kan se slik ut (bare de første og siste radene og de første og siste kolonnene vises):

Tabell 11.5 Eksempel på full datamatrikse av en større sensorisk analyse

Dmr	Sort	Gjt	L1	L2	L3	L4	L5	L6	.	S8	S9	S10	S11	S12	T1	T2	T3	T4	T5
1	A	1	6,5	2,9	4,6	2,5	5,3	1,0	.	2,6	2,5	2,0	6,6	5,4	5,6	5,1	4,3	4,3	7,2
1	A	1	7,2	2,9	3,1	4,9	2,6	2,0	.	3,3	3,6	2,8	3,4	4,2	7,9	2,5	1,4	2,5	8,2
3	A	1	6,5	1,0	4,7	1,0	1,0	1,0	.	4,9	1,7	5,2	5,5	2,5	3,9	4,3	1,3	4,2	7,6
4	A	1	7,7	1,0	3,8	7,2	5,3	5,5	.	2,8	2,7	6,7	5,9	6,6	5,0	3,1	2,2	2,3	5,7
5	A	1	7,1	1,0	5,0	3,7	3,2	1,0	.	4,1	2,7	5,7	6,7	6,3	5,5	2,5	1,0	2,7	9,0
6	A	1	6,4	1,9	5,6	5,2	1,0	1,0	.	2,4	2,0	3,5	4,6	5,5	5,9	5,0	1,7	4,1	6,9
7	A	1	7,2	1,9	5,5	5,2	5,8	2,7	.	4,5	1,6	4,2	6,2	2,4	3,3	3,5	1,2	4,3	7,7
8	A	1	8,5	1,1	3,4	3,4	5,7	1,1	.	4,7	1,0	1,0	1,0	9,0	6,2	1,1	1,1	2,0	9,0
1	A	2	5,7	3,6	4,2	4,5	3,3	3,0	.	2,5	2,4	2,1	6,2	1,0	5,4	4,7	3,2	3,6	6,9
2	A	2	6,1	4,1	4,6	2,2	2,0	1,0	.	3,7	3,6	2,7	4,0	3,7	8,3	2,4	1,4	2,6	8,3
3	A	2	6,9	1,0	3,7	1,0	1,0	1,0	.	3,0	1,7	5,2	6,9	2,6	5,1	3,3	1,4	5,6	7,7
.
.
3	H	2	7,1	1,0	3,8	4,0	3,2	1,0	.	3,1	2,3	2,5	5,6	1,0	3,7	2,6	1,0	5,6	7,5
4	H	2	5,7	3,2	5,2	3,3	3,2	1,0	.	3,0	2,5	3,1	5,8	5,1	4,3	4,2	4,1	4,7	5,5
5	H	2	5,8	1,0	4,8	5,0	3,3	2,8	.	3,7	2,8	5,8	5,9	6,9	5,0	3,2	2,7	3,8	9,0
6	H	2	5,7	4,1	4,0	3,5	2,3	1,0	.	3,3	1,9	3,0	5,8	2,4	5,2	4,3	2,5	4,9	7,4
7	H	2	6,0	1,3	5,1	6,2	5,2	2,1	.	4,2	2,5	4,7	6,6	3,4	5,0	4,9	2,3	4,6	6,2
8	H	2	6,3	2,8	5,0	3,8	5,1	1,1	.	2,9	2,5	1,0	4,5	6,1	3,7	1,1	1,0	5,1	9,0

I motsetning til mer klassiske statistiske tester, kan vi ikke her sette opp en enkel formel som vi bare kan sette inn dataene våre i og så få resultatet ut, gjerne i form av ett enkelt tall, for eksempel en F-verdi. I stedet må man benytte en iterativ algoritme som konvergerer etter et visst iterasjoner. Detaljene i algoritmene kan variere fra program til program – dette er en metode som aldri vil gjøres ved hjelp av



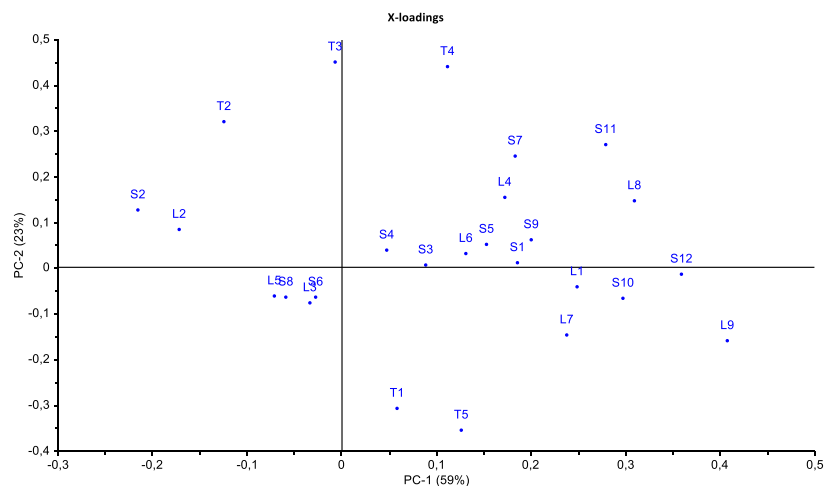


manuelle beregninger. Hovedresultatet fra en PCA er skåringsplot (scores) og ladningsplot (loadings). Ofte brukes resultatene fra en PCA som en måte å få et første oversiktsbilde av datasettet på. Bruker vi PCA på dataeksemplet over, vil vi kunne identifisere enkeltbedømmelser som skiller seg ut, såkalte outliers. Hva man i så fall gjør med slike avvikende resultater, er et annet spørsmål! Men hvis det skulle vise seg at alle bedømmelsene på stikkende lukt på prøver fra en bestemt fisk, så kan det være grunn til å undersøke om det under forbehandlingen av denne prøven har skjedd noe spesielt som gjør at den bør utelukkes fra analysene. Er man fornøyd med datasettet slik det er, er det vanlig å beregne gjennomsnitt over både gjentak og dommere når man gjør en PCA. Da vil datasettet redusere seg til:

Tabell 11.6: Eksempel på redusert datamatrix av et større datasett

Prøve	L1	L2	L3	L4	L5	L6	L7	.	S8	S9	S10	S11	S12	T1	T2	T3	T4	T5
A	6,55	2,23	4,33	3,91	3,51	1,84	3,19	.	3,69	2,49	3,53	5,16	4,32	5,45	3,46	1,89	3,66	7,73
B	6,76	2,16	4,37	4,48	3,65	1,83	2,86	.	3,57	2,46	3,63	5,58	3,64	5,52	3,08	2,02	4,13	7,63
C	5,74	2,94	4,01	4,13	3,46	1,73	1,86	.	3,41	2,44	2,94	5,51	3,45	4,81	4,03	3,03	4,95	7,02
D	5,63	2,73	4,66	3,33	4,16	1,48	2,21	.	4,14	1,54	2,16	3,81	2,67	5,25	3,72	1,94	3,58	6,96
E	5,99	2,76	4,36	4,14	3,73	1,53	2,13	.	3,83	1,91	2,38	5,15	2,93	4,38	4,66	3,21	4,70	5,99
F	5,94	2,64	4,89	3,35	3,97	1,44	2,85	.	3,76	1,64	2,53	4,86	2,71	4,34	4,43	2,46	3,83	6,63
G	7,03	1,84	4,61	4,38	3,78	2,30	3,68	.	3,75	2,63	3,58	5,90	4,59	4,53	4,07	2,95	5,03	6,59
H	6,23	2,34	4,51	3,89	3,72	1,36	2,64	.	3,61	1,93	3,02	5,68	3,76	4,78	3,73	2,14	4,29	7,13

En PCA kjørt ved hjelp av Unscrambler X ga følgende resultat:

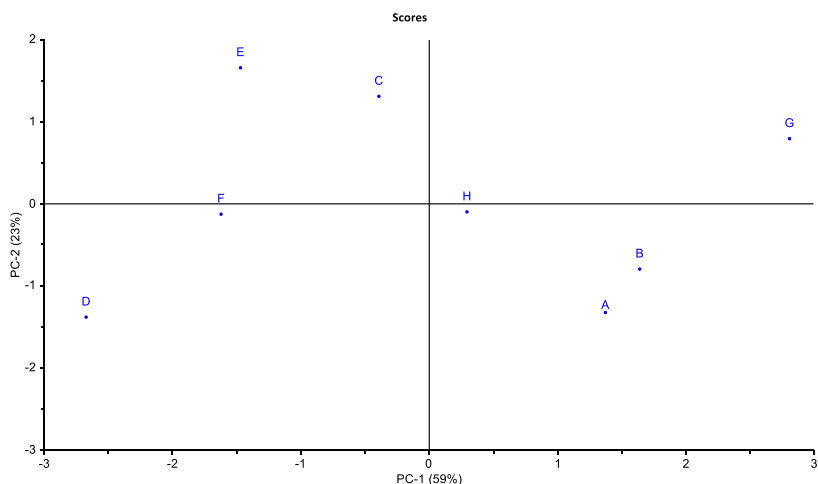


Figur 11.1: PCA plot med markering av produktegenskaper (loadings)





Ladningsplottet viser hvordan egenskapene plasserer seg når vi ser på deres plassering i det 2-dimensjonale planet definert av de to viktigste dimensjonene. Disse representerer $59\% + 23\% = 82\%$ av all variasjonen i datasettet. Her kan vi røpe at S2 og L2 er syrlig smak og lukt, og L8, L9, S11, S12 har med stram og emmen å gjøre. Den viktigste dimensjonen er altså en akse fra syrlig til emmen. T1 og T3 er hardhet og saftighet, så disse 2 egenskapene er viktige når det gjelder å definere prinsipalkomponent 2.



Figur 11.2: PCA plot med markering av prøveplassering (scores)

I skåringsplot'et ser vi hvordan prøvene plasserer seg langs de samme dimensjonene. D og F ser dermed ut til å være de som har skåret høyest på syrlig smak og lukt, mens G er mest emmen. Litt verre er det å plassere prøvene langs en hardhet-saftighet-akse (dimensjon 2), utover at C ser ut til å være den som scorer høyest på saftighet.

I noen tilfeller kan det også være aktuelt å se på flere dimensjoner og plotte dem mot hverandre i tillegg.

PCA opptrer i litteraturen (særlig den litt «eldre» – før 1980...) også under andre navn: Singular Value Decomposition (SVD) eller egenvektor-dekomposisjon.





11.6 Dypdykking i datamaterialet

En sensoriker vil ofte føle et behov for å gå nærmere inn i datamaterialet enn det en som bare er interessert i å sammenlikne sorter eller prøver er. Kvalitetskontroll av det sensoriske panelet er en viktig del av den daglige driften, og i forbindelse med trening og rekruttering av enkeltdommere, er det også viktig med en form for kvalitetskontroll. Et verktøy som etter hvert har fått en viss status i det sensoriske miljøet er programmet PanelCheck, utviklet i et samarbeid mellom sensorikere og statistikere ved Nofima og Danmarks Tekniske Universitet og med støtte fra forskningsråd og industripartnere i Norge og Danmark. PanelCheck kan lastes ned gratis fra www.panelcheck.com.

PanelCheck er delt inn i 4 hoveddeler:

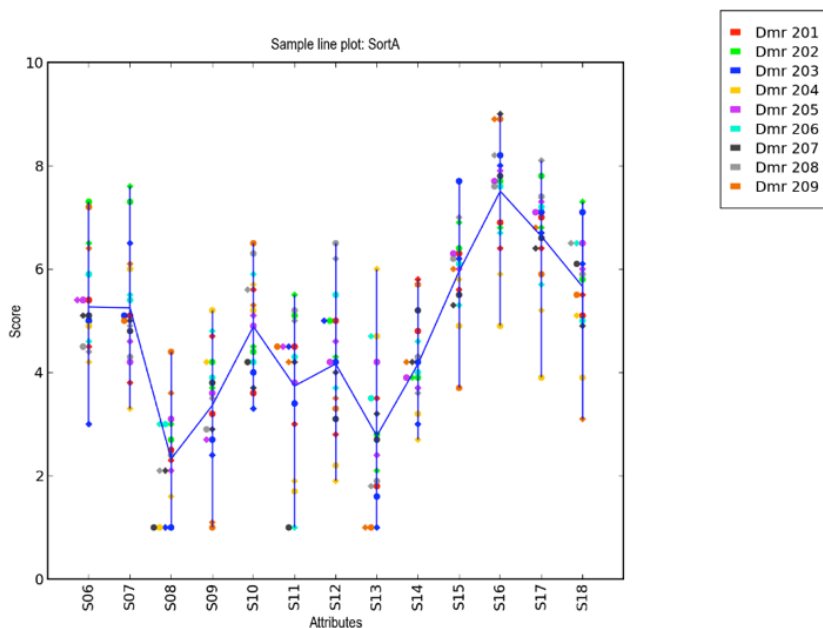
- Univariate
- Mutivariate
- Consensus
- Overall

Noen av de ovenstående punktene inngår i forskjellige indekser for enighet mellom enkeltdommere og panelet (AGRPROD), hvordan dommerne er enige med panelet når det gjelder korrelasjoner mellom egenskapene (AGRATT), hvordan dommerne er enige med seg selv over gjentakene (REPPROD), hvordan dommerne er enige med seg selv over gjentak når det gjelder egenskapene (REPATT) og antall egenskaper som dommerne kan skille fra hverandre på nivå 5% (DIS). Disse indeksene er basert på RV2-koeffisienten (Tomic og flere (2013)). Indeksene vil etter hvert bli lagt inn i nye versjoner av PanelCheck.

Data kan importeres fra flere standardformater, så som rene tekstfiler (i diverse varianter) og Excel-filer.

