



NORGES LANDBRUKSHØGSKOLE

MEIERIINSTITUTTET

---

Forelesninger

om

SENSORISK KVALITETSKONTROLL

av

K. Steinsholt

ÅS-NLH, 1974

## Side.

1. Innledning .....	1
2. LUKT OG SMAK .....	2
2.1. Smakssansen .....	2
2.2. Luktesansen .....	11
2.2.1. Lukteepitelet .....	11
2.2.2. Luktenerven .....	12
2.2.3. Luktehjernen .....	12
2.2.4. Primær mekanisme .....	18
2.2.5. Studier av prosesser i luktecellene .....	22
3. KVANTITATIVE MÅLEMETODER FOR LUKT OG SMAK. ....	27
4. SYNSINNTRYKK. ....	28
5. BRUK AV FØLELSEN. ....	30
6. HØRSELEN.....	36
7. UTSTYR VED BEDØMMELSE. ....	36
8. SAMMENSETNING AV TESTPANEL ....	37
9. UTVELGELSE OG TRENING AV DOMMERNE .....	37
10. GENERELT OM BEDØMMELSER. ....	38
11. EN DEL VANLIGE BEDØMMELSESSOPPLEGG. ....	39
12. ENKELT-PRØVER. ....	40
13. PAR-PRØVER, DUOTRIO PRØVER ETC. ....	40
14. TRIANGEL-TEST .....	44
15. SEKVENSIELL ANALYSE FOR BESTEMMELSE AV ANTALL PRØVER	47
16. RANGORDNINGS-PRØVER .....	54
16.1. Direkte rangeringstest .....	54
16.2. Indirekte rangeringsserier. ....	63
17. SKILNADS TEST. ....	66
18. GRADING ELLER SCORINGSTEST. ....	76
19. PROFILERING .....	80

## 1 INNLEDNING.

Det er to vesensforskjellige egenskaper ved et næringsmiddel som bestemmer dets verdi. For det første er det middelets energiinnhold sammen med innhold av stoff som ikke kan syntetiseres av den menneskelige organisme. Vi kaller dette for næringsverdien av næringsmiddelet, og til å bestemme dette har vi tallrike kjemiske, fysiske og fysikalske metoder.

Den andre gruppen av egenskaper kan vi kalle nydelsesverdien av næringsmiddelet. Dette er en egenskapsgruppe som er bestemt av mange faktorer og er også situasjonsbetinget. Det legges lite vekt på denne gruppen hvis energiinntaket ligger nær minimum, og betydningen stiger jo lettere det er å tilfredsstille energibehovet dvs. jo høyere levestandarden er. Selv om det har vært arbeidet meget med å finne analytiske- "objektive"- metoder for å kvantifisere de enkelte egenskapene som samlet utgjør nydelsesverdien, nemlig produktenes farge, konsistens og lukt og smak, er det ennå ikke noen metode som kan erstatte den organoleptiske eller sensoriske bedømmelsen.

Begrepet "organoleptisk" stammer fra de to greske ordene "organon" og "lambainein" som betyr å gripe eller å få tak i. Ordet har derfor betydningen: "Det som organene eller sansene kan få tak i."

Selv om bedømmelser av f.eks. vin har pågått i mange århundre, er det først ved utviklingen av statistiske metoder til å behandle bedømmelsesresultatene at den sensoriske kvalitetsbedømmelsen har blitt en nødvendighet for produsenter av næringsmidler for:

1. å velge de riktige råvarene for bestemte produksjoner,
2. å etablere kvalitetsstandarder for eksisterende produkter,
3. å sammenlikne egne varers kvalitet med konkurrentenes, og
4. å utvikle nye produkter eller forbedre eksisterende produkter.

Valg av metodikk er avhengig av hva vi vil oppnå med bedømmelsen og også av hvor mange prøver vi har til rådighet, og hvor mange penger vi mener det er tilrådelig å legge ned i bedømmelsen.

Det er som regel fire egenskaper som en legger vekt på ved sensorisk bedømmelse, nemlig: Produktets utseende, dets konsistens, lukt og smak. Mer spesifiserte egenskaper innenfor de fire hovedegenskapene kan også ha interesse for en del produkter f.eks. seighet, sandethet, nedsmelting o.s.v.

Vi skal innledningsvis se litt på de 4 hovedgruppene, litt om virkemåten av smaks- og luktesansen, og kort om enkelte "objektive" metoder til å bestemme kvalitetsegenskaper, før vi går over til selve prøvemetodene som er hovedtemaet i dette kurset.

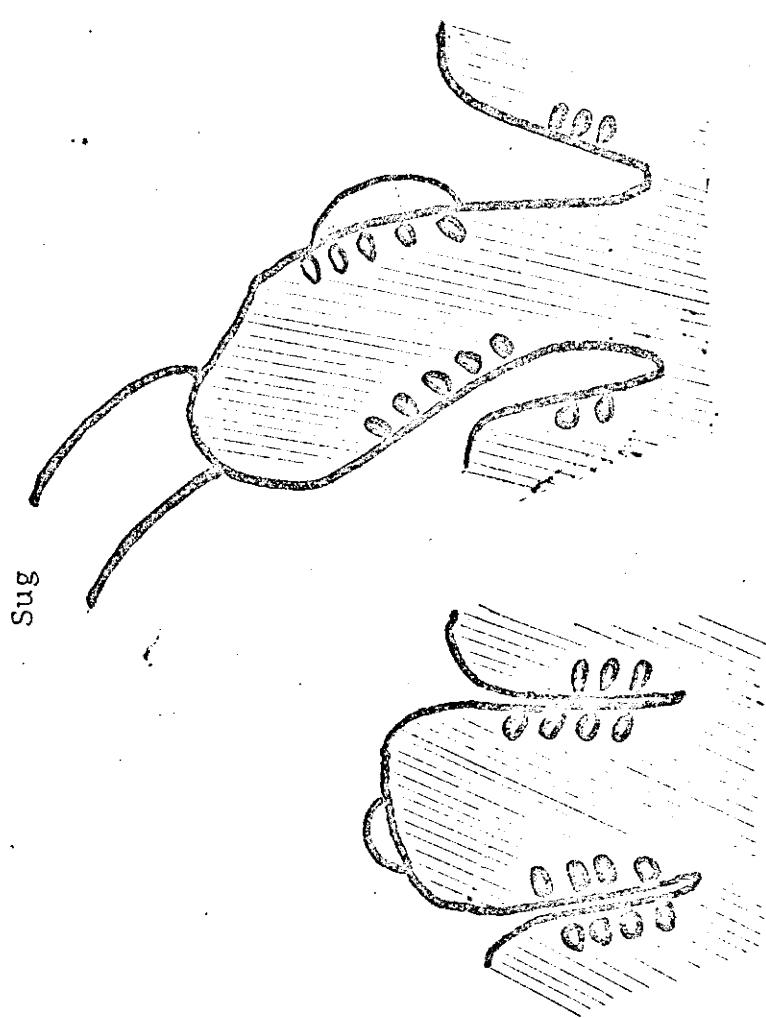
## 2. LUKT OG SMAK.

### 2.1. Smakssansen.

Smak er det følelsesinntrykket som en får ved stimulering av reseptorcellene som ligger i smaksløkene på tunga. Det er enighet om at det bare finnes fire forskjellige smaksretninger: Søtt, salt, surt og bittert. Typiske eksempler er: sukrose, koksalt, sitronsyre og kinin. Smaksstoffene må være i oppløsning for å nå fram til smakscellene for så å gi en elektrisk utløsning som sendes gjennom nervefibrene for å bli registrert i hjernen.

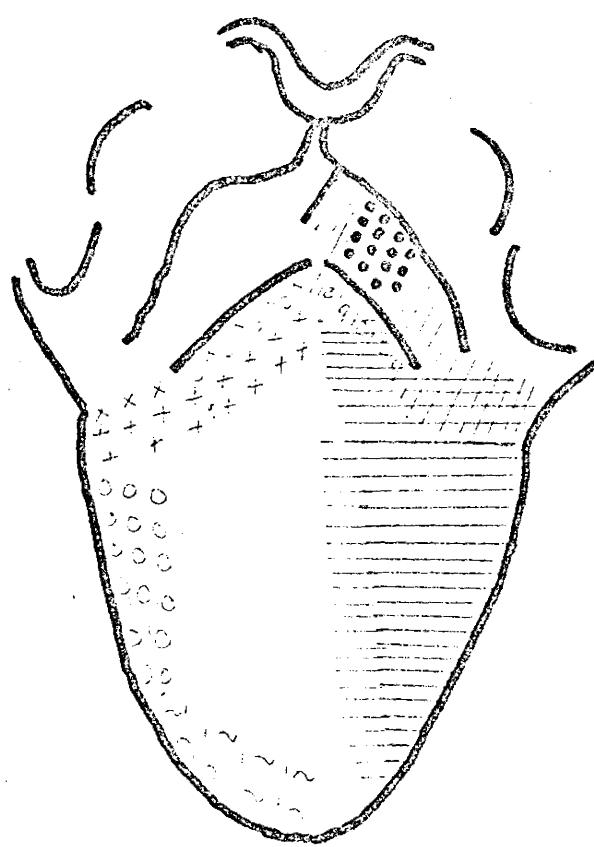
Den fysiske kjemien bak stimulering og utløsningsmekanismen inkluderer trolig en svak adsorbsjon av den kjemiske stimulatoren eller smaksstoffet til smakscellenes motaker, noe som resulterer i et jonebytte med elektrisk polarisering som gir nerveimpulsen. Disse elektriske strømningene er registrert ved å operere små ledere inn i smakscellene. Videre er det observert (von BEKESY (2.3)) at elektriske impulser til smaksløkene kan gi smaksfølelser (Fig. 1.)

En smaksløk inneholder mange mottakerceller som igjen har mange reseptorer. En celle kan observere flere smakstyper, men det er likevel konstatert bl.a. ved elektriske impulser at cellene er gruppert på tunga etter sin evne til å reagere sterkest på en enkelt smaksretning (fig. 2.). Søthet smakes best av tungespissen, saltsmaken noe lenger <sup>bak</sup> på tungeranden. På denne, men



Figur 1. Undersøkelse av smakreceptorenes plassering. ( Etter BEKESY 3.)

+ = bitter  
0 = sur  
/ = salt  
~ = søtt



Figur 2.  
Smaksfølelse i forskjellige områder på tunga.

lengre bak, reagerer cellene sterkest på sur smak, mens en bitter smak føles best helt bak på tunga.

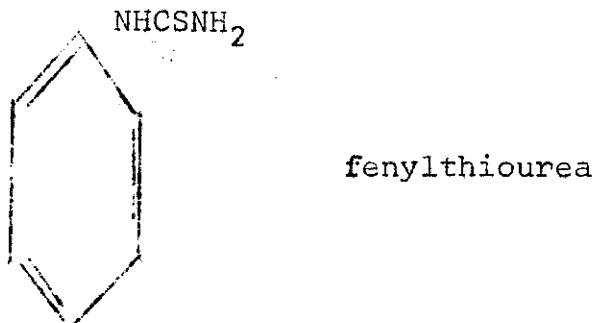
Det er forøvrig to teorier for reseptorenes selektive virkemåte, som begge kan forklare den sonemessige følsomheten for de fire smaksretningene på tunga:

1. Smaksakseptorene reagerer spesifikt på stofftyper av omtrent samme konfigurasjon. Hver reseptør gir en type impuls som svarer til en av de fire smaksretningene.
2. Søtt blir oppfattet av de mest følsomme reseptorene, salt og bittert oppfattes av mindre følsomme reseptorer og surt først når alle reseptorene er stimulert. (Spektrumteorien).

På samme måten som de øvrige sansene, reagerer smaken ikke bare kvalitativt, men også kvantitativt på stimuli. Evnen til å smake små konsentrasjoner av de fire smaksretningene er forskjellige fra individ til individ og det er rapportert om forskjeller p.g.a. kjønn, alder, røykevaner etc. Den laveste grensen som et individ kan smake, kalles for treskelverdien og det regnes med at en gjennomsnittlig treskelverdi ligger på omtrent:

Søthet:	Sukrose	0,4 %
Salt:	Natriumklorid	0,18 %
Sur:	Vinsyre	0,017 %
Bittert:	Koffein	0,0038 %

Det viser seg også at evnen til å registrere smak kan være genetisk bestemt. Stoffet fenylthiocarbamat (PTC) smaker vondt for 70 % av befolkningen mens 30 % vanskelig kan observere dette. Smaksevnen skylles et dominant gen omtrent på samme måten som brunfarge på øynene. PTC har vært mye anvendt i studier av bl.a. raser og spesielle isolerte populasjoner.

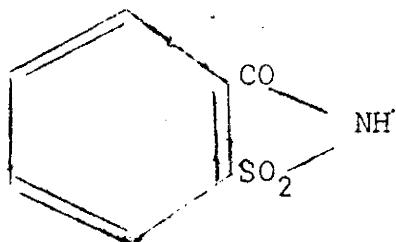


Det er også observert ved elektriske impulser at hvis en f.eks. skyller munnen med sukkervann, så vil en få en sterkt øket terskelverdi for de andre tre hovedretningene av smak.

#### 2.1.1. Kjemisk basis for de fire forskjellige smaksretningene.

1. Med salt forstår vi vanligvis NaCl, men også andre klorider kan gi typisk saltsmak. Det er derfor antatt at det er Cl<sup>-</sup> heller enn katjonene i et klorid som er ansvarlig for salt-smaken. Nå er NaCl salt, NaBr bittert og salt, NaI salt og bittert mens Na<sub>2</sub>SO<sub>4</sub> har en smak som kombinerer de fire smakstypene salt-bitter-sur-søt. Na<sup>+</sup> har derfor også en viss betydning. Det ble omtalt allerede i 1922 av KIONKA og STRÄZ (15) at katjonene bestemte styrken av saltmaken mens anjonene bestemte smakskarakteren. Det er videre funnet at salter med lav molekylvekt smaker salt mens høymolekylære salter ofte har en bitter smak.
2. En sur stimulus er alltid en følge av H<sub>3</sub>O<sup>+</sup>-joner. Imidlertid har også koncentrasjonene av udiissosiert syre betydning for smakens styrke og muligens også karakteren av anjonet.
3. Det er langt værre å fastslå hvilken kjemisk konfigurasjon som gir søt smak enn hva tilfeldet var for de salte og sure smaksretningene. En mengde organiske stoff gir søtsmak, og de kan variere mye i den kjemiske oppbyggingen. Imidlertid er alle som regel ikke ioniserte.

Glyserol og sakkarose som begge smaker søtt, er polyalkoholer. Derimot har sakkarin, som er enda søtere, en helt annen oppbygging:



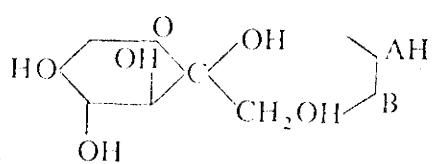
Benzenringen gir ofte en bitter smak, og en anelse av en slik smak finnes også i sakkarin.