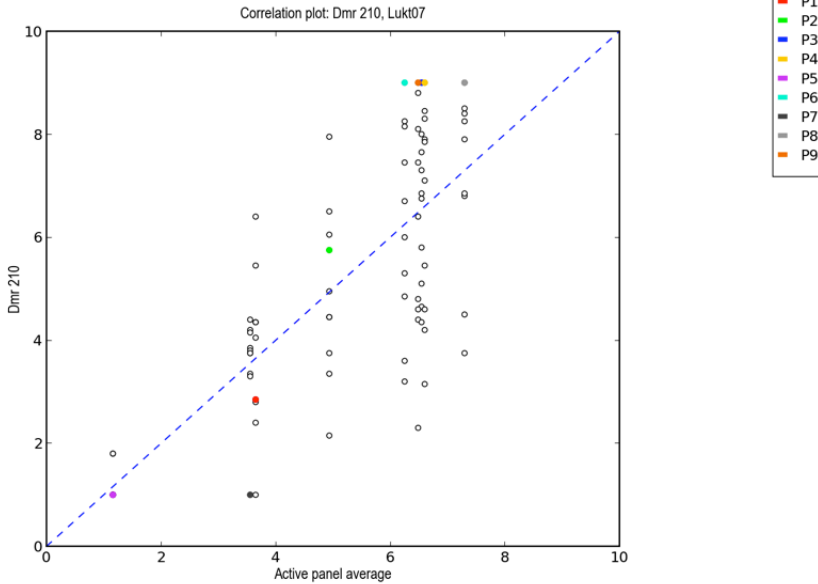


Line Plots er nettopp det: enkle figurer som for hver sort (her: Sort 9) viser enkeltgjentakene mot egenskap (her: S1 – S10), pluss panelets gjennomsnitt (blå linje).



Figur 11.3: Line plot som viser alle enkeltbedømmelsene for Sort A for utvalgte egenskaper (S06-S18)

For å se på enkeltdommeres innsats har PanelCheck også andre måter å vise resultatene på: middelerverdi over prøver (har dommerne en tendens til å ligge lavt eller høyt på skalaen?), og hvordan ligger en dommer i forhold til de andre dommerne. En enkeltdommers bedømmelser sammenliknet med de andre dommeres bedømmelser vises som et Correlation plot:

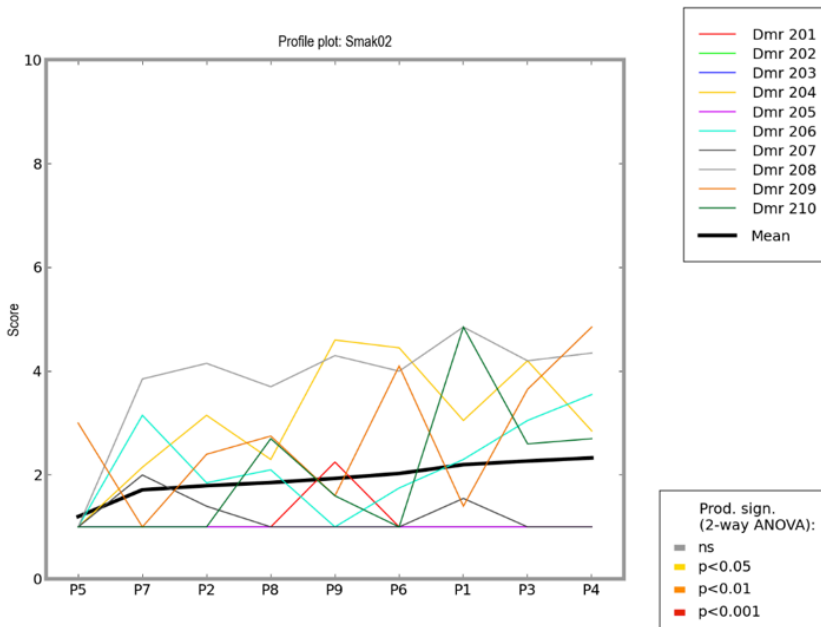


Figur 11.4: Correlation Plots viser enkeltdommernes bedømmelser sammenliknet med panelets gjennomsnitt.

I eksemplet over er det dommer 210 som vises ved de fargede punktene (som i sin tur representerer prøve P1, P2, ..., P9 i henhold til kodinene øverst til høyre). En dommer som er 100% enig med panelgjennomsnittet, vil ha alle punktene på den stiplede linja. En dommer som stort sett ligger over linja har en tendens til å gi høyere bedømmelser enn panelet, en som stort sett ligger under har en tendens til å gi lavere bedømmelser enn panelet.



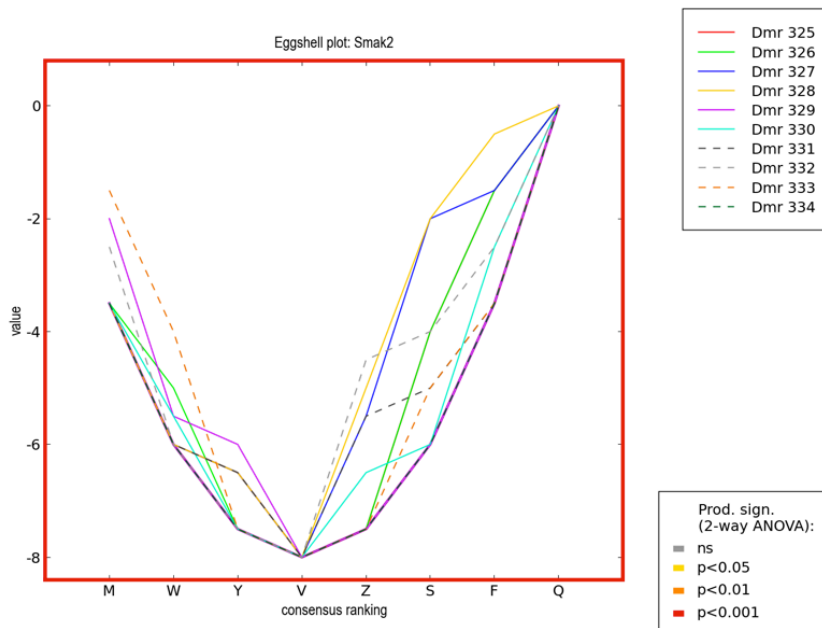
Profile Plots viser enkeltdommernes resultater (midlet over gjentak) sammen med panelets middelværdi. X-aksen utgjøres av de forskjellige prøvene (her: P1, P2, ..., P9) ordnet i stigende rekkefølge etter panelmiddelværdiene (andre rekkefølger kan velges). Eksemplet under viser ganske stor spredning mellom dommerne, og sammenliknet med denne er det liten forskjell mellom laveste verdi (P5 = 1,20) og høyeste verdi (P4 = 2,33). At prøvene ikke kan sies å være signifikant forskjellige, sees også ved at rammen rundt plottet er grå.



Figur 11.5: Profilplot for Smak02. Det er liten forskjell mellom prøvene, og panelet har heller ikke funnet signifikante forskjeller. Enkeltdommerens bedømmelser spriker en god del.



Eggshell Plots er en måte å sjekke i hvilken grad dommerne er enige når det gjelder rangering av prøvene. Samspillseffekten Dommer \times Prøve i en variansanalyse vil også være ett uttrykk for dette, men samspillet er et lite håndfast begrep for de fleste. Dette vises bedre med figurer:



Figur 11.6: Eggeskallplot for egenskapen Smak2.

Dommerne er stort sett enige om rangeringen mellom prøvene, og panelet har funnet signifikante forskjeller mellom prøvene. For en egenskap hvor det er mer uenighet, vil de forskjellige kurvene sprike mye mer, og ofte vil heller ikke panelet ha funnet signifikante forskjeller

Ideelt sett skal alle dommerne falle sammen med den nederste kurven, men at alle dommerne skulle 100% enige i rangeringene av alle prøvene i et datasett, er urealistisk.





Når dommernes bedømmelser spriker ganske mye, så kan det ha flere årsaker. Det kan godt tenkes at prøvene skiller seg veldig lite fra hverandre, og da er det heller ikke noe poeng i at dommerne ikke er enige i rangeringene av dem. Eggeskallplottet tar ikke hensyn til om prøvene ligger i området 1,3 til 8,7 eller 5,6 til 5,9. At dommernes bedømmelser spriker mye, kan i seg selv bidra til at forskjellene ikke blir signifikante: nevneren i F-testen blir for stor.

Om en dommer skulle avvike fra resten av panelet, er det også viktig å være klar over at dette kan skyldes at akkurat denne dommeren er spesielt sensitiv overfor denne egenskapen, og at det faktisk er denne som har «rett», mens alle de andre har «feil». Den ideelle bruken av et eggeskallplot er situasjoner hvor fasiten er kjent, for eksempel prøver tilsatt kjente mengder sukker som skal bedømmes for søthet. Da kan eggeskallplot benyttes for å vise dommernes prestasjoner.

F & p Plots og MSE Plots lager oversikter over F-, p- og MSE-verdiene fra variansanalysen. Denne typen figurer er godt egnet til å få en rask oversikt om det er enkelte egenskaper som stiller seg ut i en eller annen retning.

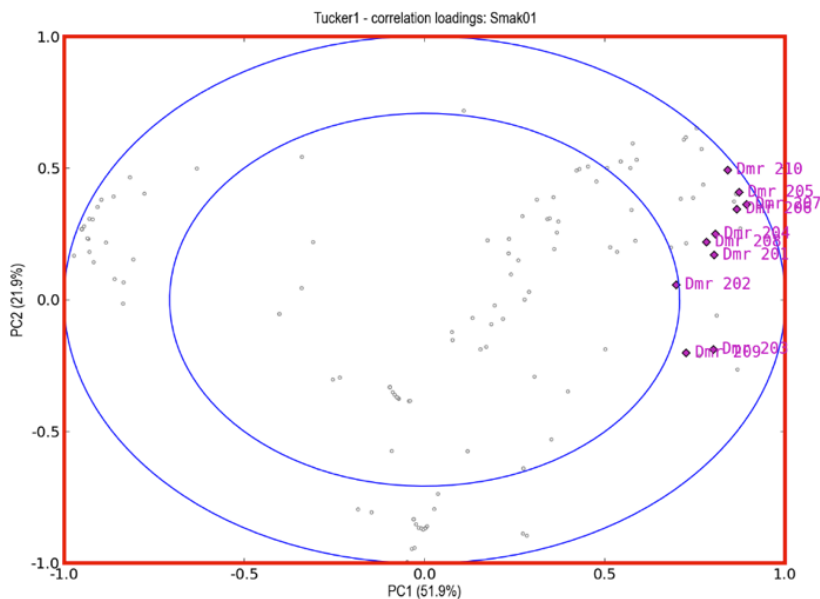
p-MSE Plots er et annet plot som kan hjelpe til å fortelle hvordan dommeren har prestert. Det tar utgangspunkt i at det gjøres en variansanalyse for hver dommer separat. MSE er et uttrykk for hvor godt dommeren har gjentatt seg selv, altså om vedkommende har bedømt forskjellige gjentak fra samme prøve omtrent likt. Samtidig er det av interesse å få vite om dommeren har kjent forskjell mellom prøvene, p er resultatet fra F-testen. Hvis p er liten, har dommeren kjent forskjell. Også disse plot'ene må tolkes med forsiktighet: en variansanalyse basert på et veldig lite datasett, som er vanlig når vi analyserer dommerne hver for seg, er i sin natur ustabil. Men det gir allikevel en viss pekepinn på hvordan dommeren har prestert.





Tucker-1 Plots

Datsettet «brettes ut» ved at matrisene for enkeltdommernes bedømmelser (antall rader=antall prøver, antall kolonner=egenskaper) settes ved siden av hverandre og analyseres ved hjelp av PCA. Et Correlation loading plot avslører i hvilken grad dommerne tolker egenskapene forskjellig.



Figur 11.7: Tucker1-plot for Smak01

Her ligger dommerne godt samlet og stort sett nær den ytterste ellipsen; et tegn på at panelet fungerer godt for denne egenskapen. Hvis dommerne hadde vært mer uenige, ville de ligget spredd mer eller mindre ut over hele ellipsen. Med få prøver er Tucker-1 mindre egnet, siden dommerne ville havne nær den ytre ellipsen uansett. Derfor anbefales minst 7 prøver hvis denne analysen skal tillegges noe vekt.





Manhattan Plots

Dette er et annet alternativ for se på forskjeller mellom dommerne: her ser vi på forklart varians for hver dommer/egenskap-kombinasjon. Mange lyse felter i figuren forteller oss at dommerne har forklart mesteparten av variansen allerede etter de første prinsipalkomponentene. Forklarer dommerne en mindre del av variansen blir det desto flere mørke partier.

Consensus – Original – Standardized – STATIS

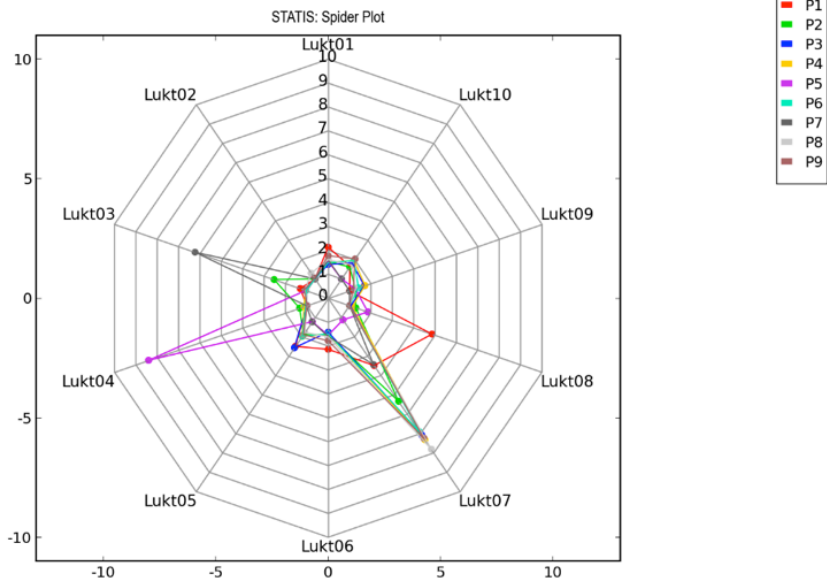
Disse 3 variantene gjør stort sett det samme: PCA enten på original-data eller på standardiserte data (alle verdiene for en egenskap divideres med standardavviket over alle verdiene), dermed sikrer man at alle egenskapene får samme standardavvik. I sensorikk er det ikke så vanlig at dataene standardiseres, siden de i utgangspunktet er målt på samme skala. Standardisering benyttes heller i situasjoner hvor dataene består av målinger langs forskjellige skalaer, for eksempel hvis de sensoriske dataene skal inngå i en matrise sammen med vekt i gram, vannprosent og liknende. I tillegg til PCA kan STATIS vise enkelt-dommernes bidrag til vektingen som STATIS benytter i de øvrige beregningene og et spider-plot, som er en form for linjeplot for middelverdiene av egenskapene.

Under fanen Overall kan PanelCheck utføre variansanalyser i noen utvalgte enkle, men ofte forekommende, modeller. I versjonen tilgjengelig per april 2015, er for eksempel følgende 3 ANOVA-modeller tilgjengelige:

- 2-way ANOVA (1 rep): Toveis variansanalyse uten gjentak i henhold til terminologien i denne boka
- 2-way ANOVA: Toveis variansanalyse med gjentak i henhold til terminologien i denne boka
- 3-way ANOVA: To faste effekter og dommer-effekt, alle faktorer krysset, ingen gjentak.

PanelCheck gir de samme resultatene som tradisjonelle statistikkprogram gir, men utskriften egner seg lite som vedlegg i rapporter.





Figur 11.8: Spider plot fra STATIS-delen



Referanser:

Daniel M Ennis, Benoît Rousseau, Technical Report: Identifying and removing sources of bias in product tests and surveys. The Institute for Perception, Newsletter Spring 2015.

David Hirst, Tormod Næs, A graphical technique for assessing differences among a set of rankings. *J. Chemometrics* 1994, 8, pp 81-93

Yosef Hochberg, Ajit C Tamhane, *Multiple Comparison Procedures*. New York: John Wiley & Sons, 1987. ISBN 0-471-82222-1

Daniel Kahneman, *Thinking, fast and slow*. New York: Farrar, Strauss and Giroux, 2011. ISBN 978-0374275631
Norsk utgave: *Tenke, fort og langsomt*. Oslo: Pax forlag, 2012. ISBN 978-82-530-3552-9

Per Lea, Marit Rødbotten, Tormod Næs, Measuring validity in sensory analysis. *Food Quality and Preference* 1995, 6, pp 321-326

Oliver Tomic, Ciaran Forde, Conor Delahunty, Tormod Næs, Performance indices in descriptive sensory analysis – A complimentary screening tool for assessor and panel performance. *Food Quality and Preferences* 2013, 28, pp 122-133

www.panelcheck.com.





Foto: Julie H. Gjermundsen





SENSORISK ANALYSE AV NON-FOOD

Liv Bente U. Strandos, Elopak AS

Sensorikk er per definisjon «en vitenskapelig disiplin brukt for å overvåke, måle, analysere og tolke reaksjoner av de karakteristikkene av mat og materialer slik de er oppfattet ved bruk av de menneskelige sansene syn, smak, lukt, berøring og hørsel». I tillegg til næringsmidler, er dette gjeldende for alle typer produkter, inkludert kosmetikk, rengjøringsmidler, emballasje, tekstiler, trevirke, tobakksprodukter, legemidler, audio- og videoprodukter, transportmidler med flere. Alt som har en sensorisk karakteristikk oppfattet av en eller flere av de menneskelige sansene kan måles.

Sensorisk analyse av non-food er et område i rask vekst. Stadig flere produsenter ser nytten av å ha kunnskap om hvordan forbrukerne opplever deres eget produkt. Hensikten med dette kapittelet er å synliggjøre at sensorisk analyse er mye mer enn å måle sensoriske egenskaper av ulike næringsmidler. Nedenfor er eksempler hvor sensoriske analysemetoder brukes innenfor ulike non-food-bransjer.

12.1 Emballasjeindustrien

Liv Bente U. Strandos, Manager Sensory Services, Elopak AS

Emballasje har mange funksjoner. En er å beskytte produktet og øke holdbarheten i tillegg til å ivareta produktets kvalitet. Emballasjen har også i oppgave å selge produktet og gi selskapet en positiv image. Den skal også være lett å transportere, håndtere, lagre, lett å åpne samt gi informasjon om f. eks innhold og opprinnelse av produktet.





Ulike sensoriske metoder kan brukes, alt avhengig av hvilke egenskaper av emballasjen som skal vurderes, enten det er lukt, smak, visuelt eller taktilt (berøring). Innenfor analyse av lukt og smak, finnes det flere relevante og aktuelle metoder som:

- NMKL-prosedyre nr. 19, 2007 Riktlinjer för sensorisk bedömning av livsmedelsförpackningar
- NS-EN 1230-1:2009 Papir og papp beregnet på å komme i kontakt med næringsmidler - Sensorisk analyse - Del 1: Lukt
- NS-EN 1230-2:2009 Papir og papp beregnet på å komme i kontakt med næringsmidler - Sensorisk analyse - Del 2: Smakoverføring
- ASTM E460 Standard Practice for Determining Effect of Packaging on Food and Beverage in Products During Storage. (2012)
- ASTM E619 Standard Practice for Evaluating Foreign Odors in Paper Packaging. (2012)
- ASTM E1870 Standard Test Method for Odor and Taste Transfer from Polymeric Packaging Film. (2011)
- ASTM 2609 Standard Test Method for Odor or Flavor Transfer or Both from Rigid Polymeric Packaging. (2008)



Figur 12.1 Drikkekartonger

Foto: Elopak AS





«Som produsent for emballasje som skal komme i direkte kontakt med næringsmidler er vi underlagt streng lovgivning. EU-direktivet pålegger oss å forsikre at emballasjen ikke forårsaker helseskade eller endrer produktets sensoriske egenskaper. Sensorisk analyse er i så måte en veldig viktig del av arbeidet innen produktutvikling, prosessoptimalisering og kvalitetskontroll. Det interne panelet er spesielt godt trent i å detektere og karakterisere ulike smaker typiske for pappbasert emballasje. Vår erfaring er at sensorisk analyse ofte gir viktig informasjon hvor kjemiske målinger kommer til kort, selv om en kombinasjon av de to disipliner i mange tilfeller er en klar fordel.»

– Liv Bente U. Strandos, Manager Sensory Services – Elopak AS



Figur 12.2 Såpeskum og børste

Foto: Kjell Merok

218 Sensorikk – Måling med menneskelige sanser





12.2 Hygieneprodukter og vaskemidler

Margit Anette Bleg Hegna, Lilleborg (part of Orkla Group)

Sensorikk i form av duft, farge, utseende og konsistens er viktige emosjonelle verdier for et hygieneprodukt eller et vaskemiddel.

Duften gir produktet dets karakter og er på mange måter produktets signatur. For eksempel er duften av Zalo for mange forbrukere definisjonen på hvordan et håndoppvaskemiddel skal lukte. Tilsvarende assosiasjoner eksisterer også for Milo tøyvaskemiddel i forhold til ulltøy.

Duften av et produkt påvirker humøret, og god duft av et rengjøringsmiddel gjør vaskejobben mer lystbetont. Den er også knyttet til velværefølelsen i dusjen når en benytter en dusjsåpe eller en sjampo. Duften av renhet er i tillegg et viktig bevis for at ting faktisk er rent. For eksempel så skal nyvasket tøy både se rent ut og dufte godt for at forbrukeren opplever at det er rent.

Det finnes ingen offisielle standarder eller krav i forhold til utvikling eller bedømming av duft for et hygieneprodukt eller et vaskemiddel. Ved utvikling av en ny duft for et tøyvaskemiddel vil en velge mulige duftretninger utfra trender og forbrukerpreferanser. For å evaluere duften vil en typisk lage fysiske prøver av produktet med ulike aktuelle parfymmer for deretter å utføre en brukstest med forbrukere i målgruppen for produktet. Det er viktig å ha tilstrekkelig med respondenter for å få valide resultater. Forbrukerne skal da evaluere både duften av selve produktet, duften av vått tøy rett ut av maskinen og duften av tørt tøy. En kan også benytte trenede personer for å bistå i utvelgelsen av dufter for eksempel parfymører fra parfymehusene.

Non-food er definert som noe som ikke er mat, men allikevel solgt i dagligvarebutikker, som for eksempel husholdningsartikler, personlig pleie produkter og kontorekvisita.



12.3 Lyd og billede

Søren Bech, Bang & Olufsen a/s,

Bang & Olufsen a/s har brugt sensorisk analyse i udviklingsarbejdet siden 1976. Det videnskabelige grundlag og metodikken, blev for alvor udviklet i forskningsprojektet Adonis «Perceptual image quality of television displays» (1994 – 99) i samarbejde med Philips NatLab og Philips TV. Den udviklede metode RaPID og brug af metoden var banebrydende inden for bedømmelse af billedkvalitet. Efterfølgende blev metoden videreudviklet til brug for bedømmelse af lyd-kvalitet og der blev etableret permanente teams til lytte- og kiggeforsøg. Medlemmerne af teams blev udvalgt efter en række fysiologiske og perceptuelle tests og ud af ca. 100 ansøgere blev 20 udvalgt. Bang & Olufsen har løbende brugt sensorisk analyse i forbindelse med dels forskningsprojekter og dels i forbindelse med produktudvikling og benchmark analyser af konkurrerende produkter. Sensorisk analyse bruges også

i forbindelse med intern dokumentation af opfyldelse af specifikationer for produkter. Brugen af sensorisk analyse til detaljerede undersøgelser af «experience» er vigtig for Bang & Olufsen fordi samspillet mellem mennesker (brugeren) og vores produkter er og har altid været central i design og udviklingsprocessen – reflekteret i den nyeste udgave af vores mission statement «Bang & Olufsen exists to move you with enduring magical experiences».



Figur 12.3 Høytaler Foto: Bang & Olufsen



12.4 Trelast- og trevareindustrien

Anders Q. Nyrud, Norsk Treteknisk Institutt

Anders Roos, SLU

Treindustrien produserer hovedsakelig byggematerialer og byggevarer. For estetiske produkter, som gulv, innvendig panel, listverk og lignende, eksisterer det kvalitetsstandarder som stiller krav til materialenes utseende. Disse tar utgangspunkt i fysiske virkesegenskaper som kvistsetting, årringbredde, margens plassering og sprekkdannelse. Kriteriene er basert på industriens egne erfaringer og vurdering av produktenes visuelle egenskaper.

Materialvalg påvirker brukeres oppfatning av innemiljø, og trematerialer kan bidra til både å heve opplevd velvære og det visuelle inntrykket av bygg. Produktegenskaper som farge eller utseende, i enkelte tilfeller også taktile egenskaper og lukt, er viktige for hvordan valg av byggemateriale påvirker et innemiljø. Industriens kvalitetskriterier gir ikke tilstrekkelig informasjon til å forstå hvordan virkesegenskaper påvirker treoverflaters visuelle inntrykk. Kvalitetskriteriene gir heller ikke grunnlag for å beskrive taktile egenskaper og lukt.

Det har vært gjennomført undersøkelser for å kartlegge sensoriske egenskaper som kan forklare treoverflaters visuelle og taktile inntrykk. Nyrud med flere (2008) brukte i sin analyse et sensorisk panel som var godt trent i bedømmelse av varer ved hjelp av beskrivende metode (se kapittel 4.3), men som tidligere ikke hadde bedømt produkter av tre. Dommerpanelet analyserte visuelt ulike terrassematerialer og de sensoriske resultatene ble sammenholdt med forbrukeres preferanser. Atten visuelle og taktile attributter ble identifisert – over halvparten av disse inngår ikke i industriens kvalitetskriterier, men viste seg hensiktsmessige i en mer grundig beskrivelse av treoverflater. Studier viser at forbrukerne foretrekker materialer med et homogent visuelt uttrykk og moderat fargeintensitet. Visuell homogenitet kan beskrives med sensoriske variabler, materialavhengige variabler og produksjonsavhengige variabler.

I tillegg til å få en sensorisk beskrivelse av trematerialer er sensorisk profilering også brukt i sammenheng med studier av forbrukeres





Figur 12.3 Test av treoverflater

Foto: Kjell Merok

assosiasjoner og tolkninger av trematerialer. Tre foretrekkes framfor tre-plastkompositter, antagelig fordi det oppleves som mer naturlig og behagelig som igjen kan komme av at det oppleves som varmere og mykere. I en studie med fokus på taktile egenskaper ble sensorisk analyse kombinert med produktsemantikk på materialer av tre og tre-kompositt. En lignende studie hvor istedet bare visuelle egenskaper ble benyttet viser at treslag gir ulike opplevelser og assosiasjoner hos forbrukere. Ask oppleves eksempelvis som eksklusivt mens gran og furu er mer naturlig.

Kunnskap som følge av sensoriske studiene av materialer kan gi grunnlag for materialvalg for interiør når vi ønsker å gi rommet en spesiell atmosfære.

222 Sensorikk – Måling med menneskelige sanser





Referanser:

S. Bech, R. Hamberg, M. Nijenhuis, C. Teunissen, H. Looren de Jong, P. Houben, S. K. Pramanik "The RaPID

Perceptual Image Description Method (RaPID)", SPIE Conference "Human Vision and Electronic Imaging", 29 January - 1 February, California, 1996, paper 2657

M. Nijenhuis, R. Hamberg, C. Teunissen, S. Bech, H. Looren de Jong, P. Houben, S. K. Pramanik "Sharpness, sharpness related attributes, and their physical correlates" SPIE Conference "Very High Resolution and Quality Imaging II", 10 - 11 February, California, 1997, paper 3025

Jonsson O, Lindberg S, Roos A, Hugosson M, Lindström M. 2008. Consumer perceptions and preferences on solid wood, wood-based panels and composites – A repertory grid study. *Wood and Fiber Science* 40(4): 663-678.

Høibø, O., Nyrud, A.Q. 2010. Consumer perception of wood surfaces: the relationship between stated preferences and visual homogeneity. *Journal of Wood Science* 56(4):276-283.

Lindberg S., Roos A., Kihlstedt A., Lindström M. 2013 A product semantic study of the influence of the sense of touch on the evaluation of wood-based materials. *Materials and Design* 52 (2013):300–307.