Fra tungespissen på kveg og gris er det ekstrahert proteiner som danner komplekse forbindelser med sukker. Reaksjonen er ikke tidsavhengig og stabiliteten avhenger av sukkerets koncentration. DASTOLI og PRICE (7) har funnet at forandringen i den fri energien ved "reaksjonen" er større desto mer søtt sukkeret er. Forandringen er imidlertid så liten at det tyder på en virkelig kjemisk reaksjon.

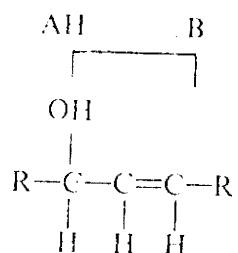
$\Delta F$  for fruktose-protein reaksjonen er bestemt til å ligge svært nær  $\Delta F$  for dannelsen av hydrogenbindinger ( $\pm 5$  kcal/mol).

Dette styrker teorien om at søt smak kommer av dannelsen av hydrogenbindinger mellom det søte stoffet og receptor-molekylet. SHALLENBERGER og ACREEL ( 24 ) mener at et søtt stoff har to komponenter A og B. A og B er elektronegative atomer som er adskilt av en distanse mellom 2,5 og 4 Å.  $H^+$  binder de to elektronnegative ionene med en kovalent binding. AH er da en protondonator mens B er protonakseptor.

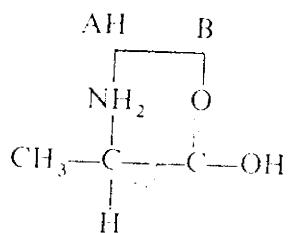
## Eksempler



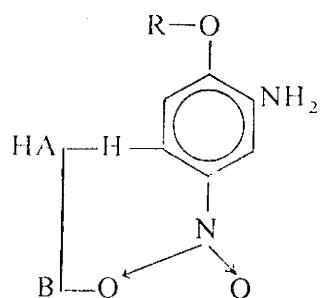
Fruktsukker



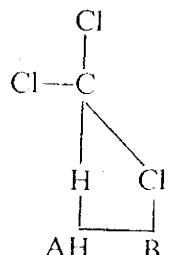
Umettet syre



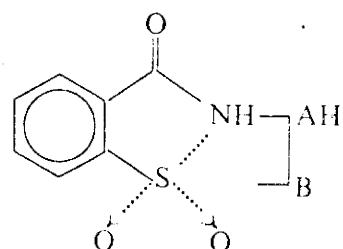
Alanin



2-amino-4-nitrobenzen

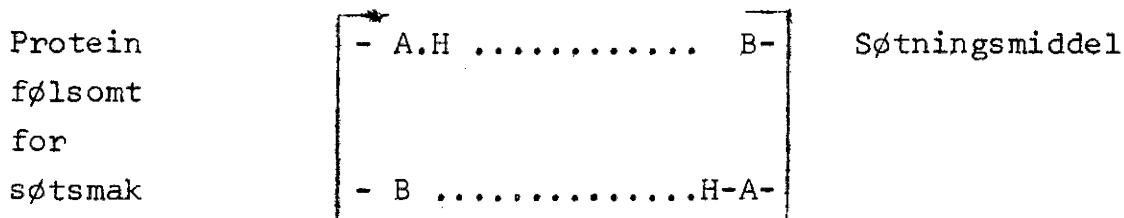


Kloroform



Sakkarin

Proteinet i reseptoren som er følsomt for söt-smak, har en tilsvarende oppbygning slik at:



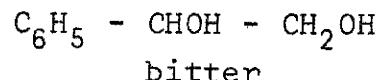
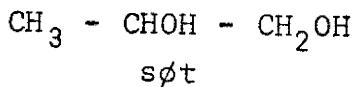
Det er funnet (7) at det söt-ømfintlige proteinet er homogent og har en molekylvekt på ca. 150000. Det har et stort antall kationaktive grupper, og er ganske varmeresistent.

4. Som nevnt i forbindelse med omtalen av salt-smak, har ofte salter med en høy molekylvekt en bitter smak. CsCl er f.eks. sterkt bittert. Det ser også ut til å være en liknende sammenheng mellom sott og bittert. I homologe serier er ofte de lavmolekulære stoffene sote mens de høymolekulære er bitre.

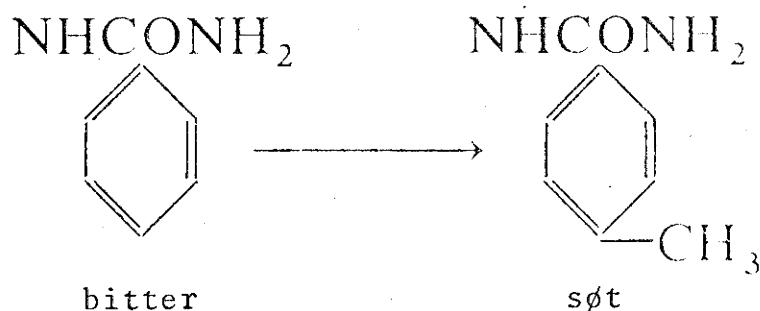
$\text{NO}_2$ -grupper gir ofte bitter smak. Symmetriske forbindelser har ofte også en bitter smak:



Fenylgrupper kan gi bitter smak:



Metylering kan gi motsatt virkning:



Det er satt fram en teori om at bittere stoffer har A og B komponenter som søte forbindelser. Her er imidlertid avstanden mellom gruppene 1,5 Å eller halvparten av hva den var mellom gruppene i stoff som smaker søtt.

På samme måten som for stoff som smaker søtt, er det funnet proteiner som reagerer med bitre smaksstoffer, og antagelig her også gjennom hydrogenbindinger.

## 2.2. Luktesansen.

Selv om en ved kombinasjon av de fire primære smaksretninger kan få et vist spektrum av smaken, er det klart at noen særlig kulinarisk opplevelse kan ikke smakssansen alene gi oss. Luktesansen, derimot, er langt mer mangfoldig og også langt mer følsom. F.eks. er følsomheten for etanol ca. 24000 ganger større for luktesansen enn for smakssansen. Merkaptaner f.eks. (Mossekjønnslukt) kan observeres i en mengde av 0,001 mg pr.  $m^3$  luft. Silkespinnerhannen kan lukte 30-40 molekyler pr.  $cm^3$  av kjønnsferomonet til hunndyret (10 trans- 12 -cis -heksadekadienol). Men så har også silkespinneren lukteorganene på antenner og kan antagelig bare kjenne denne lukta. Han lukter da også det annet kjønn på 3-4 km avstand.

Da luktesansen er særdeles viktig i all organoleptisk bedømmelse, skal jeg her ta med litt om oppbygginga av systemet og teorier for hvordan en mener det virken. Det meste av stoffet er hentet fra et foredrag av ELLEF VOLD ved NINF fra publikasjoner av DØVING (10), OTTOSON og SHEPHERD (21).

Luktesystemet består av lukteepitelet i neseslimhinna, luktnerven og lukthjernen.

### 2.2.1. Lukteepitelet.

Hos mennesket ligger lukteorganet regio olfactoria i taket på nesehulrommet og dekker omtrent  $2 \frac{1}{2} cm^2$  på begge sider. (Fig. 3.) Organet ligger i ei bakevje og kommer ikke i kontakt med hovedstrømmen av den lufta vi puster inn. Det er derfor viktig ved lukting at en puster raskt inn (sniffer).

Lukteepitelet er  $150-200 \mu$  tjukt og består av 3 lag celler: Lukteceller, støtteceller og basalceller. Det er omtrent  $2 \times 10^7$  lukteceller i hver nesehule.

Luktecellene er bygd av et hule- eller eggforma midtstykke med kjerne og en perifer og en sentral forsats (Fig. 4.)

Den perifere forsatsen har i enden noen fine hår som kalles luktehår eller cilier ( $9,2 \mu$  tykke og  $10-100 \mu$  lange). Det er sannsynlig at membranen rundt ciliene er den delen som reagerer på luktestoffene.

Den sentrale forsatsen er en nervetråd hvor impulser blir sendt videre til sekundære lukteceller.

Støttecellene inneholder ovale kjerner og et gult pigment. De ender i tynne hår. Ved roten er støttecellene flatttrykte og forsynt med nisjer som er utfylt av luktecellene. Håra eller ciliene til de to typene av celler danner et tett nettverk på overflata av r. olfactoria.

I overgangen mellom lukteespitelet og det underliggende vevet ligger stjerneformede basalceller.

I bindevevet ligger kjertler som produserer et slimlag som er meget viktig for en normal funksjon av lukteorganet.

R.olfactoria har et gult pigment som skiller den fra det vevet som ligger rundt. Ingen vet hvilken misjon, om noen-, dette fargestoffet har.

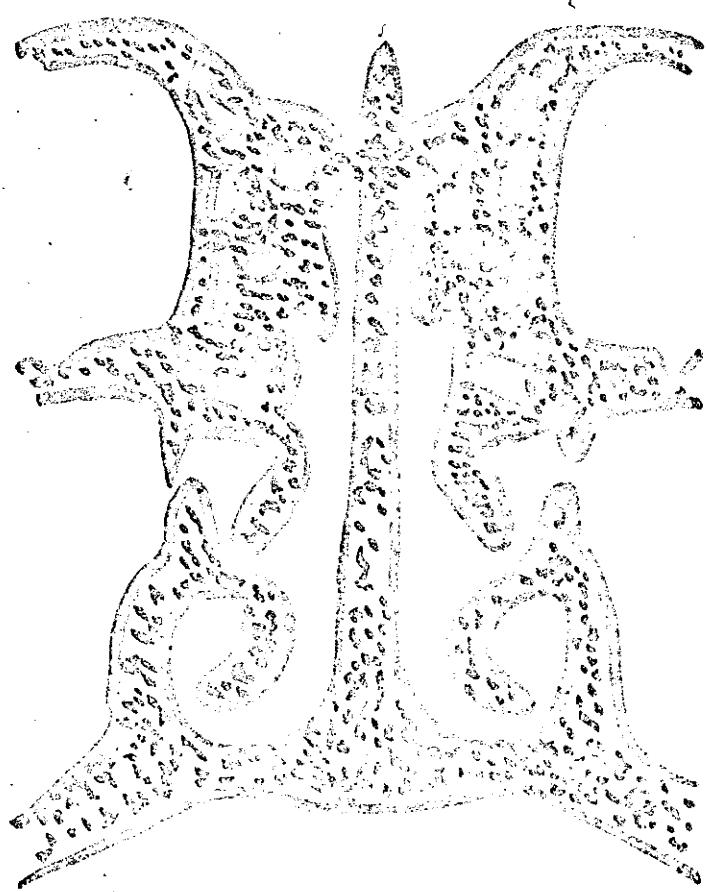
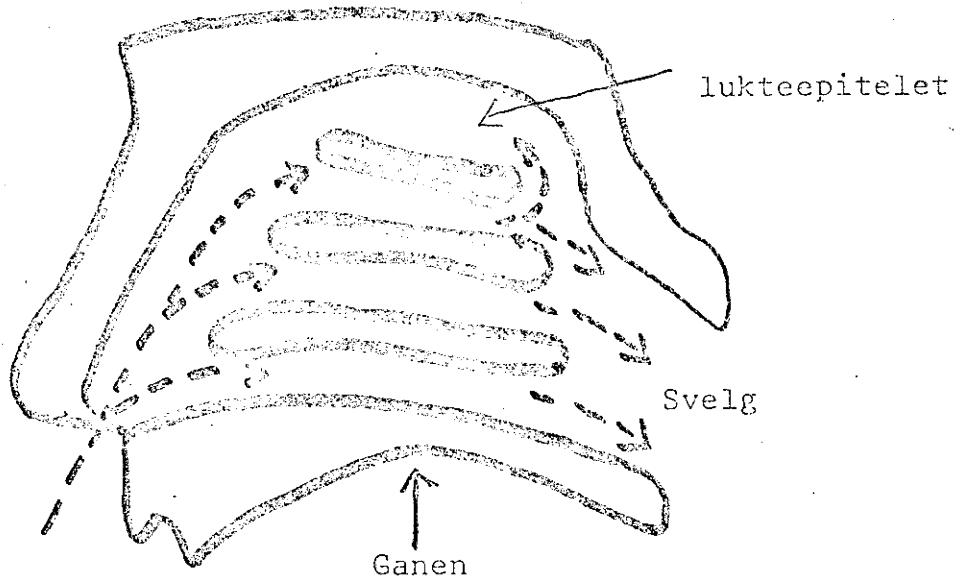
### 2.2.2. Luktnerven.

Nervetrådene fra grupper av lukteceller på flere tusen samler seg i bunter omgitt av gliaceller. Buntene (fila olfactoria) går gjennom silebeinet og til luktehjernen.

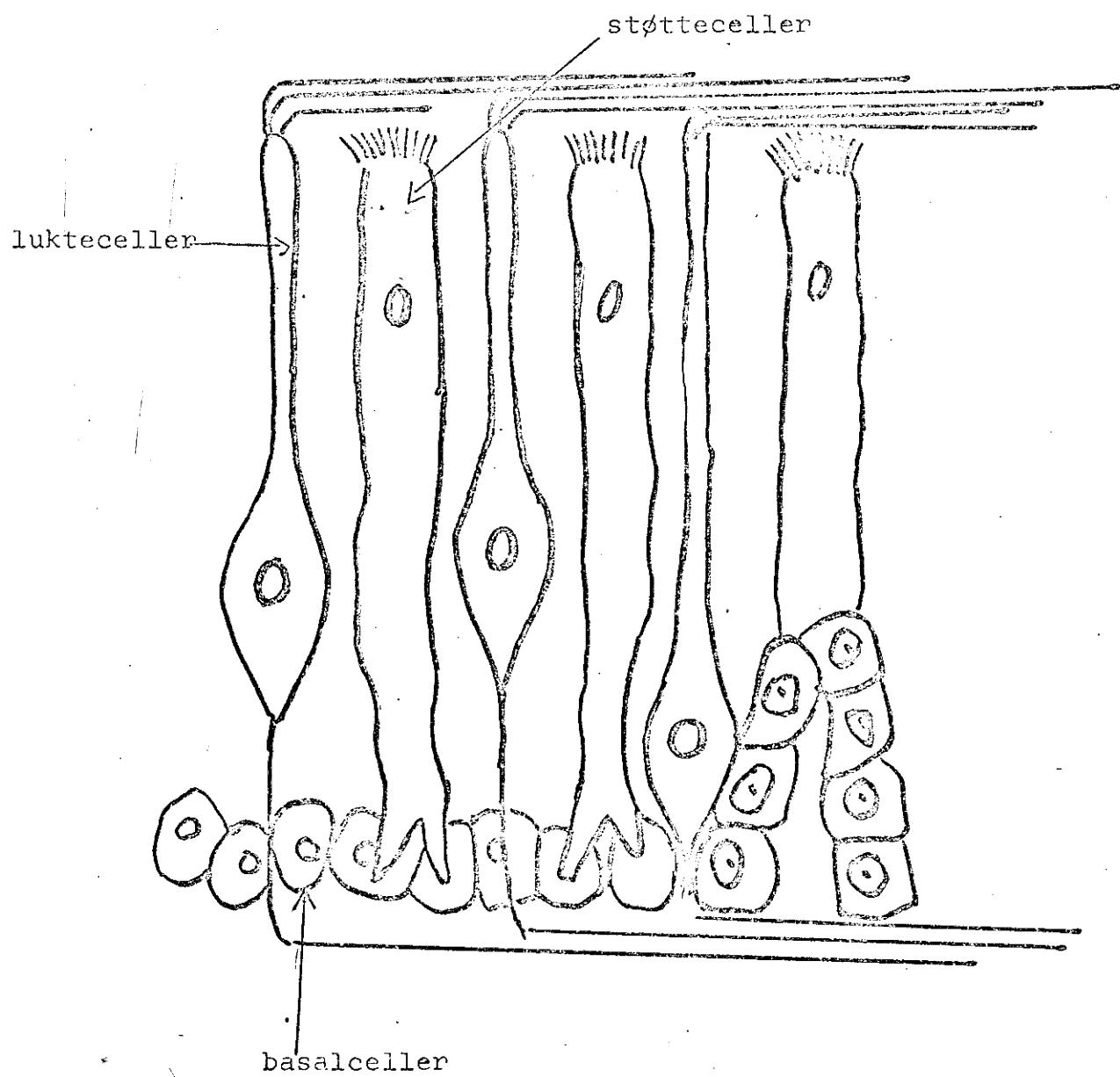
### 2.2.3. Luktehjernen.