Nyrud, A.Q., Roos, A., Rødbotten, M. 2008. Product attributes affecting consumer preference for residential deck materials. *Canadian Journal of Forest Research* 38(6):1385-1396.

Roos A., Lindberg S., Kihlstedt A. 2013. A product semantic study of the influence of the sense of vision on the evaluation of wood-based materials. *Wood and Fiber Science* 45(4):1-10.





SENSORISKE NETTVERK

Behovet for å diskutere sensoriske problemstillinger og å hente og formidle kunnskap om faget sensorikk har ført til etablering av flere nettverk. Deltakere med ulik faglig bakgrunn, men med samme ønske om faglig utvikling innen sensorikk møtes for forpliktende eller mer uforpliktende samarbeid. I de nordiske landene og i mange andre land har fagpersoner opprettet nasjonale nettverk.

Sensorisk StudieGruppe (SSG) ble opprettet i Norge i 1972. Pådriverne var personer innen norsk næringsmiddelindustri, forskningsinstitusjoner og læresteder. I løpet av de årene SSG har eksistert har medlemstallet økt og det er fortsatt åpent for nye medlemmer som har en arbeidssituasjon som krever kunnskap om sensoriske analyser.

I tillegg til møter med faglig innlegg har SSG arrangert flere kurs for personer som har og/eller skal ha ansvaret for sensoriske bedømmelser i sin bedrift. Tema på disse kursene har vært grunnleggende praktisk arbeid med oppretting av sensorisk panel, metodegjennomføring og statistiske analyser for resultatbehandling. Den første norske læreboka i sensorisk analyse ble utgitt i 1977 med medlemmer i SSG som forfattere. Boka ble revidert første gang i 1997 og nå for andre gang i 2015 .





Informasjon om Sensorisk Studiegruppe i Norge: www.sensorikk.no

Informasjon om Sveriges Sensoriska Nätverk finnes på: www.ssn.nu

Informasjon om Dansk Sensorik Selskab (D2S) finnes her: <http://ida.dk/netvaerk/dansk-sensorik-selskab-d2s>

For informasjon om internasjonale nettverk og konferanser henviser vi til adresser på internett.

Nordisk Metodikkomité for Næringsmidler (NMKL):
www.nmkl.org

European Sensory Network (ESN): www.esn-network.com

European Sensory Science Society (E3S): www.e3sensory.eu

The Sensometric Society: www.sensometric.org

European Conference on Sensory and Consumer Research:
www.eurosense.elsevier.com

Nordic Workshop in Sensory Science. – De nordiske landene skifter på å arrangere denne konferansen som vanligvis holdes hvert annet år. Ingen fast nettsadresse.

Pangborn Sensory Science Symposium. – Dette symposiet arrangeres hvert annet år og arrangørsted skifter fra gang til gang.
Ingen fast nettsadresse





NMKL-PROSEDYRER, ISO-STANDARDER, ASTM-STANDARDER

NMKL (Nordisk Metodikkomite for næringsmidler)

www.nmkl.org

Metoder:

183 – Kvalitetskontrolltest av drikkevann (NMKL 183, 2005. Amd.2007,2013)

Prosedyrer:

- 6 – Generelle retningslinjer for kvalitetssikring af sensoriske laboratorier. (NMKL-prosedyre nr 6, 1998)
- 11 – Sensorisk bedømmelse av drikkevann. (NMKL-prosedyre nr 11, 2010)
- 14 – SENSVAL: Retningslinjer for egenkontroll i sensoriske analyselaboratorier (NMKL-prosedyre nr 14, 2004)
- 19 – Riktlinjer för sensorisk bedömning av livsmedelsförpackningar, (NMKL-prosedyre nr 19, 2007)
- 21 – Guide for sensory analysis of fish and shellfish (NMKL-prosedyre nr 21, 2008)
- 27 – Måleusikkerhet i sensoriske analyser (NMKL-prosedyre nr 27, 2013)
- 28 – Guidelines for reporting sensory data (NMKL-prosedyre nr 28, 2014)
- 29 – Sensorisk analyse av kjøtt og kjøttprodukter (NMKL-prosedyre nr 29, 2015)

ISO (International Organization for Standardization), Genève, Switzerland.

www.iso.org

Standarder under gruppen ICS- 67.240 – TC 34 - à jour per mars 2015.

ISO 707:2008 (IDF 50 : 2008) Milk and milk products – Guidance on sampling

ISO 3103:1980 Tea – Preparation of liquor for use in sensory tests

ISO 3591:1977 Sensory analysis – Apparatus – Wine-tasting glass.

ISO 3972:2011 Sensory analysis – Methodology – Method of investigating sensitivity of taste.

ISO 4120:2004 Sensory analysis – Methodology – Triangular test.

ISO 4121:2003 Sensory analysis – Methodology – Evaluation of food products by methods using scales.

ISO 5492:2008 Sensory analysis – Vocabulary.

ISO 5494:1978 Sensory analysis – Apparatus – Tasting glass for liquid products.

226 Sensorikk – Måling med menneskelige sanser





ISO 5495:2005	Sensory analysis – Methodology – Paired comparison test.
ISO 5496:2006	Sensory analysis – Methodology – Initiation and training of assessors in the detection and recognition of odours.
ISO 5497:1982	Sensory analysis – Methodology – Guidelines for the preparation of samples for which direct sensory analysis is not feasible.
ISO 6564:1985	Sensory analysis – Methodology – Flavour profile methods.
ISO 6658:2005	Sensory analysis – Methodology – General guidance.
ISO 6668:2008	Green coffee – Preparation of samples for use in sensory analysis.
ISO 7304:1985	Durum wheat semolinas and pasta – Estimation of cooking quality of spaghetti by sensory analysis
ISO 7304-2:2008	Alimentary pasta produced from wheat semolina – Estimation of cooking quality by sensory analysis
ISO 8586:2012	Sensory analysis – General guidance for the selection, training and monitoring of selected assessors and expert sensory assessors
ISO 8587:2006	Sensory analysis – Methodology – Ranking.
ISO 8588:1987	Sensory analysis – Methodology – «A»-«not A» test.
ISO 8589:2007	Sensory analysis – General guidance for the design of test rooms.
ISO 10399:2004	Sensory analysis – Methodology – Duo-trio test.
ISO 11035:1994	Sensory analysis – Identification and selection of descriptors for establishing a sensory profile by a multidimensional approach.
ISO 11036:1994	Sensory analysis – Methodology – Texture profile.
ISO 11037:2011	Sensory analysis – Guidances for sensory assessment of the colour of products
ISO 11056:1999	Sensory analysis – Methodology – Magnitude estimation method
ISO 11132:2012	Sensory analysis – Methodology – Guidelines for monitoring the performance of a quantitative sensory panel
ISO 11136: 2011	Sensory analysis – Methodology – General guidance for conducting hedonic tests with consumers in a controlled area
ISO 13299:2003	Sensory analysis – General guidance for establishing a sensory profile.
ISO 13300-1:2006	Sensory analysis – General guidance for the staff of a sensory evaluation laboratory – part 1 : Staff responsibilities
ISO 13300-2:2006	Sensory analysis – General guidance for the staff of a sensory evaluation laboratory – part 2 : Recruitment and training of panel leaders
ISO 13301:2002	Sensory analysis – Methodology – General guidance for measuring odour, flavour and taste detection thresholds by a three-alternative forced-choice (3-AFC) procedure
ISO 13302:2003	Sensory Analysis – Methods for assessing modifications to the flavour of foodstuffs due to packaging
ISO 16000-30:2014	Indoor air – Part 30: Sensory testing of indoor air
ISO 16657:2006	Sensory analysis – Apparatus- Olive oil tasting glass
ISO 16820:2004	Sensory analysis – Methodology- Sequential analysis
ISO 22308:2005	Cork stoppers – Sensory analysis





- ISO 22935-1:2009 (IDF 99-1: 2009) Milk and milk products – Sensory analysis – Part 1: General guidance for the recruitment, selection, training and monitoring of assessors
- ISO 22935-2:2009 (IDF 99-2: 2009) Milk and milk products – Sensory analysis – Part 2: Recommended methods for sensory evaluation
- ISO 22935-3:2009 (IDF 99-3: 2009) Milk and milk products – Sensory analysis – Part 3: Guidance on a method for evaluation of compliance with product specifications for sensory properties by scoring
- ISO 29842:2011 Sensory analysis – Methodology – Balanced incomplete block designs

ASTM – Sensory Evaluation Standards, ASTM Headquarters, PA. USA

www.astm.org/standards

Standarder à jour per mars 2015

Food and Beverage Evaluation

- E1083-00(2011) Standard Test Method for Sensory Evaluation of Red Pepper Heat
- E1346-90(2010) Standard Practice for Bulk Sampling, Handling, and Preparing Edible Vegetable Oils
- E1395-90(2011) Standard Test Method for Sensory Evaluation of Low Heat Chillies
- E1396-90(2011) Standard Test Method for Sensory Evaluation of Oleoresin Capsicum
- E1627-11 Standard Practice for Sensory Evaluation of Edible Oils and Fats
- E1810-12 Standard practice for Evaluating Effects of Contaminants on Odor and Taste of Exposed Fish
- E1871-10 Standard Guide for Serving Protocol for Sensory Evaluation of Foods and Beverages
- E1879-00(2010) Standard Guide for Sensory Evaluation of Beverages Containing Alcohol
- E2892-15 Standard Test Method for Odor and Flavor Transfer from Materials in Contact with Municipal Drinking Water

Fundamentals of Sensory

- E544-10 Standard Practices for Referencing Suprathreshold Odor Intensity
- E679-04(2011) Standard Practice for Determination of Odor and Taste Thresholds By a Forced-Choice Ascending Concentration Series Method of Limits
- E1432-04(2011) Standard Practice for Defining and Calculating Individual and Group Sensory Thresholds from Forced-Choice Data Sets of Intermediate Size
- E1885-04(2011) Standard Test Method for Sensory Analysis – Triangle Test
- E2139-05(2011) Standard Test Method for Same-Different Test
- E2164-08 Standard Test Method for Directional Difference Test
- E2263-12 Standard Test Method for Paired Preference Test
- E2610-08(2011) Standard Test Method for Sensory Analysis – Duo-Trio Test
- E2943-14 Standard Guide for Two-Sample Acceptance and preference Testing with Consumers





Personal Care and Household Evaluation

- E1207-14 Standard Guide for Sensory Evaluation of Axillary Deodorancy
- E1490-11 Standard Guide for Two Sensory Descriptive Analysis Approaches for Skin Creams and Lotions
- E1593-13 Standard Guide for Assessing the Efficacy of Air Care Products in Reducing the Perception of Indoor Malodor
- E2049-12 Standard Guide for Quantitative Attribute Evaluation of Frafrance/Odors for Shampoos and Hair Conditioners by Trained Assessors
- E2082-12 Standard Guide for Descriptive Analysis of Shampoo Performance
- E2346/ Standard Guide for Sensory Evaluation of Household Hand Surface-Cleaning
- E2346M-14 Products with Emphasis on Spray Triggers

Sensory Applications – General

- E460-12 Standard Practice for Determining Effect of Packaging on Food and Beverage Products During Storage
- E619-9 Standard Practice for Evaluating Foreign Odors in Paper Packaging
- E1870-11 Standard Test Method for Odor and Taste Transfer from Polymeric Packaging Film
- E1958-12 Standard Guide for Sensory Claim Substantiation
- E2299-13 Standard Guide for Sensory Evaluation of Products by Children and Minors
- E2454-05(2011) Standard Guide for Sensory Evaluation Methods to Determine the Sensory Shelf Life of Consumer Products
- E2609-08 Standard Test Method for Odor or Flavor Transfer or Both from Rigid Polymeric Packaging

Sensory Theory and Statistics

- E1697-05(2012) Standard Test Method for Unipolar Magnitude Estimation of Sensory Analyses
- E1909-13 Standard Guide for Time-Intensity Evaluation of Sensory Attributes
- E2262-03(2014) Standard Practice for Estimating Thurstonian Discriminant Distances

Terminology

- E253-15 Standard Terminology Relating to Sensory Evaluation of Materials and Product





REFERANSER OG RELEVANT LITTERATUR

- Albinson, B., Wendin, K., Åström, A. (2013). Handbok i sensorisk analys. SIK rapport 470. ISBN 978-91-7290-322-7
- Alm, S., (2010). Å måle barns holdninger. Økonomisk Fiskeriforskning 19/20 : 24-35
- Almli, V. L., Næs, T., Sulmont-Rossé, C., Enderli, G., Issanchou, S., & Hersleth, M. (2011). Consumers' acceptance of innovations in traditional cheese. A comparative study in France and Norway. *Appetite*, 57, 110-120.
- Almli, V.L. & Hersleth, M. (2012). Salt replacement and injection salting in smoked salmon evaluated from descriptive and hedonic sensory perspectives. *Aquaculture International*, (DOI) 10.1007/s10499-012-9615-4.
- Amerine, M.A., Pangborn, R.M., Roessler, E.B. (1965). *Principles of Sensory Evaluation of Food*. Academic, New York
- Bartoshuk L.M. (2000). Comparing Sensory Experiences Across Individuals: Recent Psychophysical Advanced Illuminate Genetic Variation in Taste perception. *Chem Senses* 25: 447-460,
- Bech, S., Hamberg, R., Nijenhuis, M., Teunissen, C., Looren de Jong, H., Houben, P., Pramanik, S. K. (1996). «The RaPID Perceptual Image Description Method (RaPID)», SPIE Conference «Human Vision and Electronic Imaging», 29 January - 1 February, California, 1996, paper 2657
- Bitnes, J., Martens, H., Ueland, Ø., Martens, M., (2007). Longitudinal study of taste identification of sensory panelists: Effect of aging, experience and exposure. *Food Quality and Preference*, 18, 230-241.
- Bitnes, J., Rødbotten, M., Lea, P., Ueland, Ø., Martens, M., (2007). Effect of product knowledge on profiling performance comparing various sensory laboratories. *Journal of Sensory Studies*, 22, 60-80
- Bitnes, J., Ueland, Ø., Møller, P., Martens, M., (2009). Reliability of sensory assessors: Issues of complexity. *Journal of Sensory Studies*, 24, (1),
- Broman, N. O., (1995). Visual impressions of features in Scots pine wood surfaces: A comparative study. *Forest Products Journal* 45 (3), 61-66
- Civille, G.V., Lyon, B. (1996). *Aroma and flavour lexicon for sensory evaluation: terms, definitions, References, and examples*. ASTM data series publication; DS 66, ASTM, West Conshohocken, PA, USA
- Cooper, R.G., (1993). *Winning at New Products. Accelerating the Process from Idea to Launch*. Second Ed. Addison-Wesley Publishing Company, Inc., 95-120
- Dairou, V., Sieffermann, J.M., (2002) A comparison of 14 jams characterized by conventional profile and a quick original method, the Flash profile. *Journal of Food Science* 67 (2), 826-834
- Ennis, D.M., (1993). The Power of Sensory Discrimination Methods, *Journal of Sensory Studies* 353-370
- Ennis, D.M., Rousseau, B. (2015). *Technical Report: Identifying and removing sources of bias in product tests and surveys*. The Institute for Perception, Newsletter Spring.
- Forsgren, G., Frisell, H., Ericsson, B., (1999). Taint and odour related quality monitoring of two food packaging board products using gas chromatography, gas sensors and sensory analysis. *Nordic pulp and Paper Research Journal*, 14 (1), 5-16





- Forskrift 1993-12-21-1381 om materialer og gjenstander i kontakt med næringsmidler (matkontaktforskriften)
- Forskrift 2001-12-04-1372 om vannforsyning og drikkevann (drikkevannsforskriften)
- Forskrift 2004-10-04-1316 om naturlig mineralvann og kildevann
- Forskrift 2008-12-22-1623 om næringsmiddelhygiene (næringsmiddelhygieneforskriften)
- Forskrift 2008-12-22-1624 om særlige hygieneregler for næringsmidler av animalsk opprinnelse (animaliehygieneforskriften)
- Forskrift 2011-06-06-668 om tilsetningsstoffer til næringsmidler
- Forskrift 2013-06-28-844 om kvalitet på fisk og fiskevarer,
- Gawel, R., A. Oberholster og I. L. Francis (2000). «A 'Mouth-feel Wheel': terminology for communicating the mouthfeel characteristics of red wine." *Australian Journal of Grape and Wine Research* 6(3): 203-207.
- Hall, G., Wendin, K., (2007). Sensory design for Foods for elderly. *Annals of Nutrition and Metabolism*. Vol 51, issue 6, 502-503
- Heiniö, R.-L., Oksman-Caldenty, K.-M., Latva-Kala, K., Lehtinen, P., Poutanen, K., (2001). Effect of drying treatment conditions on sensory profile of germinated oat. *Cereal Chemistry* 78 (6), 707-714
- Heiniö, R.-L., Lehtinen, P. Oksman-Caldenty, K.-M., Poutanen, K., (2002). Differences between sensory profile and development of rancidity during long-term storage of native and processed oat. *Cereal Chemistry* 79 (3), 367-375
- Helgesen, H., Solheim, R., Næs, T (1997). Consumer preference mapping of dry fermented lam sausages. *Food Quality and Preference* 8 (2), 97-109.
- Hersleth, M., Ilseng, M.A., Martens, M., Næs, T. (2005). Perception of cheese - a comparison of quality ratings by experts, descriptive profiling by a trained panel and consumer response. *Journal of Food Quality*, Vol 28, pp 333-349.
- Hersleth, M., Mevik, B.-H., Næs, T., Guinard, J.-X. (2003). Effect of contextual factors on liking for wine – use of robust design methodology. *Food Quality and Preference*, Vol 14, 7, pp 615-622.
- Hersleth, M., Monteleone, E., Segtnan, A. and Næs, T. (2015). Effects of evoked context on consumers' responses to intrinsic and extrinsic product attributes in dry cured ham. *Food Quality and Preference*, 40, 191-198.
- Hersleth, M., Næs, T., Guerrero, L., Claret, A., Recchia, A., Dinnella, C., & Monteleone, E. (2013). Consumer Perception of Dry-Cured Ham – A Cross-Cultural Study in Italy, Norway and Spain. *Journal of Sensory Studies*, 28(6), 450-466, doi: 10.1111/joss.12068.
- Hersleth, M., Rødbotten, M., (2009). Smak og smaksutvikling. I: Holte, A., Wilhelmsen, B.U., Mat og helse i skolen, en fagdidaktisk innføring. Fagbokforlaget. ISBN 978-82-450-0824-1
- Hirst, D., Næs, T. (1994). A graphical technique for assessing differences among a set of rankings. *J. Chemometrics* 1994, 8, pp 81-93
- Hochberg, Y., Tamhane, A.C. (1987). Multiple Comparison Procedures, *J. Chemometrics* 1994, 8, pp
- Hochberg, J., Tamhane, A.C., (1987). Multiple Comparison Procedures New York: John Wiley & Sons, ISBN 0-471-82222-1
- Høibø, O., Nyrud, A.Q. (2010). Consumer perception of wood surfaces: the relationship between stated preferences and visual homogeneity. *Journal of Wood Science* 56(4):276-283.
- Jonsson O, Lindberg S, Roos A, Hugosson M, Lindström M. (2008). Consumer perceptions and preferences on solid wood, wood-based panels and composites – A repertory grid study. *Wood and Fiber Science* 40(4): 663-678.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Strauss and Giroux, ISBN 978-0374275631. Norsk utgave: Tenke, fort og langsomt, Oslo: Pax forlag, 2012
- Kilcast, D.m (2010). *Sensory Analysis for Food and Beverage Quality Control. A Practical Guide*. Woodhead Publishing Ltd. ISBN: 978-1-84569-476-0





- Koch, I. S., M. Muller, E. Joubert, M. van der Rijst og T. Næs (2012). «Sensory characterization of rooibos tea and the development of a rooibos sensory wheel and lexicon.» *Food Research International* 46(1): 217-228.
- Kommisjonsforordning (EF) nr. 2073/2005 av 15. november 2005 om mikrobiologiske kriterier for næringsmidler
- Kroll, B. J. (1990). Evaluation rating scales for sensory testing with children. *Food Technology*, 44 (11), 78-80.
- Land D.G., (1983). What is sensory quality? I: Sensory quality in foods and beverages: definition, measurement and control, Editor: Williams & Atkin, Chichester: Ellis Horwood
- Larssen, W. E. og M. Hersleth (2014). «Sensorisk kvalitet på marine oljer.» *Møreforskning*. MA 14/16 51
- Lawless H., Klein B.B., (1991). *Sensory Science Theory and Practice*, Ellis Horwood, Chichester, England
- Lawless H., Malone G., (1986). The discriminative efficiency of common scaling methods, *J. Sensory Studies*, 1, pp.85-98
- Lawless, H.T., Heymann, H. (2010). *Sensory Evaluation of Food. Principles and practices*. Second Ed. Springer. ISBN 978-1-4419-6487-8
- Lea, P., Næs, T., Rødbotten, M., (1998). Analysis of variance for sensory data. Wiley & Sons. ISBN 0-471-96750-5
- Lea, P., Rødbotten, M., Næs, T., (1995). Measuring validity in sensory analysis. *Food Quality and Preference*, 1995, 6, pp 321-326
- Lindberg S., Roos A., Kihlstedt A., Lindström M. (2013). A product semantic study of the influence of the sense of touch on the evaluation of wood-based materials. *Materials and Design* 52 (2013):300-307.
- Little. A.D., (1958). *Flavor Research and Food Acceptance*, Inc., ed. Reinhold, N.Y.
- Lundgren B. (1991). *Handbok i sensorisk analys*. SIK-rapport Nr 470
- Lyon D., Francombe M.A., Hasdell T.A., Lawson K. (1992). *Guidelines for Sensory Analysis in Food Product Development and Quality Control*, Chapman & Hall, London
- Lyon, D. H. og M. P. Watson (1994). «Sensory profiling: a method for describing the sensory characteristics of virgin olive oil.» *Grasas y aceites* 45(1-2): 20-25.
- MacFie, H. J., Bratchell, N., Greenhoff, K., & Vallis, L. V. (1989). Designs to balance the effect of order of presentation and first-order carry-over effects in hall tests. *Journal of Sensory Studies*, 4, 129-148.
- Martens, M., (2009). Maten er ikke sunn før den er spist. *Norsk Tidsskrift for Ernæring*, 3, 4-8
- Martens, M. (1999). A philosophy for sensory science. *Food Quality and Preference*, 10, 4-5,233-244
- Martens, M., (2009). Mat, fysiologi og psykologi: Hva styrer det vi spiser? *Meieriposten*, 3, 82-85
- Martens, M., Skaret, J., Lea, P., (2010). Sensory perception of Food Products affected by different music genres. Poster for Fourth European Conference in Sensory and Consumer Research: A sense of Quality. Spain
- Martens, M., Veflingstad, S.R., Plathe, E., Bertrand, D., Martens, H., (2010). A sensory scientific approach to visual pattern recognition of complex biological systems. *Food Quality and Preference*, 21, 977-986
- Martens, M., Risvik, E., Martens, H., (1994). Matching sensory with instrumental analysis. In: *Understanding Natural Flavours* (J.R. Piggott and A. Patterson, eds) Blackie Academic & Professional, UK, 60-76
- Meilgaard M.C., Civille B., Carr T. (2006). *Sensory Evaluation Techniques*, (4 ed.) Boca Raton CRC Press, Inc.
- Meiselman, H. L. (1996). The contextual basis for food acceptance, food choice and food intake: the food, the situation and the individual. In H. L. Meiselman & H. J. H. MacFie (Eds.), *Food choice acceptance and consumption* (pp. 139-263). Glasgow: Blackie Academic and Professional
- Mojet, J. og S. de Jong (1994). «The sensory wheel of virgin olive oil.» *Grasas y aceites* 45(1-2): 42-47.
- Monteleone, E. og S. A. Langstaff (2014). *Olive Oil Sensory Science*. First edition. Wiley Blackwell.





- Moruzzi G. and Magoun H.W (1949), Brain stem reticular formation and activation of EEG. *Electroencephalography and Clinical Neurophysiology* Vol 1, issue 1 – 4, pp 455 -473
- Munoz A.M., Civille G.V., Carr B.T. (1992). *Sensory Evaluation in Quality Control*, VanNostrandReinhold, New York, New York: John Wiley & Sons, ISBN 0-471-82222-1
- Nijenhuis, M., Hamberg, R., Teunissen, C., Bech, S., Looren de Jong, H., Houben, P., Pramanik, S.K. (1997). «Sharpness, sharpness related attributes, and their physical correlates» SPIE Conference «Very High Resolution and Quality Imaging. II», 10 - 11 February, California, 1997, paper 3025
- Nyrud, A.Q., Roos, A., Rødbotten, M. (2008). Product attributes affecting consumer preference for residential deck materials. *Canadian Journal of Forest Research* 38(6):1385-1396.
- Næs, T. (1990). Handling individual differences between assessors in sensory profiling. *Journal of Food Quality and Preferences*, 2, 187-199.
- Næs, T., Brockhoff, P., Tomic, O. (2010). *Statistics for sensory and consumer science*. Wiley, UK
- Næs, T., Langsrud, Ø. (1998). Fixed or random assessors in sensory profiling. *Food Quality and Preference*. Vol 9. Issue3, 145-154
- PanelCheck: www.panelcheck.com
- Paulsen, M.T., Ueland, Ø., Nilsen, A.N., Öström, Å., Hersleth, M., (2012). Sensory perception of salmon and culinary sauces – An interdisciplinary approach. *Food Quality and preference* 23, 99-109
- Paulsen, M.T., Næs, T., Ueland, Ø., Rukke, E.-O., Hersleth, M., (2013). Preference mapping of salmon-sauce combinations: The influence of temporal properties. *Food Quality and preference* 27, 120-127
- Peryam, D. R. & Pilgrim, F. J. (1957). Hedonic scale method of measuring food preferences. *Food Technology*, 11. 9-14.
- Piggott, J.R. (1991). Selection of Terms for Descriptive Analysis. In: *Sensory Science - Theory and Applications in Foods*. Ed. Lawless & Klein. (IFT basic symposium series) 920005, MerceL Dekker Inc.
- Reed D.R. and Knaapila A. (2010), Genetics of taste and Smell; Poisons and Pleasures. *Prog. Mol Biol Transl Sci* 2010; 94: 213-240.
- Risvik, E. (1994). Sensory Properties and Preferences. *Meat Science*, 36,67-77
- Risvik, E., McEwan, J.A., Colwill, J.S., Rogers, R., and Lyon, D. H. (1994). Projective mapping: A tool for sensory analysis and consumer research. *Food Quality and Preference* 5: 263-269.
- Roos A., Lindberg S., Kihlstedt A. (2013). A product semantic study of the influence of the sense of vision on the evaluation of wood-based materials. *Wood and Fiber Science* 45(4):1-10.
- Rødbotten, M., Kubberød, E., Lea, P., Ueland, Ø. (2004). A sensory map of the meat universe. Sensory profile of meat from 15 species. *Meat Science*, 68, 137-144
- Sander K. (2014). Hvordan tolker vi sansinntrykkene (Persepsjon og kognisjon). Kunnskapssenteret.com
- Shepherd G.M. (2012). *Neurogastronomi*, Columbia University Press, kap 1
- Shepherd, G.M., (2006). Review article, Smell images and the flavour system in the human brain. *Nature* 444, 316-321
- Stevenson R.J. 2012. The role of attention in flavor perception. *Flavour* 1:2
- Stone H., Sidel J.L. (1993). *Sensory Evaluation Practices*, Academic Press, London
- Swahn, J., Mossberg, L., Öström, Å., Gustafsson, I-B. (2012). Sensory description labels for food affect consumer product choice. *European Journal of Marketing*. Vol 46, issue 11/12, 1628-1646
- Tomic, O., Luciano, G., Nilsen, A., Hylidig, G., Lorensen, K., Næs, T. (2010). Analysing sensory panel performance in a proficiency test using the PanelCheck software. *European Food Research and Technology*. Volume 230, Issue 13, 497-511
- Tomic, O., Forde, C., Delahunty, C., Næs, T., (2013). Performance indices in descriptive sensory analysis – A complimentary screening tool for assessor and panel performance. *Food Quality and Preferences*, 28, pp 122-133





Undheim, J.O., (1991). Innføring i statistikk for samfunnsvitenskapelige fag. Universitetsforlaget. ISBN 82-00-07564

Varela, P. & Ares, G. (2014). *Novel Techniques in Sensory Characterization and Consumer Profiling*, CRC Press, Tylor & Francis Group, Boca Raton.