Denne kan på sett og vis sies å være tredelt med en forluktehjerne (bulbus olfactorius) og et sambandsledd (tractus olfactory) mellom denne og de forskjellige luktesentra i hjernen (rhinencephalon). (Fig. 5.)

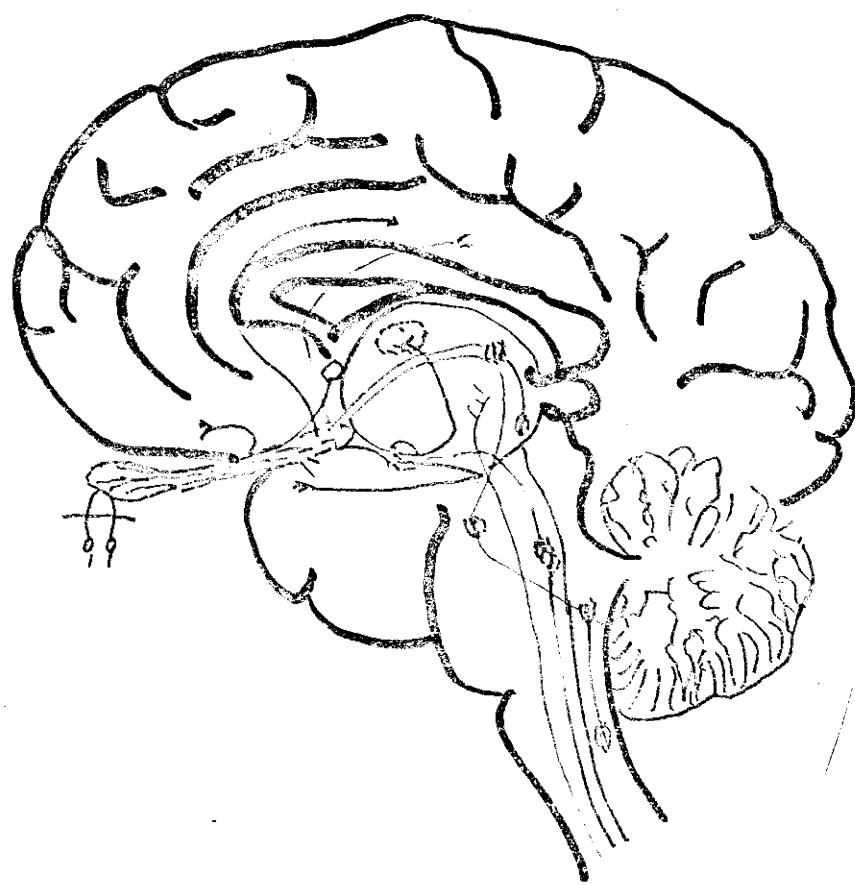
Figur 3. Lukteepitelets plassering.



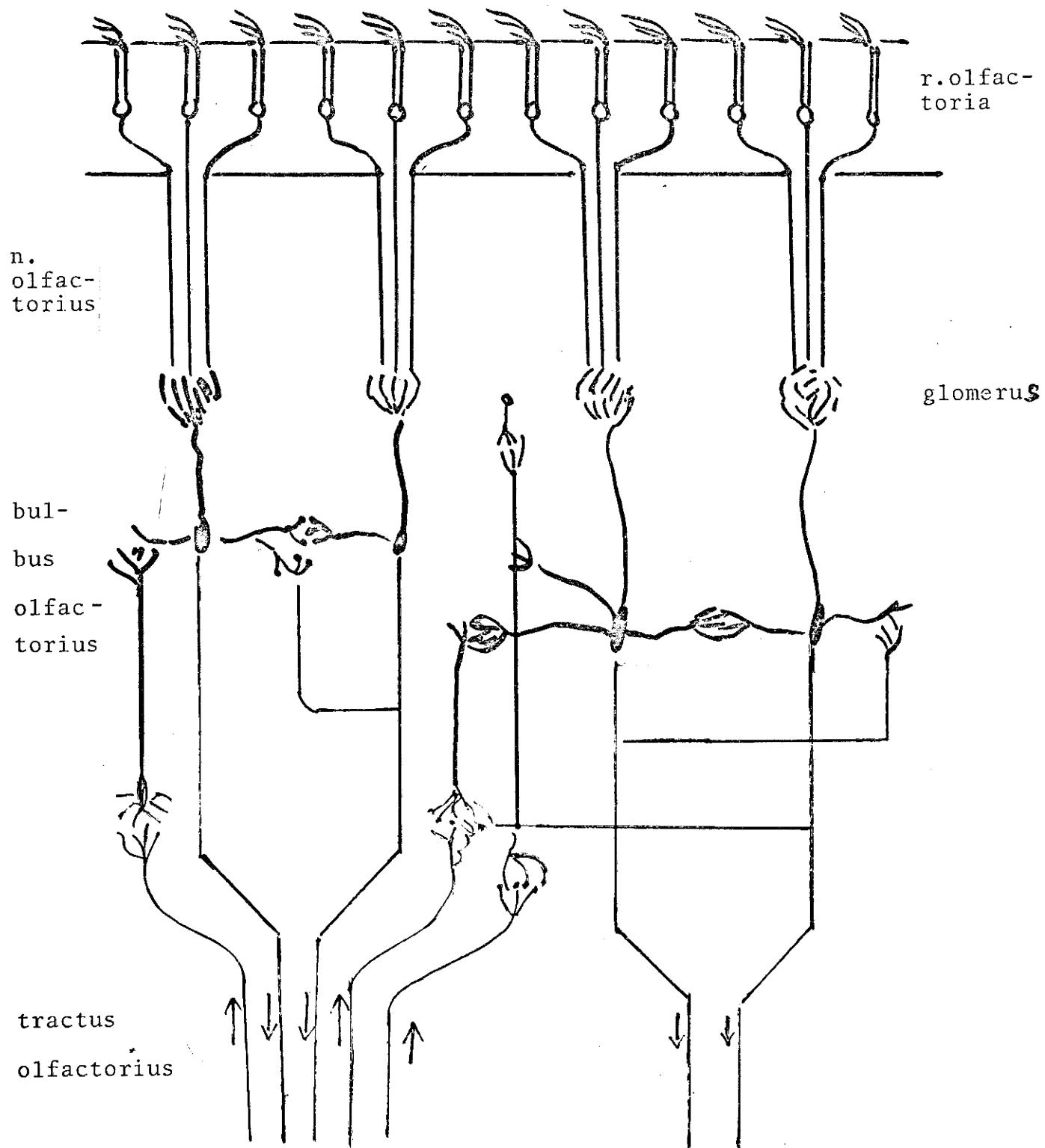
Figur 4. Celler i luktepitelet.



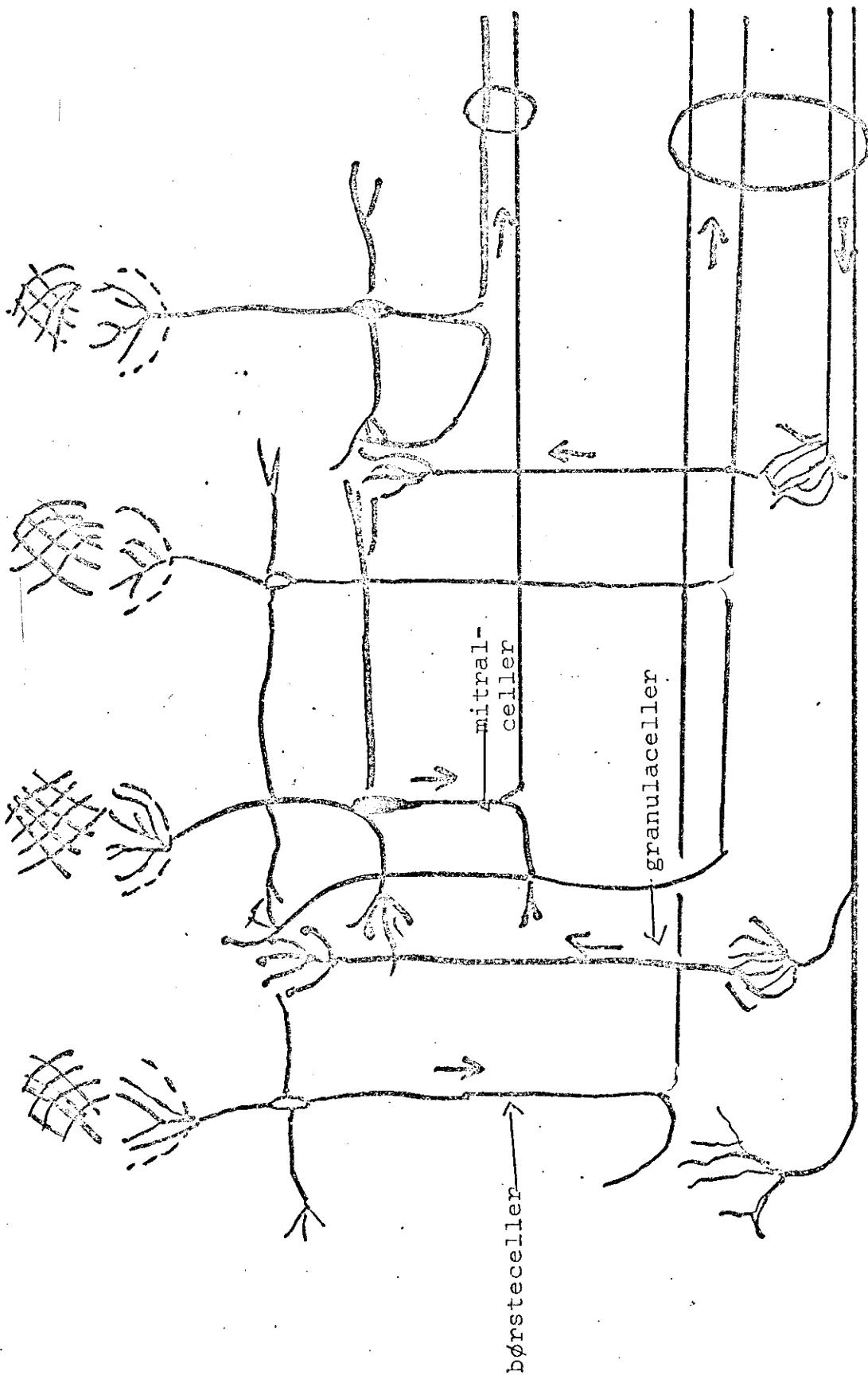
Figur 5. Luktenervenes plassering i hjernen.



Figur 6. Skisse over systemet i forlukthjernen.



Figur 7. Impulsretning i bulbus olfactorius.



Nervetrådene fra de primære luktecellene ender i små floker, glomeruli, i b.olfactorius (fig.6.). Tråden for hver enkelt av luktecellene står i en slags transformatorkontakt (synaptisk kontakt) med forgrenede utløpere fra de sekundære luktecellene. Etter DØVING (8) kan utløperen (primær dendrit) fra de sekundære luktecellene i en glomerus motta impulser fra omtrent 1000 primære lukteceller. De sekundære luktecellene har videre sekundære utløpere slik at flere av disse cellene står i synaptisk kontakt med hverandre. Det har vist seg at kontaktpunktene kan ha forskjellig polaritet som medfører at en impuls bare følger et bestemt rute. (Fig. 7.).

Nervetråder fra de sekundære luktecellene samler seg i det indre av b.olfactorius og går gjennom tractus olfactorius til luktehjernen. Her finnes lukteceller av tertiær karakter som kan gruppertes på grunnlag av sambandsmåten med nervene fra b. olfactory. Første gruppen står i synaptisk kontakt med tråder fra de sekundære luktecellene, den andre står i kontakt med nerver fra celler fra den første gruppen, og den tredje står bare i kontakt med celler fra den andre gruppen.

Ellers er det svært lite som er kjent med hensyn til hvordan de forskjellige luktesentra i luktehjernen virker,

#### 2.2.4. Primær mekanisme.

Med luktesystemets primære mekanisme forstår vi hvordan kjemiske stoff i den lufta vi puster inn kan virke på ciliene til de primære luktecellene slik at det blir utløst impulser der som går videre til de sekundære og tertiære lukteceller.

Den stereokjemiske teorien er blitt lansert av MONGRIEFF (14). Den går ut på at det er den stereokjemiske oppbygningen og størrelsen (diameter) i flere plan hos et molekyl som er årsaken til at det har en bestemt lukt. Luktehåra i de primære luktecellene har etter denne teorien, innhogg spm passer til molekyler av helt spesiell utforming. Bare når et slikt molekyl kommer i innhogget blir det utløst en impuls slik at hver primær luktecelle har hår som bare kan oppfange et slags molekyler.

Organoleptiske undersøkelser har indikert at det finnes 7 basislukter: Kamfer, moskus, blomster, peppermynte, eter, stram og rotten. AMOORE (1) har kommet frem til mål og utforming av innhogga ved å studere oppbygningen av en rekke luktestoffer. (Fig. 8). Det er også på forhånd postukert med spesille lukter av syntetiske molekyler utfra deres form, og spådommene har slått til.

Punkteringsteorien etter DAVIES (6) går ut på at molekyla til luktstoff vil bli selektivt adsorbert til den fettaktige cellemembranen til ciliene, og tvinge seg gjennom denne. På denne måten vil de direkte kunne skape en impuls, eller lette jonebytte på begge sider av membranen, og derved skape en impulse.

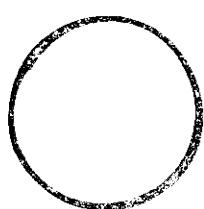
Vibrasjonsteorien hevder at flyktighet og fett- og vannoppløselighet bestemmer styrken til et luktestoff mens vibrasjonsfrekvensen i molekyla bestemmer kvaliteten eller luktsretningen. Bare svingninger vinkelrett på overflata av en reseptør i r. olfectoria vil gi reaksjoner. Molekyla i luktestoffet blir styrt på plass av spesielle funksjonelle grupper.

Bindingsteorien etter RANDENBROCK (18) går ut på at stoff med lukt har tendens til å danne hydrogenbindinger, Van der WAALSKE bindingstyper eller er utpregede elektrondonator-akseptorer. Slike stoff kan da være alifatiske grupper med dobbelte eller tredobbelte bindinger, benzenringer, halogen, alkohol- eller etergrupper, aldehyd, keton- eller estergrupper, nitro-, cyan-, sulfhydryl- eller tioetergrupper. Reseptorene i lukteepitelet må da være tilpasset de forskjellige luktestoffene både i ladning, red.oks. potensial og strukturell utforming. Teorien går videre ut på at det i reseptorene finnes strukturer med interne svingninger som blir moderert av de spesielle gruppene i luktestoffene. Derved oppstår det svingninger som er spesifikke for luktestoffet og disse gir da den afferente nerveimpulsen fra r. olfectoria.

En fjerde teori går ut på at reseptorene virker som sendere og mottakere slik som f.eks. en radareller flaggermusens styringsmekanismer. Reseptorene sender etter denne teorien ut bølger som blir modulerte spesifikt av de forskjellige luktestoffa. Disse modulerte bølgene blir så oppfattet av reseptorene som så sender impulsen videre.

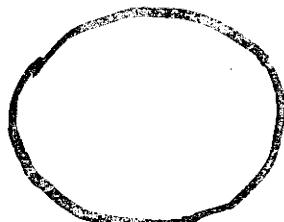
Figur 8. Sterokjemisk luktteori (Amoore(1)).

A



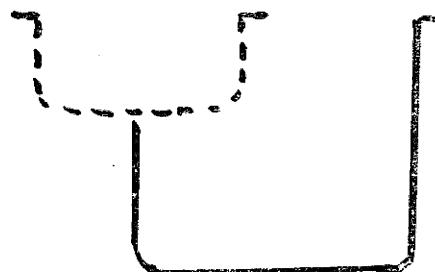
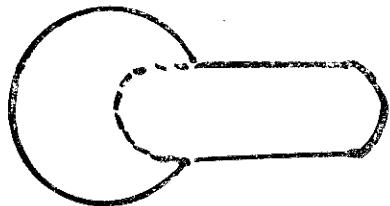
Kamfer

B



moskus

C



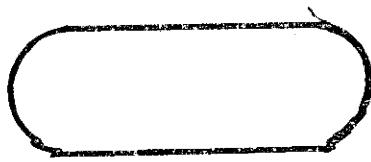
blomster

D



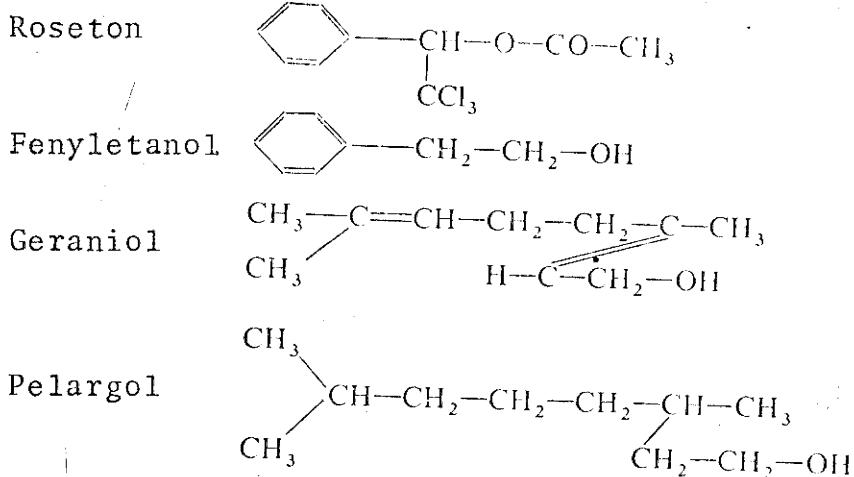
peppermint

E

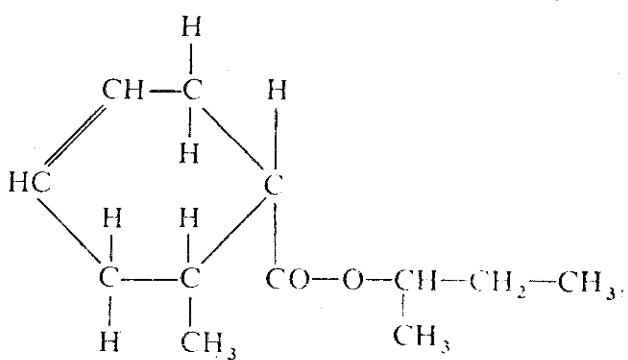


eterisk

I samsvar med teoriene for lukt, kan molekyler av forskjellig oppbygging ha samme luktretning. F.eks. lukter disse stoffene som roser.



Kjønnsferomet til en flue i Middelhavsområdet, *Ceratistis copitata*, har følgende cis-konfigurasjon:



mens trans-formen er fullstendig uvirksom.

Alle disse teoriene som hver for seg neppe kan forklare alle luktefenomen, indikerer muligens at det er mange egenskaper hos et luktestoff som kan indusere **impulser i reseptorene** og gi både kvalitative og kvantitative inntrykk av lukt. Det er regnet med at en normal luktesans kan skjelne mellom ca. 2000 lukter, mens en trenet vinsmaker kan differensiere mellom ca. 10000.

#### 2.2.5. Studier av prosesser i luktecellene.

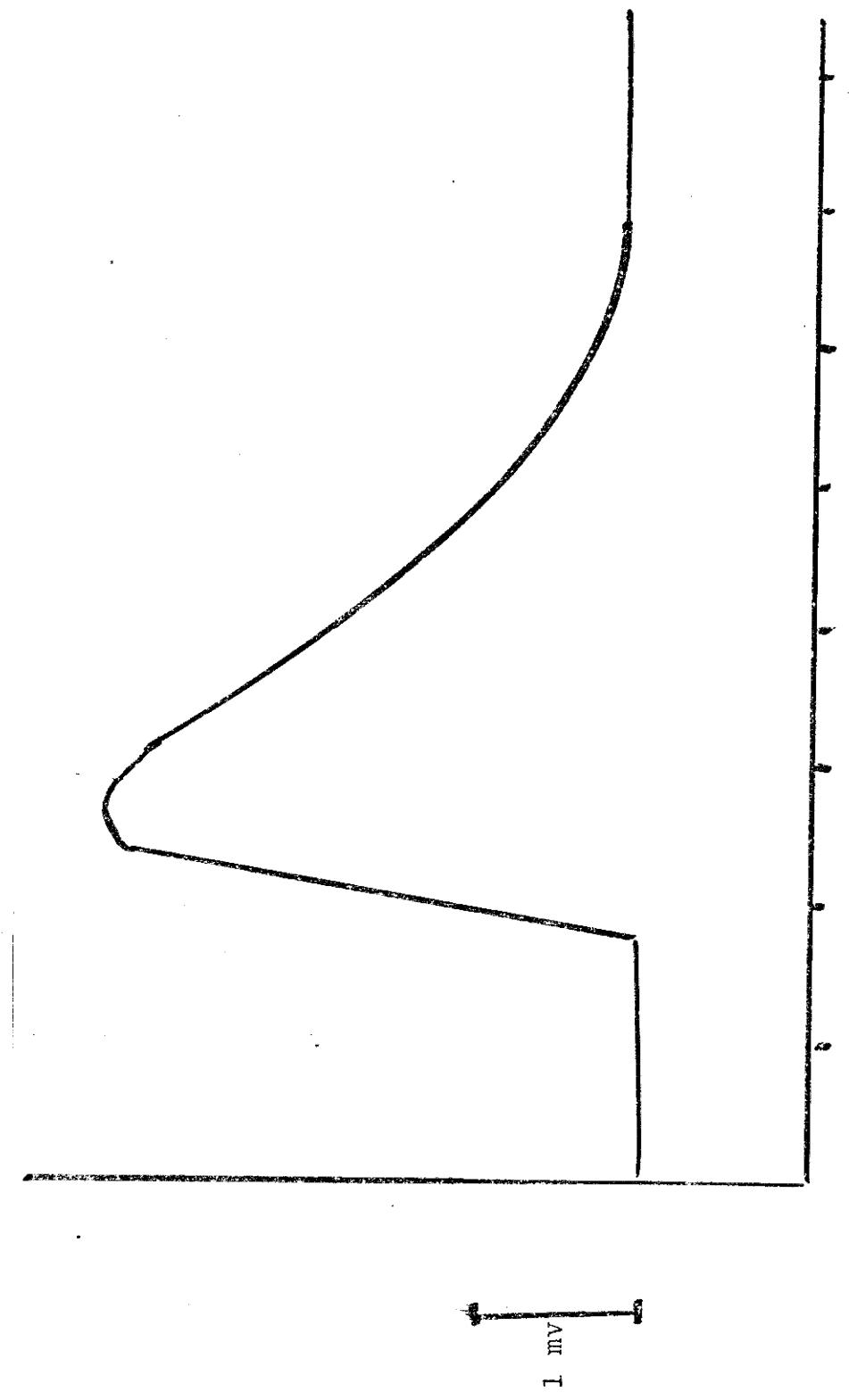
Studier av luktecellenes virkemåte krever som regel operative inngrep og har for det meste vært utført på dyr som frosk og kanin. Selv om resultatene ikke uten videre kan overføres til mennesker, kan en regne med at funksjonene prinsipielt er svært like hos de forskjellige hvirveldyr. Vi skal imidlertid være oppmerksom på at enkelte dyr har et lukteorgan (*organon vomeronasale*) i munnhola som kan tjene som organ for bl.a. sporsansen. Dette organet er rudimentært hos menneskene.

Studiene av luktecellenes virksomhet er gjort ved å plassere mikroelektroder i lukteorganene og studere utslag av luft med luktestoff i. Som regel brukes butanol.

OTTOSON (19,20) fant at det oppsto en negativ forandring i det elektriske potensialet i r.*olfactoria* 3/10 sekund etter at luktestoffet nådde cellene. Potensialet nådde et maksimum etter meget kort tid og ble borte etter få sekunder. En grafisk fremstilling med tida som absisse og potensialet som ordinat kalles et elektroolfatogram (EOG), (fig. 9.). Ved plastbelegging av r.*olfactoria* eller når ciliene på luktecellene fjernes, blir det ikke observert noe potensial. Kokain blokkerer vevet rundt resptorene men ikke selve reseptorene. Det er funnet at behandling med kokain har liten effekt på EOG.

Potensialet økte proporsjonalt med logaritmen til konsentrasjonen. Ved konstant konsentrasjon økte utslaget med volumet som blir "sniffet" inn. (Fig. 10.).

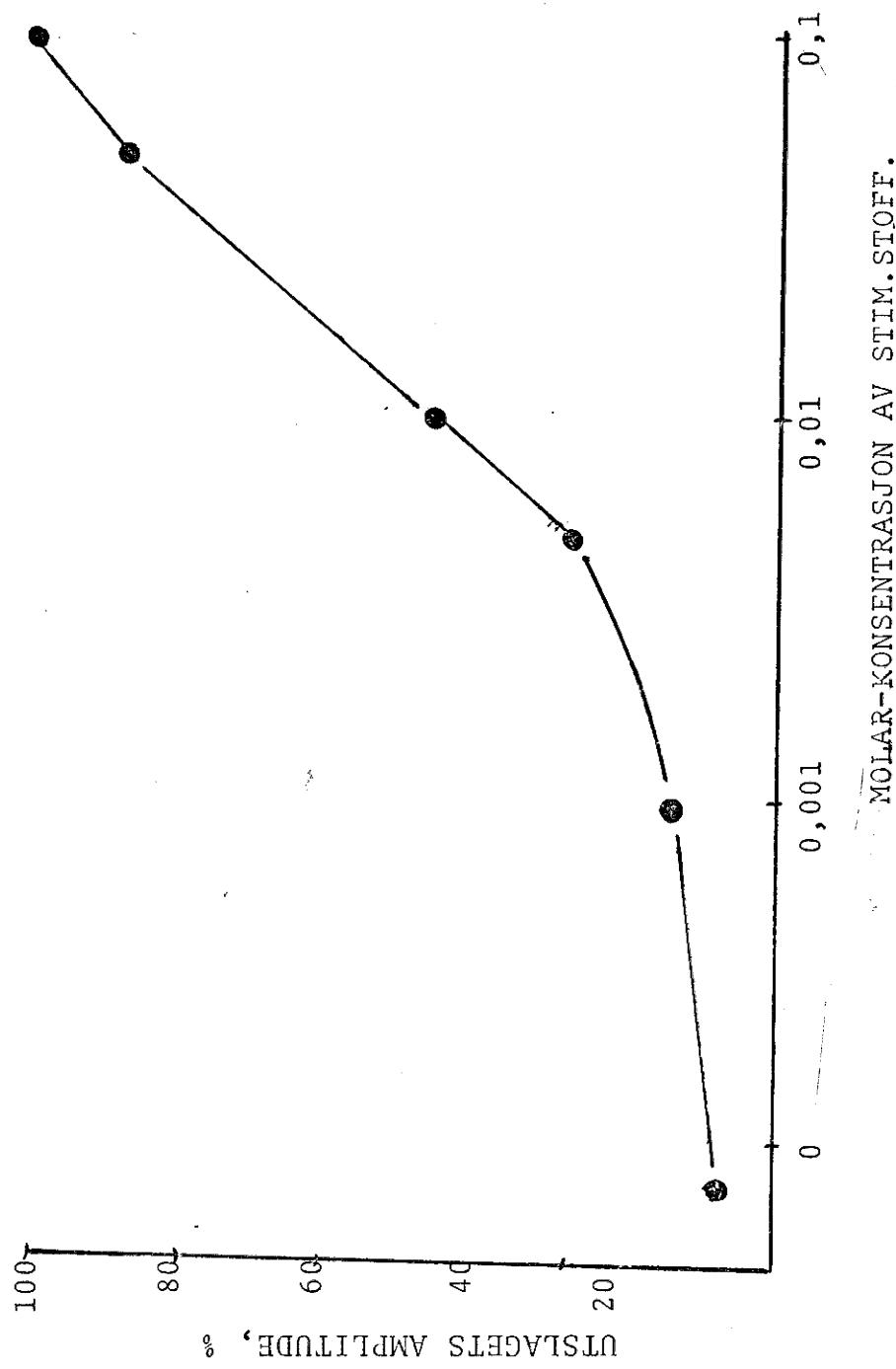
Av stor interesse for den sensoriske bedømmelsen er det forholdet at vedvarende stimulering førte til at utslaget ble borte. Utslaget ble også mindre og mindre når stimuleringen ble gjentatt med korte mellomrom. Virkningen var større desto større konsentrasjonen av luktestoffene var. (Fig. 11.). Ved studier av EOG er