Wibeck, V. (2010). Fokusgrupper: om fokuserade gruppeintervjue som undersøkingsmetode. Lund. ISBN 978-91-44-05856-6

www.qim-eurofish@com.

www.lovdatab.no

www.mattilsynet.no

www.panelcheck.com.

Wysocki C.J. and Beauchamp G.K. (1984). Ability to smell androstenone is genetically determined. *Proc. Natl. Acad. Sci* Vol 81 pp 4899-4902

Øvrum, A., Alfnes, F., Almli, V. L., & Rickertsen, K. (2012). Health information and diet choices: results from a cheese experiment. *Food Policy*, 37, 520–529.





VEDLEGG

Vedlegg 16a Skjema

Skjemaene som følger er eksempler som kan benyttes ved analyse av smak i produkter. De samme skjemaene kan også brukes ved de aller fleste typer av produkter, eventuelt med en endret instruksjon til dommerne om de skal lukte, se, føle eller høre på prøven avhengig av egen-skapen som skal bedømmes.

Eksempler på koding av prøver er delvis skrevet inn i noen av skjemaene.





IDENTIFIKASJON AV GRUNNSMAKENE

Navn: _____ Dato: _____

Dommer nr.: _____

Du får først 6 kjente prøver:

Vann, søtt, surt, salt, bittert og umami

Skyll munnen med vann før du smaker på første prøve og deretter mellom hver prøve.

Smak så på de navngitte prøvene og forsøk og huske hvordan de smaker.

Skyll munnen før du smaker på neste prøve! Spytt ut både prøven og skyllevannet, så blir du ikke så fort trøtt.

Sett så disse prøvene til side og gå ikke tilbake for å smake på dem igjen.

Deretter får du kodede prøver. Smak på disse prøvene, én om gangen i den rekkefølgen de står i tabellen under. Angi for hver prøve om de smaker vann, søtt, surt, salt, bittert eller umami. Når du er ferdig med én prøve, skyller du munnen og går videre til neste.

OBS! Du får ikke gå tilbake til foregående prøve for å smake på den, og du får ikke forandre ditt svar

Prøvene er servert i tilfeldig rekkefølge, så det lønner seg ikke å forsøke å finne noe system.

Kode	Smak	Kode	Smak	Kode	Smak

Anmerkninger:





For panelleder

TRIANGELTEST

Bedømmelse av Dato.....

Prøve A =

Kode 357 og 453

Prøve B =

Kode 831 og 107

KAP. 13 SENSORISKE NETTVERK

Dommer	Serverings- ordning	Koder	Rett/feil
1	A-A-B	357-453-831	
2	B-A-B	831-357-107	
3	B-B-A	831-107-357	
4			
5			
6			
7			
8			
9			
10			
11			
12			
	N:..... Antall rett:..... % rett:..... Signifikans:.....		





TRIANGELTEST

Navn: _____ Dato: _____

Dommer nr.: _____

Prøveomgang nr:

Skyll munnen med vann før du begynner smakingen.

Du får tre prøver hvorav to er like. Se, lukt og smak på dem, fra venstre til høyre. Skyll munnen med vann mellom hver smaking. Du skal ikke gå tilbake til forrige prøve for en ny smaking. Marker med et kryss for den prøven du mener er ulik de to andre. Dersom du ikke merker forskjell, må du gjette.

Prøvekode	357	107	831
Hvilken prøve er ulik de to andre?			





For panelleder

PARTEST

Bedømmelse av Dato.....

Prøve A = Kode 556

Prøve B = Kode 390

KAP. 13 SENSORISKE NETTVERK

Dommer	Serverings rekkefølge	Koder	Svar	Kommentarer
1	A-B	556-390		
2	B-A	390-556		
3	A-B	556-390		
4	B-A	390-556		
5				
6				
7				
8				
9				
10				
11				
12				
	N:..... Antall A:..... Antall B:..... Signifikans:.....			





PARTEST

Dommer nr.: _____

Dato: _____

Du får servert to prøver som skal bedømmes mot hverandre.

Skyl munnen med vann før du begynner å smake og mellom hver smaking. Smak på prøvene (dersom det er smak som skal registreres).

Dersom du ikke merker forskjell, må du gjette.

Kryss av for den prøven du oppfatter som..... (*søtest*)

KODE
556KODE
390



For panelleder

DUO-TRIO-TEST (balansert referansemetode)

Bedømmelse av Dato.....

Prøve A = Kode 556

Prøve B = Kode 390

KAP. 13 SENSORISKE NETTVERK

Dommer	Serverings rekkefølge	Koder	Svar	Kommentarer
1	AR-A-B	556-390		
2	AR-B-A	390-556		
3	BR-A-B	556-390		
4	BR-B-A	390-556		
5	AR-A-B	556-390		
6				
7				
8				
9				
10				
11				
12				
	N:..... Antall A:..... Antall B:..... Signifikans:.....			





DUO-TRIO-TEST

Dommer nr.: _____

Dato: _____

Skyll munnen med vann før du begynner smakingen.

Du får tre prøver hvorav to er like. Den første prøven er merket REF og de to andre er merket med kodenummer. Se, lukt og smak på dem, først på referansen (REF) og deretter på de to andre, fra venstre til høyre. Skyll munnen med vann mellom hver smaking. Du skal ikke gå tilbake til forrige prøve for en ny smaking.

Dersom du ikke merker forskjell, må du gjette.

Kryss av for den prøven du oppfatter som ulik referansen.

KODE
390

KODE
556

Hva er det som utgjør forskjellen mellom denne prøven og referansen?

Den ulike prøven er:
.....
.....
.....





For panelleder

RANGORDNINGSTEST

Bedømmelse av Dato.....

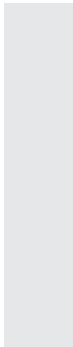
Prøve	Koder for gjentak		
	1	2	3
A: 10%	448	312	885
B: 15%	357	400	724
C: 20%	118	567	273
D: 25%	569	021	356
E: 30%	333	894	270

KAP. 13 SENSORISKE NETTVERK

1. Gjentak

Dommer	Serverings rekkefølge	Koder
1	C-B-A-E-D	118-357-448-333-569
2	E-A-D-B-C	333-448-569-357-118
3	D-B-C-A-E
4	C-E-B-D-A
5		
6		
7		
8		
9		
10		
11		
12		





RANGORDNINGSTEST

Navn: _____ Dato: _____

Dommer nr.: _____

Ranger prøvene etter økende grad av.....(*søthet*)

Smak på prøvene fra venstre mot høyre, deretter kan du smake flere ganger på prøvene.

Ordning	1 Minst	2	3	4	5 Mest
Kode					

Kommentarer:





For panelleder

RANGORDNINGSTEST – RESULTATER

Dommer	Produkt					Sum rangordning for hver dommer
	1-A	2-B	3-C	4-D	5-E	
1	2	1	3	5	4	
2	1	3	2	4	5	
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
Sum rang						

KAP. 13 SENSORISKE NETTVERK





KVALITETSKONTROLLTEST

BEDØMMELSE AV OLJE

Prøve nr.:

Navn:

Dato:

Bedøm prøven og angi eventuelle sensoriske «feil». Sett kryss i den/de rutene som angir hva som er feil ved prøven (eks.: et kryss ved mindre klar, betyr at du mener denne prøven har mindre klar farge enn ønsket. Et kryss ved for lite syrlig betyr at denne prøven er mindre syrlig enn ønsket i henhold til spesifikasjonen).

Sett deretter ring rundt tallet som angir din totale kvalitetsvurdering av prøvens utseende, lukt, smak og konsistens.

9-7 poeng = akseptabel kvalitet (9 poeng = «riktig/perfekt» kvalitet),

6-4 poeng = vurderes for "andresortering"

3-1 poeng = uakseptabel kvalitet.

Utseende	Lukt	Smak	Konsistens
<input type="checkbox"/> mindre* klar	<input type="checkbox"/> mindre bitter	<input type="checkbox"/> mindre bitter	<input type="checkbox"/> mindre tykk
<input type="checkbox"/> mindre gul	<input type="checkbox"/> mer bitter	<input type="checkbox"/> mer bitter	<input type="checkbox"/> mer tykk
<input type="checkbox"/> mer gul	<input type="checkbox"/> mer harsk	<input type="checkbox"/> mer harsk	<input type="checkbox"/> mindre jevn
	<input type="checkbox"/> mindre syrlig	<input type="checkbox"/> mindre syrlig	<input type="checkbox"/> mer jevn
	<input type="checkbox"/> mer syrlig	<input type="checkbox"/> mer syrlig	
	<input type="checkbox"/> mindre metall	<input type="checkbox"/> mindre metall	
	<input type="checkbox"/> mer metall	<input type="checkbox"/> mer metall	

*Mindre og mer av en egenskap bedømmes i forhold til spesifikasjonen (referanse).

UTSEENDE

1	2	3	4	5	6	7	8	9
maks ulik referansen								lik referansen

LUKT

1	2	3	4	5	6	7	8	9
maks ulik referansen								lik referansen

SMAK

1	2	3	4	5	6	7	8	9
maks ulik referansen								lik referansen

KONSISTENS

1	2	3	4	5	6	7	8	9
maks ulik referansen								lik referansen





Prinsippet ved dette skjemaet kan nyttes ved Just-About-Right og beskrivende- og kvalitetskontroll-metode

BEDØMMELSE AV OLJE

Dommer: _____ Dato: _____

Kode: _____

Bedøm prøvene i forhold til referansen.

-3 = mye mindre enn referansen

R = som referansen

+3 = mye mer enn referansen

LUKT

Syrlig 1

-3 -2 -1 R +1 +2 +3

Fisk 2

-3 -2 -1 R +1 +2 +3

Harsk 3

-3 -2 -1 R +1 +2 +3

SMAK

Syrlig 4

-3 -2 -1 R +1 +2 +3

Metall 5

-3 -2 -1 R +1 +2 +3

Fisk 6

-3 -2 -1 R +1 +2 +3

Harsk 7

-3 -2 -1 R +1 +2 +3





BESKRIVENDE TEST

BEDØMMELSE AV LEVERPOSTEI

Egenskapsforklaringer

Hvithet	Farge bedømt på overflaten av et ferskt snitt etter NCS-systemet Ingen intensitet = ingen hvithet, prøven har svarttone eller max kulørt Tydelig intensitet = hvit
Fargetone	Farge bedømt på ferskt snitt etter NCS-systemet «ingen» intensitet = gulaktig = gul Y50 R «tydelig» intensitet = rødaktig = gul Y70R
Krydderlukt	Lukt av krydder og urter, for eksempel pepper, muskat Ingen intensitet = ingen krydderlukt Tydelig krydderlukt = tydelig lukt av krydder
Saltsmak	Relatert til grunnsmaken salt (NaCl) Ingen saltsmak = ingen smak av salt Tydelig saltsmak = tydelig smak av salt
Leversmak	Smak av svin /storfelever Ingen leversmak = ingen smak av lever Tydelig leversmak = tydelig smak av lever
Fethet	Fornemmelse av fett i munnen Ingen fethet = ingen fornemmelse av fett i munnen Tydelig fethet = tydelig fornemmelse av fett i munnen





BESKRIVENDE TEST

BEDØMMELSE AV LEVERPOSTEI

Dommer: _____ Dato: _____

Kode: _____

Hvithet	1	----- Ingen	----- tydelig
Fargetone	2	----- Gulaktig	----- Rødaktig
Krydderlukt	3	----- Ingen	----- tydelig
Saltsmak	4	----- Ingen	----- tydelig
Leversmak	5	----- Ingen	----- tydelig
Fethet	6	----- Ingen	----- tydelig

KAP. 13 SENSORISKE NETTVERK





Vedlegg 16b Statistiske tabeller

Til utregning av verdiene i de følgende tabeller er benyttet SYSTAT for DOS Version 6.0 fra SYSTAT Software, Inc., 1735 Tehnology Dr., Ste. 430, San Jose, CA 95110, USA. Følgende funksjoner ble benyttet: For normalfordelinga: funksjonene ZCF og ZIF, for χ^2 -fordelinga XIF, for t-fordelinga TIF, for F-fordelinga FIF. (Nyere versjoner av programmet er tilgjengelige)

Tabell 1 – Binomisk fordeling

Minste antall korrekte identifikasjoner som kreves for å forkaste H_0 : Prøvene er like. N = Antall «forsøk» (vanligvis: antall dommere eller forbrukere)

P1: Partest (ensidig) på nivå 0,10.

P2: Partest (tosidig) på nivå 0,10 eller ensidig på nivå 0,05.

P3: Partest (tosidig) på nivå 0,05 eller ensidig p nivå 0,025.

T1: Triangeltest på nivå 0,10.

T2: Triangeltest på nivå 0,05.

For Duo-trio-test benyttes tabellen for ensidig partest.





Tabell 1 Binomisk fordeling

N	P1	P2	P3	T1	T2
5	5	5		4	4
6	6	6	6	5	5
7	6	7	7	5	5
8	7	7	8	5	6
9	7	8	8	6	6
10	8	9	9	6	7
11	9	9	10	7	7
12	9	10	10	7	8
13	10	10	11	8	8
14	10	11	12	8	9
15	11	12	12	8	9
16	12	12	13	9	9
17	12	13	13	9	10
18	13	13	14	10	10
19	13	14	15	10	11
20	14	15	15	10	11
21	14	15	16	11	12
22	15	16	17	11	12
23	16	16	17	12	12
24	16	17	18	12	13
25	17	18	18	12	13
30	20	20	21	14	15
35	22	23	24	16	17
40	25	26	27	18	19
45	28	29	30	20	21
50	31	32	33	22	23
60	36	37	39	26	27
70	41	43	44	29	31
80	47	48	50	33	35
90	52	54	55	37	38
100	57	59	61	40	42
110	63	65	66	44	46
120	68	70	72	48	50
130	73	75	77	51	53
140	79	81	83	55	57
150	84	86	88	58	61





Tabell 2: Studentisert variasjonsbredde (Studentized Range) for uendelig mange frihetsgrader og 3 nivåer av α . Antall prøver som sammenliknes er k.

K	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,10$
3	4,120	3,315	2,902
4	4,403	3,633	3,240
5	4,603	3,858	3,478
6	4,757	4,030	3,661
7	4,882	4,170	3,808
8	4,987	4,286	3,931
9	5,078	4,387	4,037
10	5,157	4,474	4,129
11	5,227	4,552	4,211
12	5,290	4,622	4,285
13	5,348	4,685	4,351
14	5,400	4,743	4,412
15	5,448	4,796	4,468
16	5,493	4,845	4,519
17	5,535	4,891	4,568
18	5,574	4,934	4,612
19	5,611	4,974	4,654
20	5,645	5,012	4,694





Tabell 3: Kji-kvadrat-fordelinga for 3 nivåer av α . I forbindelse med Friedman's test er antall frihetsgrader (Fr.gr. i tabellen), antall prøver (eller sorter) minus 1.

Eksempel: Har vi 5 sorter, må vi inn i tabellen for Fr.gr. = 4.

Fr.gr.	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$
1	2,7055	3,8415	6,6349
2	4,6052	5,9915	9,2103
3	6,2514	7,8147	11,3449
4	7,7794	9,4877	13,2767
5	9,2364	11,0705	15,0863
6	10,6446	12,5916	16,8119
7	12,0170	14,0671	18,4753
8	13,3616	15,5073	20,0902
9	14,6837	16,9190	21,6660
10	15,9872	18,3070	23,2093
11	17,2750	19,6751	24,7250
12	18,5493	21,0261	26,2170
13	19,8119	22,3620	27,6883
14	21,0641	23,6848	29,1412
15	22,3071	24,9958	30,5779
16	23,5418	26,2962	31,9999
17	24,7690	27,5871	33,4087
18	25,9894	28,8693	34,8053
19	27,2036	30,1435	36,1909
20	28,4120	31,4104	37,5662





Vedlegg16c Enkel sannsynlighetsregning

Lesere som enten er fortrolige med elementær sannsynlighetsregning, eller er villige til å akseptere utsagn om sannsynligheter som blir framsatt seinere i kapitlet, kan med god samvittighet hoppe over dette delkapitlet.

I sin enkleste form er sannsynlighet definert som antall gunstige tilfeller dividert med antall mulige tilfeller:

$$\text{Sannsynlighet} = \frac{\text{Antall gunstige tilfeller}}{\text{Antall mulige tilfeller}}$$

Dette er vanlig å demonstrere med eksempler fra myntkast, terningspill og kortspill, siden dette er begreper de fleste har lett for å forholde seg til. Denne tradisjonen er det ingen grunn til å gi slipp på, og vi bruker terningkast i det etterfølgende. Hva er sannsynligheten for å få 4 øyne i et kast med én terning? Hvis vi antar at terningen er en ordinær, 6-sidet, helt balansert terning, finnes det i utgangspunktet 6 mulige utfall; «1», «2», «3», «4», «5», «6». Bare ett av disse er gunstige når det gjelder å få 4 øyne, nemlig «4». Sannsynligheten for å få 4 øyne i et kast med en terning er derfor 1/6. Sannsynligheten for «1» er også 1/6, det samme er sannsynligheten for «2», «3», osv.

Det er vanlig at vi snakker om begivenheter i slike sammenhenger, vel vitende om at dette ikke er helt i tråd med bruken av ordet i dagligtale. I vår sammenheng er en begivenhet rett og slett noe som skjer, for eksempel at vi får «Kron» i et myntkast, eller at 154 av 197 forbrukere synes at suppe A er bedre enn suppe B, altså helt blottet for den betydningen av noe høytidelig som vi ofte forbinder ordet med i dagligtalen.

Beklageligvis er det sjelden vi får så enkle sannsynligheter som i terningeksemplet over å beregne. For å beregne sammensatte sannsynligheter trenger vi noen regneregler til. En av dem er enkel:

Sannsynligheten for at to uavhengige begivenheter inntreffer, er lik produktet av sannsynlighetene for de to enkeltbegivenhetene.





At det ikke er noen forbindelse mellom en terning og en kortstokk vil de fleste kunne akseptere. Sannsynligheten for at vi skal få «2» i et kast med en terning og at vi skal trekke en spar fra en kortstokk kan derfor beregnes som produktet av de to enkeltbegivenhetene. Den første sannsynligheten er $1/6$, og siden det i en vanlig kortstokk finnes 52 kort hvorav 13 er spar, er sannsynligheten for den siste begivenheten lik $13/52 = 1/4$. Den sammensatte sannsynligheten blir derfor $1/6 \times 1/4 = 1/24 = 0,0417$.

For å forenkle notasjonen, har man introdusert en sannsynlighetsfunksjon P (noen lærebøker opererer med Pr (engelsk: Probability) i stedet for P): $P(K)$ kan f.eks. bety Sannsynligheten for å få «kron» i et kast med en mynt. Regelen om at sannsynligheten for en begivenhet er lik 1 minus sannsynligheten for den omvendte begivenheten, kan symbolsk skrives slik:

$$P(A) = 1 - P(\bar{A})$$

der \bar{A} betyr den «omvendte» begivenheten av A , eller i sannsynlighetsteoretisk terminologi: *komplementet til A*. I terningeksemplet vil summen av sannsynlighetene for alle mulighetene er lik 1:

$$1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 1.$$

Det betyr for eksempel at sannsynligheten for å få «2», «3», «4», «5» eller «6» vil være 1 mindre enn sannsynligheten for å få «1», eller i mer stringent notasjon:

$$P("2" \cup "3" \cup "4" \cup "5" \cup "6") = 1 - P("1") = 1 - \frac{1}{6} = \frac{5}{6} = 0,8333$$

En annen viktig regneregul sier at sannsynligheten for at enten A skal inntreffe, eller B skal inntreffe, eller begge skal inntreffe, er lik sannsynligheten for at A skal inntreffe pluss sannsynligheten for at B skal inntreffe minus sannsynligheten for at begge skal inntreffe samtidig. Med symboler blir dette:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

der « \cup » leses «union» og « \cap » leses «snitt». Union og snitt har omtrent samme funksjon innen mengdelære som summasjon og multiplikasjon





innen aritmetikk. Regneregler som den over lar seg relativt enkelt illustrere ved hjelp av Venn-diagrammer, se figur 1. I figur 1b ser vi hvorfor vi må trekke fra $P(A \cap B)$ («Snittet mellom A og B» i mengdelæretterminologi) i formelen ovenfor: hvis vi tar med alt som er i A og deretter alt som er i B, så får vi med oss snittet mellom A og B 2 ganger. Følgelig må vi trekke fra dette en gang.

Sannsynligheten for at vi skal trekke en spar eller en konge fra en kortstokk, kan skrives slik:

$$\begin{aligned} P(\text{«Spar»} \cup \text{«Konge»}) &= P(\text{«Spar»}) + P(\text{«Konge»}) - P(\text{«Spar»} \cap \text{«Konge»}) \\ &= \frac{13}{53} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52} = 0,3077 \end{aligned}$$

Regelen om sannsynligheten for to uavhengige begivenheter som det ble referert til tidligere i avsnittet, lar seg symbolsk uttrykke slik:

$$P(A \cap B) = P(A) \times P(B).$$

Et klassisk eksempel som benytter flere av prinsippene beskrevet over, er det såkalte fødselsdagsparadokset: «hvor mange mennesker må vi ha samlet for at sannsynligheten for at minst 2 stykker har samme fødselsdag skal være minst $\frac{1}{2}$?» Mange vil tippe at det svaret må være i nærheten av $365/2 = 182,5$ (et typisk svar i system-1-tenkning i terminologien til Daniel Kahneman (2011)). For å finne svare etter «rett-fram-metoden», må vi først bestemme oss for hvor mange personer vi skal beregne sannsynligheten for, for eksempel 100. Så må vi beregne sannsynligheten for at 2 har samme fødselsdag, men også ta med sannsynligheten for at 3, 4, 5, .. har samme fødselsdag, eller at 2 deler en fødselsdag mens 2 andre deler en annen, og så videre nesten i det uendelige, og så sjekke om dette blir over eller under $\frac{1}{2}$. Er denne sannsynligheten over $\frac{1}{2}$ (hvilket den er), prøver vi oss med 99 personer og går nedover, helt til vi finner det minste antall personer etter mange timers regning.

I stedet snur vi spørsmålet på hodet, og spør: «hvor mange personer må vi ha samlet for at sannsynligheten for at ingen av dem har felles fødselsdag, skal være minst $\frac{1}{2}$?». Et par forenklinger må riktig nok





gjøres: at alle fødselsdager i året er like sannsynlige, at det ikke eksisterer skuddår, og at vi ikke på forhånd vet at det finnes tvillinger (eller flerlinger) i utvalget av personer. Ser vi på de to første personene, så kan den første ha hvilken som helst av årets 365 dager som fødselsdag, og den neste alle uten den dagen som den første har. Sannsynligheten for at ingen av de to har samme fødselsdag, er da:

$$1 - \frac{365}{365} \times \frac{364}{365} = 1 - \frac{364}{365} = \frac{1}{365} = 0,0027$$

som er langt under $\frac{1}{2}$. Vi prøver den 3. personen: vedkommende kan ikke ha felles fødselsdag med noen av de to første, altså:

$$1 - \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} = 1 - \frac{363 \times 364}{365 \times 365} = 1 - \frac{132132}{133225} = 0,0082$$

Og slik fortsetter vi, til den 22. personen:

$$1 - \frac{365}{365} \times \frac{364}{365} \dots \times \frac{344}{365} = 0,4757$$

og den 23.:

$$1 - \frac{365}{365} \times \frac{364}{365} \dots \times \frac{344}{365} \times \frac{343}{365} = 0,5073$$

Det paradoksale svaret er altså at det bare trengs en gruppe på 23 personer samlet for at sannsynligheten er større enn $\frac{1}{2}$ for at minst 2 skal ha samme fødselsdag!

For å beregne gunstige eller mulige tilfeller i litt mer kompliserte situasjoner, trenger vi et par hjelpemidler til. Ett av disse er å beregne hvor mange måter kan vi velge ut m elementer blant n elementer på. Svaret på det er gitt ved hjelp av den binomiske koeffisient « n over m », som skrives:

$$\binom{n}{m} = \frac{n!}{(n-m)!m!}$$

Her betyr $n!$ (uttales « n fakultet») produktet av de n første tallene, slik at: $n! = 1 \times 2 \times 3 \times \dots \times n$. Eksempelvis blir $5! = 1 \times 2 \times 3 \times 4 \times 5 = 120$.





Hvis vi nå vil beregne hvor mange måter vi kan velge ut 2 elementer blant 5 på, får vi ved innsetting i formelen:

$$\binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5!}{2!3!} = \frac{1 \times 2 \times 3 \times 4 \times 5}{(1 \times 2) \times (1 \times 2 \times 3)} = \frac{4 \times 5}{1 \times 2} = 10$$

Dette resultatet kan vi bekrefte ved å sette opp bokstavene a b c d e på et papir, og så skrive ned alle mulige måter å velge ut 2 bokstaver på. Det blir: ab, ac, ad, ae, bc, bd, be, cd, ce, de, tilsammen 10 måter.

Å velge ut 0 elementer fra en gruppe på m, kan gjøres på 1 måte, noe som ved første øyekast kan virke litt rart. Men hvis man skal gjøre beregninger selv, er det altså viktig å være klar over at

$$\binom{m}{0} = 1$$

Den binomiske koeffisienten benyttes i den binomiske fordelinga: denne fordelinga forteller oss hva sannsynligheten er for at det i løpet av n «forsøk» (f.eks. 10 kast med en mynt) skal være x forekomster av en viss begivenhet (f.eks. 8 «Kron») når det er en sannsynlighet på p (1/2 i eksemplet) for at begivenheten skal skje i løpet av et gitt antall forsøk. Med symboler blir dette:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

eller i det konkrete eksemplet:

$$P(X = 8) = \binom{10}{8} \left(\frac{1}{2}\right)^8 \left(1 - \frac{1}{2}\right)^{10-8} = \frac{9 \times 10}{1 \times 2} \left(\frac{1}{2}\right)^{10} = 0.044$$





I praksis vil vi som regel ha behov for enten å beregne den kumulative sannsynligheten (sannsynligheten for at X er mindre enn eller lik en viss verdi), eller at X er større enn en viss verdi. Den kumulative sannsynligheten er gitt ved:

$$P(X \leq x) = \sum_{i=0}^x \binom{n}{i} p^i (1-p)^{n-i}$$

Når vi snakker om den binomiske fordelinga, pleier vi å erstatte $P(X \leq x)$ med $B(x; n, p)$, altså:

$$B(x; n, p) = \sum_{i=0}^x \binom{n}{i} p^i (1-p)^{n-i}$$

Hvis vi vil beregne den omvendte sannsynligheten, må vi passe på at det omvendte av «mindre enn eller lik x » ikke er det samme som «større enn eller lik x », men «større enn eller lik $x + 1$ ». Da får vi i henhold til den generelle regelen $P(A) = P(\bar{A})$:

$$B(x; n, p) = 1 - \sum_{i=x+1}^n \binom{n}{i} p^i (1-p)^{n-i}$$

I stedet for å beregne sannsynligheten for at blant 100 forbrukere så er det 35 eller flere som foretrekker suppe A framfor suppe B, så beregner vi altså heller sannsynligheten for at færre enn 35 (det vil si 34 eller færre) foretrekker A, og så trekker dette fra 1:

$$\sum_{i=35}^{100} \binom{100}{i} \left(\frac{1}{2}\right)^i (1-\frac{1}{2})^{100-i} = 1 - \sum_{i=0}^{34} \binom{100}{i} \left(\frac{1}{2}\right)^i (1-\frac{1}{2})^{100-i} = 1 - 0,0009 = 0,9991$$

Denne binomiske fordelinga inngår som en viktig del i de statistiske testene som går under fellesbetegnelsen forskjellstester. I praksis er det som regel ikke nødvendig å gjøre disse beregningene for hånd: mens vi tidligere kunne benytte statistiske tabeller som var gjen-gitt i de fleste lærebøkene i sensorikk, vil vi nå bruke regneark eller statistikkprogram.