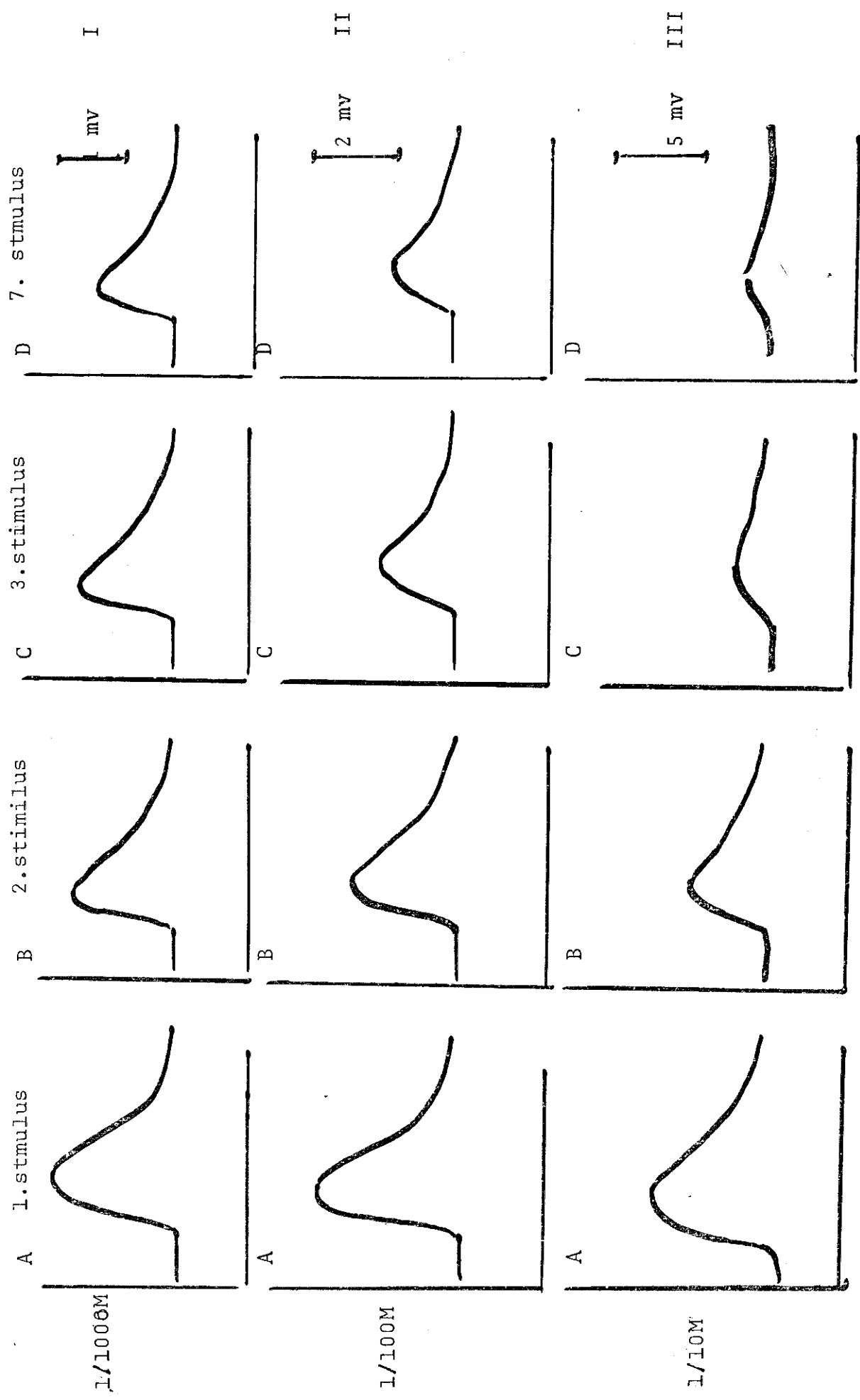
Elektroolfaktrogram.

Figur 9. Stimulering med butanholdig luft (15).

Figur 18. Virkningen av konsentrasjonen av stimulerende stoff på det relative utslagget i maksimum verdi for EOG (17).



Figur 11. Stimulering hvist 10 sekund med butanholdig luft (17).



det også funnet at den enkelte reseptør er mer følsom overfor et luktestoff enn overfor andre.

Det elektriske potensialet som oppstår i r.olefactoria når det kommer et luktestoff i kontakt med reseptorene, gir en nerveimpuls med en hastighet på ca. 0,2 m/sek. mot hjernen. Dette er sakte, sammenliknet med impulser i motoriske nerver som går ca. 1 m/sek.. Nervetrådene samler seg i bunter hvor de har en avstand på ca. 200 Å som faktisk er tilstrekkelig for å hindre overslag.

Studier av det sekundære systemet i b.olfactorius er utført prinsipielt på samme måten som for de primære luktecellene.

Resultatene kan sammenfattes slik:

Den nervøse aktiviteten fra de primære luktecellene sprer seg gjennom nervetrådene til de sekundære luktecellene via synaptiske kontakter. Mitralcellene blir da stimulert gjennom sin primære nervetråd (dendrit). Denne sender igjen en nerveimpuls til lukteceller av tertiær orden, og også gjennom sekundære nervetråder til granulacellene som sender hemmende impulser til utgangsmittalcellen hvor den vil hemme den opprinnelige impulsen. Denne hemningseffekten brer seg også til andre lukteceller enn de som har vært i virksomhet. F.eks. er det vist at en stimulering av luktecellene i det ene neseboret vil gi hemning i det andre.

Hva som skjer i de tertiære luktecellene når de mottar en impuls, er i dag så og si helt ukjent. Det er antagelig samband mellom luktesenteret i hjernen og hypofysen slik at enkelte lukter har innvirkning på andre kroppsorganer og også på adferdsmønster. F.eks. mener man at laksen lukter seg fram til den elva hvor den er født.

Det en idag mener å vite om oppbygging av virkemåte av smaks- og luktesansen, har følgende konsekvenser for den organoleptiske kvalitetsbedømmelsen:

Ved smaking må en ha næringsmiddelet så lenge i munnen at det berører hele tungen. Tørre produkter bør løses opp i vann før bedømmelsen for å sikre god kontakt med smaksløkene. En kan regne med at det første lukteinntrykket er det sterkeste og

derfor viktigste. Ved bedømmelse av prøver som lukter sterkt bør en ta en pause mellom hver prøve.

### 3. KVANTITATIVE MÅLEMETODER FOR LUKT OG SMAK.

Ved undersøkelse av smak eller lukt har vi idag ingen apparater som kan erstatte tunge og nese. Alle undersøkelser av smak- og luktestoff må derfor på et eller annet tidspunkt være knyttet til organoleptiske bedømmelser. Ved hjelp av f.eks. gasskromatografi kan en imidlertid kvantitativt bestemme et reint luktestoff når det først er konstateret at det er et luktestoff. Mer kompliserte lukts-inntrykk er det vanskelig å kvantifisere ved hjelp av analytiske metoder. Ved gasskromatografiske analyser av blandingsstoff, f.eks. vanilje, kan det være komponenter som utgjør en forsvinnende del av blandingen som er de egentlige luktestoffene. Dessuten kan enkelte stoff lukte sterkt selv i så små konsentrasjoner at de ikke kan oppdages ved analytiske metoder. Ikke desto mindre har gasskromatografi blitt et viktig instrument i utforskning av luktestoffer.

EISBERG og LEVY (1) har beskrevet en interessant og enkel måte til å bestemme terskelverdier for bestemte stoff på. (Fig. 12.). Luktestoffet blir plassert i ei flaske med vid åpning. Gjennom korken i flaska går det to glassrør. Et står i forbindelse med ei sprøyte for innsprøyting av luft, og et står i forbindelse med spesielle åpne "luktepropper". Metoden baseres på at luktestyrken er avhengig av den mengde luft som trekkes inn. Luftmengde "sprøytes" gjennom flaska inntil forsøkspersonen for det første kan si at det er et luktestoff til stede og for det andre bestemme stoffet. Personen skal holde pusten under forsøket. Den mengde luft som er tilstrekkelig for identifikasjon betegnes som luktekoeffisienten (olfactory coefficient). Denne varierer da sterkt fra stoff til stoff slik resultatene i tabell (3.1.) viser. Denne metoden kan også brukes til å få et mål for virkningen av visse sykdommer på lukteevnen. Metoden egner seg nok mindre bra for stoff som en fort "venner" seg til.

Tabell (3.1.). Luktekoeffisienter og kokepunkt (10).

Stoff	Koeffisient	Koke-punkt.
	cc	°C
Bensin	5,26	70(40-80)
Xylol	10,0	140(139-142)
Terpentinolje	10,7	158(155-162)
Smørsyre	11,43	162,3
Sitronsyre	12,75	172
Olje fra bitre mandler	13,08	180
Kreosol	14,5	212(205-220)
Kamfer	15,0	220
Anisolje	15,94	225
Kløverolje	17,22	255(250-260)

## 4. SYNSINNTRYKK.

Det første inntrykket en konsument får av en vare er ofte avgjørende. Varens utseende er derfor svært vesentlig og må også trekkes inn i en ordinær bedømmelse av et produkt. Emballasjens dekor, pakningens form og selve varens form og farge er ikke uvesentlig for varens mulighet på markedet. F.eks. pakkes ønkelte desertoster i trekantede pakninger av hensyn til det markedsførerne tror publikum liker, enda det teknisk sett er en meget uheldig pakkeform.

Det er også viktig at pakningen er ordentlig fylt og at det ikke er tilsølet utenpå.

Fargen på et produkt må også bedømmes, og det er vesentlig at farge og smak er "synkronisert".

Fargen til et produkt kan bestemmes objektivt enten ved å sammenlikne med en fargeskala (smør, brunost), ved sammenlikning med en serie farger f.eks. i et Lovibond Tintometer eller ved måling

av reflektivt lys fra overflaten av produktet.

Et standard system for beskrivelse av farger ble foreslått av Commission Internationale de l'Eclairage (CIE) i 1931. Systemet angir en farge med tre verdier, X, Y, Z - tristimulus verdier. Verdiene representerer de mengdene av rødt (X), grønt (Y) og blått (Z) som må blandes sammen for å få fargen. De kromatiske koordinatene defineres da som

$$x = \frac{X}{X+Y+Z}$$

$$y = \frac{Y}{X+Y+Z}$$

$$z = 1 - (x+y)$$

hvor en hver farge kan beskrives av x og y i et kromatisk diagram (figur 13).

CIE har valgt tre lyskilder som utgangspunkt:

- A. En wolfram-lampe fylt med gass og med en farge-temperatur på  $2854^{\circ}\text{K}$ .
- B. "Ettermiddags-lys" med farge-temperatur på  $4900^{\circ}\text{K}$ .
- C. "Gjennomsnittlig dagslys" med fargetemperatur på  $6700^{\circ}\text{K}$ .

B og C får en ved å bruke lyskilde A med spesifiserte filtre.

Fargen fra disse lyskildene har også en plass i systemet, og er for lyskilde A :

$$S_A = 0,44757(X) + 0,40745(Y) + 0,14498(Z)$$

hvor da (X) betyr "av farge X".

Energifordelingen for lyskildene gjennom det synlige spektret ( $380-780 \text{ m}\mu$ ) er satt opp i tabellform og betegnes  $P_A$ ,  $P_B$  og  $P_C$ .

Videre har CIE satt opp i tabellform distribusjonskoeffisienter for en såkalt lik energi stimulus gjennom spektret, dvs. X, Y og Z-verdier ved hver bølgelengde for en teoretisk hvit-farge hvor energien pr. enhet bølgelengde er lik gjennom hele det synlige spektret. Disse verdiene kalles  $\bar{x}_\lambda$ ,  $\bar{y}_\lambda$  og  $\bar{z}_\lambda$ ; og  $\sum \bar{x}_\lambda = \sum \bar{y}_\lambda = \sum \bar{z}_\lambda$ . Disse verdiene veid ved distribusjonskoeffisienten for energien av lyskilde A, er som eksempel satt opp i tabell (4.1). De relative verdiene av produktene er satt opp slik at  $\sum P_A \bar{x}_\lambda = 100.00$ .

Refleksjonen ( $P_\lambda$ ) fra den ukjente fargen måles nå for hver  $5 \text{ m}\mu$  mellom 380 og  $780 \text{ m}\mu$ .

Den ukjente fargens tristimulus verdier er da:

$$X = \sum P_\lambda P_A \bar{x}_\lambda$$

$$Y = \sum P_\lambda P_A \bar{y}_\lambda$$

$$Z = \sum P_\lambda P_A \bar{z}_\lambda$$

De kromatiske koordinatene ( $x$ ,  $y$  og  $z$ ) er definert som før som den ene tristimulus-verdien dividert med summen av de tre verdiene. Fargeretningen kan da beskrives av  $x$  og  $y$  i det kromatiske diagrammet, mens fargestyrken beskrives av  $Y$ -verdien.

Da  $\sum P_A \bar{y}_\lambda$  er valgt lik 100.00, får vi  $Y$ -verdiene i prosent.

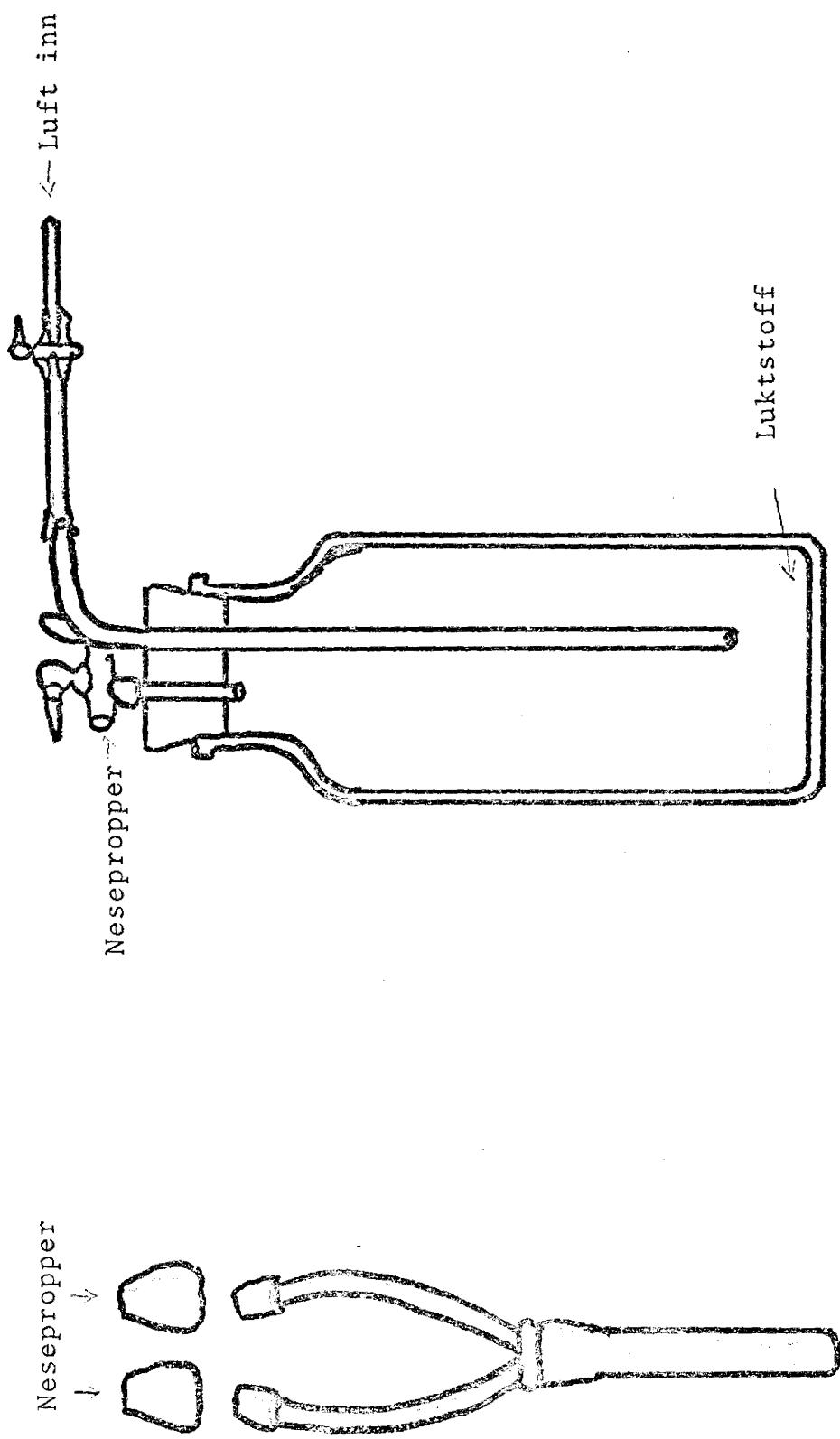
I næringsmiddelindustrien er Hunter Color and Color Difference Meter mye brukt. Fig. 14. viser hvilke måleenheter en bruker og disse a (rød  $\rightarrow$  grønn), b(gul  $\rightarrow$  blå) og L (produktets hvithet) kan overføres til CIE-systemet.

For de som måtte være interessert i en omfattende behandling av fargebestemmelser, anbefales boka til WRIGHT (26).

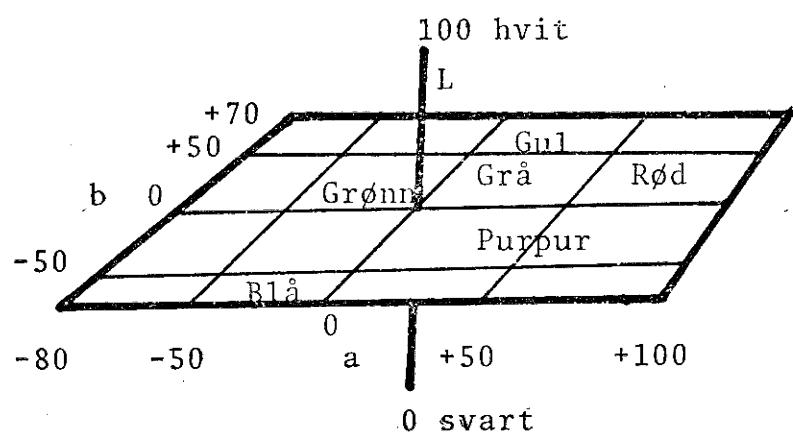
Et produkts "gloss" (blankhet) er enklere å bestemme i det totalrefleksjonen bestemmes ved bestemte infallsvinkler. Glossmålinger vil kunne ha interesse f.eks, i frukt, bær, syltetøy, etc.

## 5. BRUK AV FØLELSEN.

Vi bruker følelsessansen til vurdering av et produkts konsistens og teksturegenskaper. Fingrene kan vi bruke vesentlig for å kjenne om et produkt er fast eller mykt, men følelsen i munnen er nok viktigst. Her vil vi kjenne om produktet kliner seg til ganen, om det inneholder partikler, om det er fettaktig og om det er seigt og vanskelig å tygge osv.



Figur 12. Apparat for bestemmelse av luktekoeffisient. (22)



Figur 14. Fargeskala for Hunter Color and Color Difference Meber.

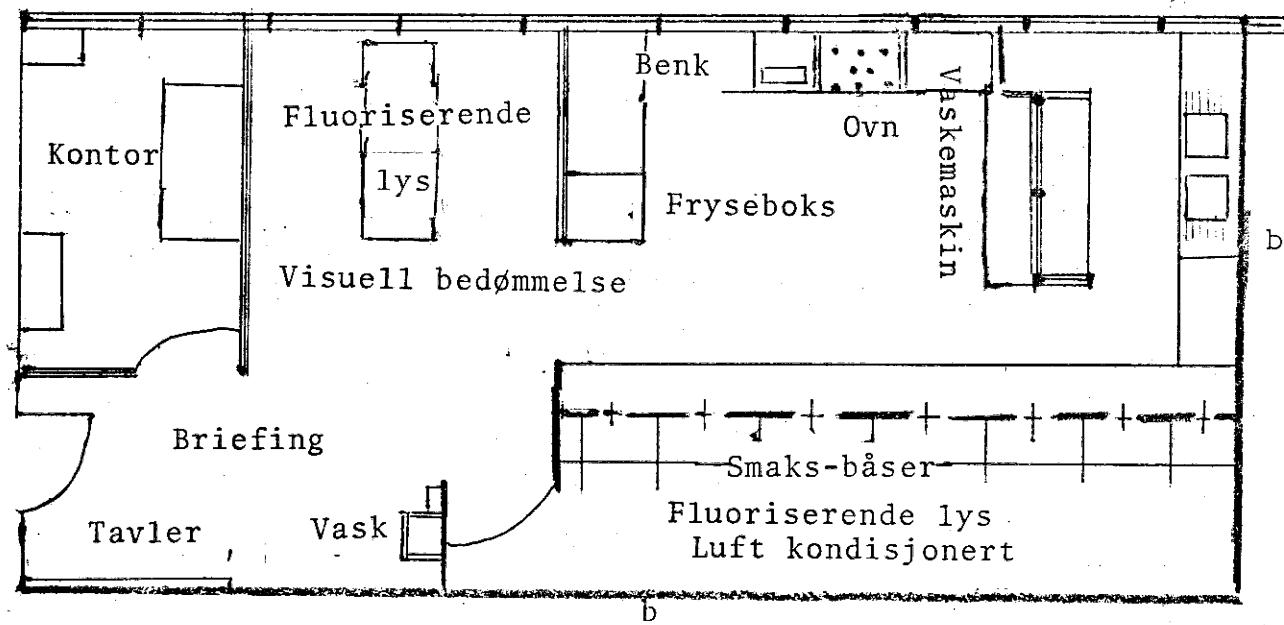
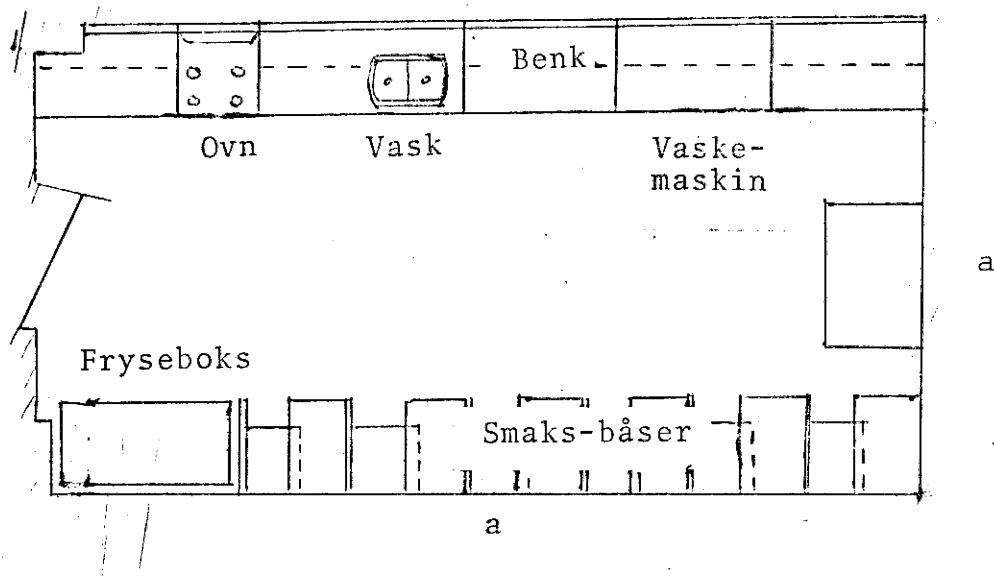
Tabell 4.1. Fordelingskoeffisienten for  $P_A \bar{x}_\lambda$ ,  $P_A \bar{y}_\lambda$  og  $P_A \bar{z}_\lambda$ 

(refleksjonene fra en perfekt hvit overflate ved lyskilde A.

(Hvitglødende lampe ved  $2848^\circ K$ ) for bølgelengde  $\lambda$ .

Bølgelengde i $\lambda$	$P_A \bar{x}_\lambda$	$P_A \bar{y}_\lambda$	$P_A \bar{z}_\lambda$	Bølgelengde i $\lambda$	$P_A \bar{x}_\lambda$	$P_A \bar{y}_\lambda$	$P_A \bar{z}_\lambda$
0.380	0.0006	0.0000	0.0029	0.580	4.8594	4.6139	0.0090
0.385	0.0011	0.0000	0.0053	0.585	5.3549	4.4668	0.0077
0.390	0.0024	0.0000	0.0113	0.590	5.7896	4.2704	0.0062
0.395	0.0047	0.0001	0.0224	0.595	6.1403	4.0379	0.0058
0.400	0.0097	0.0003	0.0463	0.600	6.3518	3.7733	0.0048
0.405	0.0174	0.0004	0.0825	0.605	6.4299	3.4855	0.0037
0.410	0.0356	0.0010	0.1699	0.610	6.3346	3.1780	0.0019
0.415	0.0694	0.0020	0.3319	0.615	6.0877	2.8622	0.0013
0.420	0.1308	0.0039	0.6283	0.620	5.6865	2.5358	0.0013
0.425	0.2269	0.0077	1.0974	0.625	5.1267	2.1901	0.0007
0.430	0.3246	0.0133	1.5840	0.630	4.4902	1.8523	0.0000
0.435	0.4055	0.0207	2.0036	0.635	3.8779	1.5529	0.0000
0.440	0.4632	0.0306	2.3236	0.640	3.2791	1.2812	0.0000
0.445	0.4976	0.0426	2.5484	0.645	2.7004	1.0344	0.0000
0.450	0.5155	0.0583	2.7173	0.650	2.1681	0.8183	0.0000
0.455	0.5230	0.0788	2.8621	0.655	1.7078	0.6372	0.0000
0.460	0.5097	0.1052	2.9254	0.660	1.3141	0.4861	0.0000
0.465	0.4690	0.1380	2.8539	0.665	0.9850	0.3625	0.0000
0.470	0.3882	0.1808	2.5581	0.670	0.7241	0.2651	0.0000
0.475	0.2998	0.2375	2.1979	0.675	0.5368	0.1958	0.0000
0.480	0.2138	0.3108	1.8179	0.680	0.4022	0.1461	0.0000
0.485	0.1372	0.4004	1.4575	0.685	0.2877	0.1041	0.0000
0.490	0.0799	0.5196	1.1622	0.690	0.2019	0.0729	0.0000
0.495	0.0387	0.6813	0.9308	0.695	0.1429	0.0515	0.0000
0.500	0.0136	0.8960	0.7545	0.700	0.1047	0.0377	0.0000
0.505	0.0070	1.1878	0.6191	0.705	0.0756	0.0271	0.0000
0.510	0.0285	1.5398	0.4843	0.710	0.0549	0.0199	0.0000
0.515	0.0934	1.9518	0.3585	0.715	0.0394	0.0144	0.0000
0.520	0.2127	2.3855	0.2627	0.720	0.0283	0.0097	0.0000
0.525	0.3849	2.7859	0.2012	0.725	0.0198	0.0069	0.0000
0.530	0.6069	3.1609	0.1547	0.730	0.0140	0.0050	0.0000
0.535	0.8631	3.4987	0.1140	0.735	0.0101	0.0041	0.0000
0.540	1.1567	3.7999	0.0809	0.740	0.0072	0.0031	0.0000
0.545	1.4904	4.0618	0.0555	0.745	0.0052	0.0021	0.0000
0.550	1.8660	4.2841	0.0375	0.750	0.0032	0.0010	0.0000
0.555	2.2887	4.4701	0.0255	0.755	0.0021	0.0010	0.0000
0.560	2.7550	4.6110	0.0181	0.760	0.0021	0.0010	0.0000
0.565	3.2564	4.6974	0.0130	0.765	0.0011	0.0000	0.0000
0.570	3.7853	4.7285	0.0104	0.770	0.0011	0.0000	0.0000
0.575	4.3259	4.7002	0.0092	0.775	0.0000	0.0000	0.0000
0.580	4.8594	4.6139	0.0090	0.780	0.0000	0.0000	0.0000
				Total	109.8472	1000000	35.5824





Figur 15. To utforminger av bedømmelsesrom.

## 6. HØRSELEN.

For et fåtall produkter kan det også være behov for å bruke ørene for å bedømme visse sider av produktet. Som eksempel kan vi ta knekkebrød, hvor vi venter bestemte knaskelyder under tyggingen, men hvor lyden ikke bør være for sjenerende. Det hender også at enkelte ostedommere banker i osten for via lyden å få et inntrykk av konsistens og hullsetning.

I følge SRINIVASAN's (27) undersøkelser er det også grunn til å tro at lyder under bedømmelser f.eks. fra maskineri etc. kan virke både kvalitativt og kvantativt på den organoleptiske respons.

## 7. UTSTYR VED BEDØMMELSE.

Som oftes foregår bedømmelse av produkter i lokaliteter som er langt fra tilfredsstillende. Ofte er en produsent helt **avhengig** av slike undersøkelser i sitt produktutviklingsarbeid. Her i huset blir det ofte brukt store summer på forsøk som bare kan gis en evaluering ved organoleptisk bedømmelse. Men vi har ikke noe rom som er utstyrt for bedømmelse.