Vedlegg 16d Noen statistiske begreper

I statistiske formler bruker vi vanligvis X til å symbolisere en måling eller et tall; har vi mange, kaller vi dem gjerne X_1, X_2, X_3 opp til X_n , hvor n =antallet målinger som inngår, og hvor n kan være spesifisert hvis vi har en veldig konkret situasjon, eller n kan rett og slett bare stå for et generelt antall. Vil vi skrive summen av disse tallene, er det tungvint å skrive for eksempel summen av 1000 målinger som $X_1 + X_2 + X_3 + \dots$ og så videre opp til måling nr. 1000: X_{1000} . Ved hjelp av summasjonssymbolet skriver vi bare

$$\sum_{i=1}^{1000} X_i$$

Et mye brukt (og misbrukt) begrep er ordet gjennomsnitt. Ofte brukes betegnelsen middelværdi, eller bare middel. Vanligvis snakker man da om det aritmetiske gjennomsnittet: summen av alle observasjonene dividert med antallet av dem. Formelen ser slik ut:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Middelværdien er en måte å beskrive sentraltendensen i et datasett på, et slags sammenfattende mål. En vanlig måte å sammenfatte et datasett hvor 12 dommere har bedømt et visst antall egenskaper for 8 sorter i 3 gjentak, vil være å bruke gjennomsnittet over både dommere og gjentak når vi på en enkel måte skal beskrive de 8 sortene. For hver egenskap og sort vil vi altså beregne gjennomsnittet over de 24 målingene som ligger bak (8 dommere \times 3 gjentak).

Opptil flere varianter av gjennomsnittet finnes, her skal vi bare kort nevne medianen, som har en viss betydning innen de såkalt ikke-parametriske metodene – som i noen tilfeller baserer seg på rangeringer heller enn bedømmelser langs en skala. Medianen er den midterste målingen, det vil si den som er slik at halvparten er målingene er større enn medianen, den andre halvparten er mindre. Mens gjennomsnittet for tallene 3, 5, 6, 11, 1000 er 205, så er medianen 6. Siden sensoriske data er målt på en begrenset skala, vil vi aldri få så store forskjeller





på gjennomsnittet og medianen i praksis. Medianen egner seg best på datasett som ikke er symmetriske og har store avvik i den ene eller andre retningen (lønninger: noen kan tjene millioner, men ingen kan tjene mindre enn 0.)

Ofte er det ikke tilstrekkelig å beskrive et datasett ved hjelp av bare middelveidien. Vel så interessant er det å vite noe om hvordan dataene ligger spredd rundt denne middelveidien. En måte å beskrive slik spredning på, er ved hjelp av varians (S^2), eller standardavvik (S):

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$

På grunn av praktiske årsaker som faren for avrundingsfeil og det ineffektive i først å beregne middelveidien og deretter danne differenser og kvadrere disse, benyttes i praksis andre formler hvis man ønsker å gjøre disse beregningene «for hånd». I programvare som beregner standardavviket (regneark, statistiske program, lommeregner), brukes formler som gjør at det er tilstrekkelig å gjennomløpe datasettet en gang:

$$S^2 = \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - \frac{1}{n} \left(\sum_{i=1}^n X_i \right)^2 \right)$$

For å bruke en slik formel er det tilstrekkelig å akkumulere

$$\sum X_i \text{ og } \sum X_i^2$$

Definisjonen av varians/standardavvik som et mål for spredning er intuitivt fornuftig. Det som inngår, er enkeltobservasjonenes avvik fra middelveidien:

$$X_i - \bar{X}$$





får vi et mål for spredningen i hele datamaterialet. Men hvis denne summen alene hadde vært brukt til å definere spredning, ville positive og negative avvik ha kunnet nøytralisere hverandre, og «spredningen» kunne blitt 0 selv om intuisjonen hadde sagt oss noe annet. Å se på tallverdier (ignorere minustegn) i stedet:

$$\sum_{i=1}^n |X_i - \bar{X}|$$

slik at en verdi på 3 enheter større enn middelveiden ville ha gitt samme bidraget til spredningen som en verdi 3 enheter mindre enn middelveiden, hadde opplagt vært en forbedring. Men å opphøye differensene i 2. potens er enda bedre: potenser er enklere å ha med å gjøre enn tallverdier når statistiske egenskaper skal utledes og bevises.

Derfor er det fornuftig å la

$$\sum_{i=1}^n (X_i - \bar{X})^2$$

inngå i definisjonen. Men også denne definisjonen har en ulempe: hvis n er «stor», kan spredningen også bli «stor» selv om enkeltavvikene fra middelveiden er «små»: summen av mange små tall kan bli et stort tall! En mulig løsning på problemet er å se på gjennomsnittlig spredning, altså:

$$\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

At heller ikke denne definisjonen er helt perfekt er det vanskeligere å gi en elementær forklaring på. La oss heller bare akseptere at det av teoretiske statistiske årsaker er fornuftig å dividere med $n-1$ i stedet for med n . Dermed er vi tilbake til definisjonen tidligere i kapitlet.

Begrepene varians og standardavvik er egentlig likeverdige: variansen er standardavviket opphøyd i 2. potens (eller på en annen måte: standardavviket er kvadratrota av variansen). Standardavviket har den åpenbare fordelen at den måles i samme enhet som verdiene i datasettet.





Er det snakk om noen ulemper, må det være at det i formelen inngår et kvadratrottegn. (Et problem av rent teoretisk interesse er at vi ikke kan finne et forventningsrett estimat for standardavviket, men det kan vi for variansen.)





STIKKORD

A

A ikke A 96
Adapsjon 45, 71
Affektiv 79, 144
Akkreditering **166-169**, 171
Aksept 79, 81-95, 111, **118-125**, 131-132, 148-149, 167
Alpha 194
Amplitude 31-33
Anova 185, 212
Aroma 138, 151

B

Balansert rekkefølge 46, 75, **94-99**, 124-128, 149
Beskrivende test 42, 67, 81-83, **90-99**, 121-123, 138, 184-185, 191, 248-248
Bitter 17, **36-37**, 50, 54, 114, 130
Bølgelengde 28-30

C

Carry-over-effect 66, 91
Check All That Apply 126, 128, 149
Conjoint 131-133
Consensus 205, 212
Correlation plot 206, 207, 211
Critical value 194

D

Datamatrikse 202, 203
Design **42-45**, 131, 139-140, 177, 200
Differansetest 79
Dipolar 82
Direkte observasjon 120
Diskret 85, 86
Dummy-prøve 125
Dunnett's test 195
Duo-trio-test 79, 86, **92-97**, 178, 180, 183, 241-242, 250

E

Ensidig 94, **181**, 182
En-til-en-intervju 120
Error 176, 187, 189, 191, 194, 199
Ettersmak 41, 66, 107
Excel 179, 181, 205
Extrinsic 39

F

Fargeblind 29, 49
Fargespekter 30
Fisher 177, 188
Flash profiling 126-128
Flavour 138
Fokusgruppe 105, **118-121**, 145, 146
Forbruker **kapittel 5**, 149, 159, 162, 163
Forbrukertest 47, 60-75, 78, 85, 131-132, 142-146, 154

Forskjellstest **49-55**, 78-80, 108, 177-178
Forsøksdesign 42, 45, 75, 131
Forventningseffekt 72, 75
Free-choice 126
Frekvens 71, 72, 132
Friedman's test **191-193**, 196, 253
Frihetsgrader 187-188, 192, 196, 252-253

G

Gjenkjenneesterskel 50-54, 58, 61
Gjentak 46, **104-105**, 175, 188, 196-197, 199-200, 202-203, 205-206, 212, 243, 260
Glorieeffekt 75
Grunnsmak 34-37, 40, **49-53**, 109, 150
Gruppedynamikk 120

H

HACCP 161
Haloeffekt 75
Hedonisk 79, 123-125
Hierarkisk effekt 198-200
Hjemmetest 144
Hjerne 11, **26**
Hormoner 18
Hypotesetest 88, 123, **176**
Hørsel 25, 26, **31**





I

Ikke-smaker 35, 57
Indeks 205
Innredning 68, **70**
Intensitetsskala 74, **82-84**
Intersubjektiv 12
Intervallskala 81
Intervju 118-120
Intrinsic 139

J

Just-about-right 85

K

Kjemoreseptorer 34, 36
Klassifikasjon 127
Kode **67**, 88
Koeffisient 184, 205, 257-258
Konsistens 34, 36, 38, 80-83, 125, 137, 142-144
Kontekst 45, **131**
Kontinuerlig 85, 124
Kontrastfeil 73
Kryss Alt Som passER 146
Krysset effekt 198-200
Kvalitative tester 118
Kvalitetskontroll 12, 109, **154**, 170, 205, 218, 246-247
Kvantitative tester 78, 120, 125

L

Laboratorium 59, **69**, 132, 159, 168
Ladningsplot 203-204
Likert-skala 81
Line plots 206
Lukt 11, 16, **38-41**, 48-51, 55, 138, 159, 161
Lukteblindhet 41
Luktesansen 17, 27, **34**, 36, 39, 41
Lyd 16, 23, **31**
Lysbølger 32

M

Markedsanalyse 128, **118**
Matopplevelse 16, **17**, 22, **131**, 150
Mean 190-199
Memory standard 154-156
Metode 12, 46, **78**, **117**, 144, 149, 161
MFA 129
Moderator 119
Motivasjon 62
Multiple sammenligninger 193
Multivariat 121
MSE 191

N

Napping 128, 145
Neofobi 151
Nerver 18, **26**
Nese 34, 38, **40**
Nettesting 133
Nominalskala 80
Non-food 56, **216**
Nullhypotese 88, 90, **kapittel 11**
Nøstet effekt 198

O

Objektiv 11, 13, 21 27, **78**, 144, 159
Observasjon 80,120
Odour 34-36
Opplevelse 11, 17, 128, 149-151
Ordinalskala 80
Orthonasal 40, 138
Outlier 203

P

Panel 4, 5, 13, 14, 17, 27, **47**, 53, 57, 61, 76, 79, 102, 144
PanelCheck **61**, 191, **205-213**
Panellerleder 51, **59-60**, 73, 100
Papiller **35**, 36
Parametriske 8

Partest 79, 87, 90, 92- **93**, 122, 146, **181**
PCA **201-205**, 212
Persepsjon **16**, 118
Poengbedømmelse 59, **81**, 109
Preferanse 13, 14, 73-74, 79, 81, 118, **121**, 132, 146, 151
Preference mapping 125, 145-146
Produktutvikling 11, 13, 118, **137**
Profile plot 208
Projective mapping 147
PROP 37, 50, 57
Prøvepreparering 42, **62-63**, 100
PTC 37

R

Random 75, **197**
Rangeringstest 79, 87, **97**, 123
RaPID 220
Ratioskala 80-81, 176
Referanse 17, 108-109
Registreringssystemer 17, 19, **22**
Rekkfølgeeffekt 46, **76**
Repertory grid 121
Repeterbarhet 46, 51, 171
Retronasal **40-41**
Respondent 122, 126, 128
«Rånelukt» 41, 50

S

Salt 20, **36-37**, **50**, 52, 54
Samspill 197
Sansecelle 16, **19-20**, **24-25-**26, 36
Sansing 16, 19, **24**
Sannsynlighet 89, 93, 96, 123, **kapittel 11**
Sensitivitetsterskel **50-53**
Sensorikk 11, **12**, 13, 185
Sensorisk StudieGruppe 4, **224**





Sentrumseffekt 76
Sertifisering 156, **166**
Serveringsrekkefølge 45,
75-76, 87, 91-93, 122
Serveringsbeholdere 64
SLP **62**, 171
Signifikansnivå 176
Signifikant 180, 194-195, 199
Skala 61, 67, **80**, 100, 111-
113, 124, 156, 206, 212
Skåringsplot **203-204**
Smaksblind **37**, 49-50, 55, **56**
Smaksløk **35-36**, 57-58, 66
Smakssansen **34**, 36
Smell **34**, 36
Smiley 125
Sneglehus **31-32-33**
Sorting 126-**127**, 149
Spesifikasjon 111, 141, 143,
146, 154, 157, 166
Spiderplot 212-**213**
Spytt 36, **38**, **66**
Spørreskjema 42, 124, 130
Spørreundersøkelse 121
SSG 4, **224**
Standard 63, 94, 109, 155,
157, 163, 166, 167, 170, 185,
189, **194-197**
Standardavvik 190, 212
Statistikkprogram 179, 186,
190, 192-195, 201, 212
Staver **24**, 28
Stemmebånd 33
Straksprøve 155
Subjektiv 12-13, 79
Supersmaker 56-**57**
Surt **36**, 50, **52**, 54
Syn 4, 26, **28-29**, 31, 40, 56
Synergieffekt 72
Syre 37
Systemeffekt 75
Søtt 18, **36**, 50, **52**

T

Taktil 48, 217, 221, 222
Tallskala 81-82, 109
Tapper 24, 28, 29
Taste 34, 51
Terskelverdi 20, 25, 33, 41,42,
50-52, 61, 71, 112
Testrom 144
Tilvenning 20, 27, 28
To-av-fem 95, 184
Tosidig 94, **181-183**
Triade 121
Triangeltest 79, **86-96**, 138,
146, 147,176, 181
Trommehinne 31, 33
Tukey's test 194, 196
Tunge 20, 24, 27, **35-36**, 57

U

Umami 36, 37, 50
Unipolar 82-84
Univariat 201, 205
Unscrambler 203

V

Validitet 120, 150
Variansanalyse 176, **185-188**,
193, 196, 200, 209-212
Vokabular 119, 149

Ø

Øret 31-33
Øyet 19, 24, **28-29**, 34