Bedømmelsesrommet skal være så nøytralt som mulig med hensyn på farger, lukt, lys og lyd. Det bør være luft-kondisjonert og meget godt ventilert. Det bør også være et krav at en dommer kan bedømme helt uavhengig av de andre ): uten mulighet for å se eller høre de andres reaksjoner. Spesielle båser burde derfor være et krav.

I fig. 15 a og b har jeg skissert et enkelt prøverom og et noe mer avansert. Fortrinnvis bør hver plass ha vannkran og små vasker, men det er ikke absolutt nødvendig. Veggene som skiller de enkelte dommere fra hverandre kan være så lave at de kan ha felles lys. Lyset bør være av dagslys-type med mulighet for å sette på rød, gult eller grønt filter hvis en av oppgavene til dommerne er å bedømme fargeforskjeller. Det bør være plass for dommerne til å foreta en samlet diskusjon etter at prøvene er bedømt. Dette er svært viktig i utdannelsen av et dommerteam.

Prøvene serveres til de enkelte dommerne, og for serveringen kreves en eller to personer.

kvalifisert leder som om mulig ikke bør være med i dommerteamet

#### 8. SAMMENSETNING AV TESTPANEL.

I mindre bedrifter må en som ~~aftest~~ ta ut dommerne blant bedriftens ansatte. Det er naturlig at laboratoriepersonell blir engasjert i dette arbeidet. Alt etter hvilke oppgaver panelet far, kan det være nødvendig med 5-20 personer. Personer som er særlig ergasjert i et produkt, bør ikke være dommere for dette produktet fordi det er fare for at de ikke vil være nøytrale og også for at de øvrige kan la seg påvirke. Personer med spesielle kjeppester bør en også vise forsiktighet med i et team. En bakteriolog f.eks. vil ofte kunne legge stor vekt på feil som han mener kan ha bakteriologiske årsaker. En produktkonsulent bør heller ikke være med i et team som skal bedømme den varegruppen han er konsulent for. o.s.v. Det er derfor mange problemer ved utvelgelsen av personer til et ideelt bedømmelses-team.

For enkelte oppgaver kan det vel også være et spørsmål om hvor god bedømmeren egentlig bør være. Skal teamet vårt representere vanlige forbrukere kan det kanskje være uheldig å ha med personer som kan skjelne de minste forskjeller mellom to prøver. Den sistnevnte, virkelige eksperten er imidlertid meget verdifull. Spørsmålet er bare hvordan og når han skal brukes. Ekspertene har trolig sin plass i kontroll av råvarer og ferdigvarer om nå denne skjer internt eller externt. og i de første trinnene av produktutvikling. Før markedsføring er det imidlertid viktig hvis en ikke har råd til et større anlagt markedstest, å plukke ut mer eller mindre uøvede dommere og anse disse som representativt for konsumentene.

#### 9. UTVELGELSE OG TRENING AV DOMMERNE.

En naturlig begynnelse ved utplukking av dommerkandidater er å bruke prøver for smaksfølsomhet, enten ved evnen til å differensiere to prøver med liten forskjell i primære smaker (salt, søtt, surt, bittert) eller i bestemmelse av terskelverdier for disse.

Viktigere enn en slik primær-test er det å finne kandidatenes evne til å skjelne mellom små forskjeller i aktuelle næringsmidler. En rekke metoder er foreslått s.s. triangeltest med sekvens analyse, Kendall's koefisient for overensstemmelse etter rangeringstest korrelasjonskoefisienter ved parvise bedømmelser, o.s.v. Disse testene vil alle bli gjennomgått seinere. Det er også viktig at dommeren har evne til å gjenkjenne en del lukter som kan være vanlige i de produktene som skal bedømmes. Dette vil især være viktig i de såkalte profileringer hvor en skal splitte opp lukt og smaksinntrykk i enkelte komponenter og vurdere deres kvalitet, kvantitet og tidspunktet når de observeres.

Det er selvfølgelig viktig at kandidatene også er interessert i bedømmelsen og har evnen til å koncentrere seg om en bestemt oppgave. Ei god helse synes også å være vesentlig fordi en rekke syklige tilstander virker inn på terskelverdiene både for lukt og smak. Tidligere var det observert at kvinner hadde bedre smaksevne enn menn, men dette er forklart f.eks. ved forskjellige røykevaner. Seinere forsøk har vist at det neppe i dag er forskjell mellom kjønn eller alder.

#### 10. GENERELT OM BEDØMMELSER.

Det er viktig at bedømmelse av næringsmidler foregår ved de samme temperaturene som næringsmiddlet vanligvis blir konsumert ved. Vanligvis bruker en en konstant temperatur og den relative fuktigheten i bedømmelsesrummet bør også være omtrentlig konstant. Det er også vanlig at prøvene har stått ved en bestemt temperatur et døgn eller mer før bedømmelsen.

Det antallet prøver som dommerne med rimlighet kan klare i ett strekk, er selvfølgelig avhengig av næringsmiddlets art, men bør neppe overstige ca. 30.

Vanligvis er det et krav om at prøvene skal serveres i tilfeldig rekkefølge. Dette kan være uheldig fordi en meget sterkt-smakende prøve da kan påvirke bedømmelsen av de som følger umiddelbart etter. En dobbelt bedømmelse med forskjellig (tilfeldig) rekkefølge i de to seriene vil være å foretrekke, alternativt at dommerne får prøvene i forskjellig rekkefølge. Begge deler vil imidlertid medføre mer arbeid ved servering og koding.

Kodingen er et kapittel for seg. For ordinære poengbedømmelse er trolig en hver koding brukbar. Imidlertid kan det oppstå vanskeligheter ved preferanse- eller rangeringstest. Hvis to prøver nummereres 1 og 2 eller A og B så er det en tendens til at prøvene blir satt i denne rekkefølgen selv om det ikke er forskjell mellom den.

Det må være gjort klart på forhånd hvordan resultatene skal noteres og hva en skal legge vekt på i bedømmelsen slik at den enkelte dommeren slipper å spekulere på selve det tekniske opplegget under bedømmelser.

## II. EN DEL VANLIGE BEDØMMELSESESOPPLEGG.

Det er vanlig å stille bedømmelsesoppleggene i to grupper:

1. Analytiske opplegg hvor dommerne skal avgjøre hvorvidt de bedømte prøvene tilfredsstiller bestemte krav, hvorvidt de er like når det gjelder en bestemt egenskap o.s.v.. Dommernes egen oppfatning bør ikke komme fram i slike typer av bedømmelser. Denne typen av opplegg krever derfor et øvet panel av dommere som i og for seg har en vanskelig oppgave. Metodene innen denne gruppen av opplegg skal også være objektive og forskjellige panel skal i prinsippet gi samme svar.
2. Den andre gruppen av opplegg er ofte brukt ved konsummenttest, og her skal dommerne gi uttrykk for hvordan de liker prøven. Metodene kalles ofte hedoniske (etter nydelsesverdi) og er subjektive. Da konsummenttest er dyre å gjennomføre, brukes et dommerteam som oftest også til dette arbeid.

Skjematiske kan da metodene systematiseres slik:

A. Analytiske eller objektive metoder.	B. Hedoniske eller subjektive metoder.
I. Differanse-tester	I. Preferanse-tester
1. Enkelt-prøver	1. Par-prøver
2. Par-prøver.	2. Triangelprøver.
3. Triangelprøver	II. Bedømmelse/med hedonisk skala Scheffé's test.
4. Duotrioprøver.	
II. Rangordning-tester.	III. Rangordning.
III. Kvalitetstester.	
1. Graderings- eller scoring test.	

## 12. ENKELT - PRØVER.

Dommeren får forelagt en prøve og blir stilt et ja- eller nei spørsmål om denne prøven s.s.: "Er det en usmak på denneosten?"

Prøven er enkel og har en stor fordel. Det er nemlig slik vi vanligvis bedømmer et produkt. Vi har sjeldent to forskjellige kjøttkaker til middag eller forskjellige vaniljeiskremtyper til dessert.

Imidlertid har prøven en innebygd skjevhetsfaktor i den som får seg forelagt en prøve lurer nok på hva det er for galt med denne prøven og vil være påvirket av det. Det er også vanskelig å gjenta bedømmelsen for hver dommer. Og hvis nå 70 % av dommerne mener at det er noe galt med prøven, hvor mange av disse svarene er rein gjetning? Prøven kan derfor bare egne seg for dommere med lang tids øvelse f.eks. profesjonelle vin- og kaffesmakere etc. Ikke desto mindre har metoden vært benyttet ved konsummenttest før markedsføring av nye produkter.

## 13. PAR - PRØVER, DUOTRIO PRØVER etc.

Denne type av test karakteriseres ved at sjansen for på slumprå få et "riktig" resultat er 0,5.

Problemstillingen ved bruk av den egentlige par-prøven kan være f.eks.: Hvilke av de to foreliggende prøvene er søtest, saltest, surest etc.? Slike spørsmål kan stilles f.eks. ved bestemmelse av terskelverdier.

Vi kan også være interessert i å bruke så store mengder av en billig råvare som mulig og vil prøve oss fram til at prøven så vidt kan skiller fra prøver tatt av vår ordinære produksjon.

Vi kan være interessert i nye tilsetninger, nye produksjonsmetoder etc. som kanskje letter produksjonen, men som vi vil sikre oss mot ikke gir utslag i dårligere kvalitet. Spørsmålet kan i disse tilfellene være: Hvilke prøve er av vår ordinære produksjon?

For konsummenttest kan vi også bruke metoden hedonisk d.v.s. stille spørsmålet: Hvilke prøve er best?

Hver dommer kan få forelagt flere par slik at hans/hennes resultat kan bli statistisk vurdert. Vanligvis benyttes flere dommere og flere par pr. dommer.

Ved den statistiske evalueringen av resultatet stiller vi en nullhypotese om at det ikke er påvisbar forskjell mellom prøvene. Denne hypotesen forkastes hvis de oppnådde resultatene er mindre sannsynlig sett utfra en rein tilfeldig fordeling enn et forkastningsnivå som vi setter på forhånd. Vanligvis velger vi 5 %. Dette betyr da at det er 5 % eller mindre sjangse for at nullhypotesen er riktig hvis vi forkaster den.

Vi kan anta at  $K$  dommere utfører i bedømmelser hver. De kan da hver ha 0, 1, 2, ... i rette svar. Setter vi  $x_r$  = antall dommere med  $r$  rette svar, vil

$$0 \leq r \leq i$$

og

$$\sum x_r = K$$

Sannsynligheten for at en dommer gir  $r$  rette svar er da:

$$P_r = \binom{i}{r} p^r q^{(i-r)} \text{ hvor } p = q = 0,5$$

Da er sannsynligheten for det resultatet vi har fått av alle  $K$  dommere:

$$P_k(x_0, x_1, \dots, x_i) = \frac{K!}{x_0! \dots x_i!} p_0^{x_0} \dots p_i^{x_i}$$

som er den multinomiale fordelingslov.

For å finne hvordan fordelingen av de riktige svarene må være for å få sannsynligheten mindre enn et gitt forkastningsnivå,  $\alpha$ , setter vi :

$$\sum \frac{K!}{x_r x_0! \dots x_i!} p_0^{x_0} \dots p_i^{x_i} \leq \alpha$$

Her setter vi inn verdiene for  $x_r$  fra i til 0 for  $x_i$  sammen med de mulige kombinasjonene for  $x_0, \dots, x_{i-1}$  og finner de kombinasjonene som gjør summen mindre enn  $\alpha$ .

Metoden tar altså hensyn til hvordan de riktige svarene er fordelt på dommerne, men krever et ikke ubetydelig regnearbeid.

Hvis ikke materialet er for lite, bruker vi heller den binomiale fordelingsloven og setter:

$$\sum_{x=m}^{x=n} \binom{n}{x} p^x q^{n-x} \leq \alpha$$

hvor n er samtlige prøver ( $K \cdot i$ ) og m er det minste antallet av riktige svar som tilfredsstiller ulikheten.

Som eksempel kan vi ta følgende: 4 dommere skal bedømme 3 par-prøver hver. Vi finner da:

$$P_0 = \binom{3}{0} 0,5^0 0,5^3 = 0,125$$

$$P_1 = \binom{3}{1} 0,5^1 0,5^2 = 0,375$$

$$P_2 = \binom{3}{2} 0,5^2 0,5^1 = 0,375$$

$$P_3 = \binom{3}{3} 0,5^3 0,5^0 = 0,125$$

Altså skal:

$$\sum_{x_r} \frac{4!}{x_0! \dots x_3!} 0,125^{x_0} 0,375^{x_1} 0,375^{x_2} 0,125^{x_3} \leq 0,05$$

Vi summerer her ovenfra d.v.s. fra maksimalt antall rette.

Følgende tabell gir resultatene:

Antall rette	$x_0$	$x_1$	$x_2$	$x_3$	$\frac{4!}{x_0! \dots x_3!}$	potensene	produktet	Sum
					$x_0! \dots x_3!$			
12	0	0	0	4	1	0,00024	0,00024	0,00024
11	0	0	1	3	4	0,00073	0,00292	0,00316
10	0	1	0	3	4	0,00073	0,00292	0,00608
10	0	0	2	2	6	0,00220	0,01320	0,01928
9	1	0	0	3	4	0,00024	0,00096	
9	0	1	1	2	12	0,00220	0,02640	
9	0	0	3	1	4	0,00659	0,02637	

Ulikheten er tilfredsstilt for 10 eller flere rette og når 1 dommer har ingen rette og 3 har tre rette og når 1 dommer har en rett, 1 dommer to rette og to dommere har tre rette. Den siste situasjonen kan forøvrig byttes med en situasjon hvor 3 dommere har to rette og 1 dommer har tre rette.

Utregnet ved den binomiale fordelingsfunksjon får vi:

Antall rette = r	$\binom{n}{x}$	$p^x q^{n-x}$	produkt	sum
12=n	1	0,00024	0,00024	0,00024
11	12	0,00024	0,00288	0,00312
10=r <sub>o</sub>	66	0,00024	0,01584	0,01896
9	220	0,00024	0,05280	0,07176

Ulikheten tilfredsstilles altså for 10 eller flere riktige:

I stedet for binomialfordelingen, kan en med store verdier for n bestemme  $r_o$  ved normalfordelingen i det

$$\frac{r_0 - np \pm 0,5}{\sqrt{npq}} < \mu_\alpha$$

hvor  $\mu_\alpha$  fåes fra tabell over normalfordelingen og 0,5 er korreksjon for overgangen fra diskontinuerlig til kontinuerlig funksjon.

Nullhypotesen forkastes hvis  $r_{obs} \geq r_o$

Vi kan også bruke  $\chi^2$  fordelingen i det

$$\chi^2 = \frac{(r_o - np)^2}{np} + \frac{[(n - r_o) - nq]^2}{nq} \leq \chi^2_{\alpha}$$

med 1 frihetsgrad. Nullhypotesen forkastes hvis  $r_{obs} \geq r_o$

#### 14. TRIANGEL - TEST.

Ved triangel-testet presenteres tre prøver for dommerne, gjerne nummerert 1,2,3 eller A,B,C, etc. To av prøvene er like, og det er dommerens oppgave å finne den av prøvene som er ulik de andre. Sjansen for å finne denne prøven uten at det finnes noen forskjell som dommeren kan merke, er  $\frac{1}{3}$ .

For beregningene brukes de samme formelene som for par-prøver, men her er da  $p = \frac{1}{3}$  og  $q = \frac{2}{3}$ . Det er værd å merke seg at dommerne skal avgive svar; hvis de ikke kan bestemme seg, skal de gjette.

Det er klart at jo flere prøver en har, desto sikrere resultat vil en få. Ved beregninger med små n kan det også være usikkert om en skal forkaste et resultat eller ikke.

Imidlertid er spørsmålet om å øke n ofte et spørsmål om ekstra omkostninger. Selve prøvene kan være kostbare som ved forskjellige kapitalvarer, og det å bedømme flere prøver tar tid og koster penger.

WALD (29) utarbeidet under krigen en metode for å finne ut hvor stort antall en trenger av prøver før en har grunnlag til å ta en bestemt avgjørelse om en kan forkaste eller godta en  $H_0$ -hypotese om et produkts brukbarhet. Denne metodikken "sekvensiell analyse" ble vurdert til å være så betydningsfull at den ble gradert som topphemmelig.

## PAR-TESTER.

Antall bedømmelser	Minimum antall korrekte svar for signifikant utslag.		
	Sannsynlighets nivå		
	0,05	0,01	0,001
7	7	7	7
8	7	8	8
9	8	9	9
10	9	10	10
11	9	10	11
12	10	11	12
13	10	12	13
14	11	12	13
15	12	13	14
16	12	14	15
17	13	14	16
18	13	15	16
19	14	15	17
20	15	16	18
21	15	17	18
22	16	17	19
23	16	18	20
24	17	19	20
25	18	19	21
30	20	22	24
35	23	25	27
40	26	28	31
45	29	31	34
50	32	34	37
60	37	40	43
70	43	46	49
80	48	51	55
90	54	57	61
100	59	63	66

## Triangel - test.

Kritiske verdier av r for p = 1/3 og  $\alpha = 0,05, 0,01$  samt 0,001.

Antal bedømmelser	Ensidig test			Dobbelsidig test		
	0,05	0,01	0,001	0,05	0,01	0,001
1	-	-	-	-	-	-
2	-	-	-	-	-	-
3	3	-	-	3	3	-
4	4	-	-	3	4	-
5	4	5	-	4	4	5
6	5	6	-	4	5	6
7	5	6	7	4	5	6
8	6	7	8	5	5	6
9	6	7	8	5	6	7
10	7	8	9	5	6	7
11	7	8	10	5	6	8
12	8	9	10	6	7	8
13	8	-	11	6	7	8
14	9	10	11	6	7	9
15	9	10	12	7	8	9
16	9	11	12	7	8	9
17	10	11	13	7	8	10
18	10	12	13	7	9	10
19	11	12	14	8	9	10
20	11	13	14	8	9	11
21	12	13	15	8	9	11
22	12	14	15	8	10	11
23	12	14	16	9	10	12
24	13	15	16	9	10	12
25	13	15	17	9	10	12
30	15	17	19	10	12	13
40	19	21	24	13	14	16
50	24	26	28	14	16	18
75	34	36	39	19	21	24
100	43	46	49	24	27	29

## 15. SEKVENSIELL ANALYSE FOR BESTEMMELSE AV ANTALL PRØVER.

I mange statistiske forsøksopplegg velger en antall gjentak i forsøksplanen uten egentlig noe kriterium på at antallet er for lite eller for stort. I første tilfelle må en kanskje gjenta hele forsøket, i det andre tilfellet har en kanskje brukt unødvendig mye tid og penger.

WALD (24) introduserte først såkalt sekvensiell analyse for å finne ut hvor mange observasjoner som var tilstrekkelig og nødvendige for å komme fram til et resultat med en bestemt sikkerhet. Hans utgangspunkt ved taklingen av problemet generelt var følgende:

Vi har en rekke med observerte variable  $(x_1, x_2, \dots, x_n)$  som vi i prinsippet kan øke i det uendelige. Hvor mange av disse variable vi trenger å observere er en del av det statistiske problemet.

Vi antar at fordelingsfunksjonen for de variable er kjent bortsett fra en ukjent parameter  $\theta$  som må ligge i et område  $\Omega$ .

Det skal nå gjøres et valg mellom to avgjørelser  $d_0$  og  $d_1$ . To underområder av  $\Omega$ , henholdsvis  $w_0$  og  $w_1$ , spesifiseres slik at

hvis  $\theta \in w_0$ , tar vi avgjørelsen  $d_0$   
hvis  $\theta \in w_1$ , tar vi avgjørelsen  $d_1$

i det  $w_0$  og  $w_1$  er preferenseområder for henholdsvis  $d_0$  og  $d_1$  mens  $\Omega - (w_0 \cup w_1)$ , er et indifferent område. Hvis  $\theta$  ligger i dette området, spiller det liten rolle hvilken avgjørelse som tas.

Etter å ha observert  $x_1, x_2, \dots, x_n$  kan vi velge mellom å:  
fortsette med å observere  $x_{n+1}$ ,  
ta avgjørelsen  $d_0$ ,  
ta avgjørelsen  $d_1$ .

Det må derfor utarbeides en regel for når en skal stoppe med å observere og ta den endelige avgjørelsen. En slik regel kalles en sekvensiell test.

Det er som regel tre krav som stilles til en slik regel.

$$1. P_r(N<\infty) = 1$$

d.v.s. vi må før eller siden få nok observasjoner.

2. Sannsynligheten for å ta avgjørelsen  $d_1$  når  $d_0$  er riktig må være liten:

$$P_r(d_1|\theta \in \omega_0) \leq \alpha$$

hvor  $\alpha$  er et lite tall.

3. Tilsvarende må sannsynligheten for å ta avgjørelsen  $d_0$  være liten hvis  $d_1$  er riktig.

$$P_r(d_0|\theta \in \omega_1) \leq \beta$$

Vi kan som et eksempel, anta at en iskremprodusent har kommet til at han vil velge mellom to typer syltetøy. Det ene syltetøyet (A) er noe dyrere enn det andre (B). Han bestemmer seg da for å bruke A (avgjørelsen  $d_1$ ) hvis det i 65 av 100 parprøver viser seg å være bedre enn B ( $p_1 = 0,65$ ). Hvis imidlertid A er bedre bare i 55 av 100 parprøver ( $p_0 = 0,55$ ) vil han bruke B (avgjørelsen  $d_0$ ). (Dette er nullhypotesen). Hvis  $p$  ligger mellom 0,55 og 0,65 regner han med at det vil bli det samme hvilket syltetøy han bruker i det gevinsten ved et bedre syltetøy oppveies av tapet ved høyere kostnad. Produsenten velger også en  $\alpha = 0,05$  hvor  $\alpha$  er sannsynligheten for å få en  $p$ -verdi som er større enn 0,65 når den i virkeligheten er mindre enn 0,55. Han velger også  $\beta = 0,10$  hvor da  $\beta$  er sannsynligheten for at  $p$  er mindre enn 0,55 når den i virkeligheten er større enn 0,65.  $\alpha$  kalles ofte for produsentens risiko mens  $\beta$  er konsumentens risiko.

For å undersøke dette problemet produseres  $n$  parprøver, og ved bedømmelsen blir  $r$  av A bedømt som best. Det har vist seg at sannsynlighetsforholdet:

$$L_n = \frac{P(r_n \cdot p_\phi)}{P(r_n \cdot p_0)}$$

er en egnet størrelse til å teste vårt problem med.

Her er telleren ( $P(r_n, p_1)$ ) sannsynligheten for r "riktige" av n under forutsetningen  $p = p_1$  mens nevneren er sannsynligheten for r "riktige" svar av n parprøver under forutsetningen  $p = p_0$ .

Hvis nå  $d_o$  er riktig, så må

$$P(r_n, p_0) \leq 1 - \alpha$$

Vi krever at  $P(r_n, p_1) \geq \beta$  og får at den største verdien for  $L_n$  hvor vi kan godta  $d_o$  blir:

$$L_{no} = \frac{\beta}{1 - \alpha}$$

Hvis  $d_1$  er sann, så er

$$P(r_n, p_1) \geq 1 - \beta$$

mens vi krever at

$$P(r_n, p_0) \leq \alpha \text{ og}$$

$$L_{nl} = \frac{1 - \beta}{\alpha}$$

blir da den minste verdien for  $L_n$  som vi kan godta.

Hvis nå  $L_n \leq L_{no}$  aksepteres nullhypotesen

$L_n \geq L_{nl}$  forkastes "

og hvis  $L_{no} < L_n < L_{nl}$

må vi ta nye observasjoner.

Under forutsetning av at observasjonene er tatt uavhengig av hverandre, er for både partest, triangeltest og duo-trio-test:

$$\frac{P(r, p_1)}{P(r, p_0)} = \frac{\binom{n}{r} p_1^r q_1^{n-r}}{\binom{n}{r} p_0^r q_0^{n-r}} \quad (q_1 = 1 - p_1) \text{ og } (q_0 = 1 - p_0)$$

hvor n er antall triangler og r er antall riktige bedømmelser.

(p, og  $p_0$  vil selvfølgelig bli valgt forskjellig etter testtypen og etter kjennskap til problemets art)..

Dette gir:

$$L_n = \left( \frac{P_1}{P_0} \right)^r \cdot \left( \frac{q_1}{q_0} \right)^{n-r}$$

Logaritmering gir:

$$\log L_n = r \log \left( \frac{P_1}{P_0} \right) + (n-r) \log \left( \frac{q_1}{q_0} \right) = n \log \left( \frac{q_1}{q_0} \right) + r \log \left( \frac{P_1 q_0}{P_0 q_1} \right)$$

Hvis  $d_0$  skal godtas må:

$$\log \frac{\beta}{1-\alpha} \leq n \log \left( \frac{q_1}{q_0} \right) + r \log \left( \frac{P_1 q_0}{P_0 q_1} \right) \text{ eller}$$

$$r \leq \log \frac{\beta}{1-\alpha} / \log \left( \frac{P_1 q_0}{P_0 q_1} \right) + n \log \left( \frac{q_0}{q_1} \right) / \log \left( \frac{P_1 q_0}{P_0 q_1} \right)$$

$$r \leq k_1 + nk_2$$

Hvis vi avsetter kumulative riktige svar (r) mot antall observasjoner n i diagram får vi altså en rett linje.

Hvis  $d_0$  skal forkastes, må:

$$\log \frac{1-\beta}{\alpha} \leq n \log \left( \frac{q_1}{q_0} \right) + r \log \left( \frac{P_1 q_0}{P_0 q_1} \right), \text{ eller}$$

$$r \geq \log \frac{1-\beta}{\alpha} / \log \left( \frac{P_1 q_0}{P_0 q_1} \right) + n \log \left( \frac{q_0}{q_1} \right) / \log \left( \frac{P_1 q_0}{P_0 q_1} \right)$$

$$r \geq k_3 + nk_2$$

$$\text{hvis } (k_1 + nk_2) < r < (k_3 + nk_2)$$

må vi ta flere observasjoner.

Hvis vi derfor tegner opp to parallelle rette linjer med stigningsforholdet  $k_2$  og skjæringspunkter med y-aksen henholdsvis  $k_1$  og  $k_3$  får vi et interval. Hvis plotting av r mot n gir et punkt på eller ovenfor den øverste linja, forkastes 0-hypotesen. Får vi et punkt på eller under den laveste linja godtar vi 0-hypotesen. Faller punktet mellom linjene fortsetter vi med ny prøveomgang. (Sett inn verdiene fra eksemplet foran !)

Hvis vi skal bruke sekvensiell analyse i par eller triangel-test må vi derfor på forhand bestemme oss for:

$p_0$  = den største sannsynligheten for riktige svar som vi ikke vil godta.

$p_1$  = den minste sannsynligheten for riktige svar som vi vil akseptere.

$\alpha$  = Sannsynligheten for å forkaste  $H_0$ -hypotesen når den er riktig.

$\beta$  = Sannsynligheten for å godta  $H_0$ -hypotesen når den er feil.

Før vi bestemmer oss for verdiene for  $p_0, p_1$ ,  $\alpha$  og  $\beta$  kan det være nyttig å beregne gjennomsnittlig antall prøver vi må ha for å få utslag. Hvis  $E_{p(n)}$  er gjennomsnittlig antall prøver for en dommer med treffsikkerhet  $p$ , vil vi etter Wald (29) ha for:

$P = P_0$  (maksimum uakseptabel treffsikkerhet)

$$EP_0(n) = \frac{(1-\alpha)\log\frac{\beta}{1-\alpha} + \alpha\log\frac{1-\beta}{\alpha}}{P_0 \log\frac{P_1}{P_0} + q_0 \log\frac{q_1}{q_0}}$$

\* og for

$P = P_1$  (minimum akseptabel treffsikkerhet)

$$EP_1(n) = \frac{\log\frac{\beta}{1-\alpha} + (1-\beta)\log\frac{1-\beta}{\alpha}}{P_1 \log\frac{P_1}{P_0} + q_1 \log\frac{q_1}{q_0}}$$

For  $p = 0$  (påvisningen av ingen treffsikkerhet) har vi en forventning,

$$E_0(n) = \frac{\log\frac{\beta}{1-\alpha}}{\log\frac{q_1}{q_0}} \quad \text{og for}$$

$p = 1$  (påvisning av ufeilbarlighet),

$$E_1(n) = \frac{\log\frac{1-\beta}{\alpha}}{\log\frac{P_1}{P_0}}$$

Den sekvensielle analysen kan vi også bruke for å beregne hvor mange prøvesmakinger som er nødvendige for å forkaste eller godkjenne en 0-hypotese om selve prøvene.

Som eksempel skal vi se på resultatet fra et triangelttest hvor 15 dommere skulle bestemme om det var forskjell mellom appelsinsaft tilsatt henholdsvis 8 og 9 % sukker. De enkelte "omgangsresultatene" går fram av tabellen.

Prøveomgang:	Riktige svar:	Prøveomgang:	Riktige svar:
1	9	10	7
2	6	11	4
3	4	12	9
4	4	13	3
5	7	14	9
6	5	15	8
7	7	16	3
8	5	17	6
9	8		

Setter vi nå:::

$$\alpha = 0,05, \beta = 0,10, p_0 = 0,40, p_1 = 0,70$$

får vi:

1. for godkjenning av nullhypotesen:

$$r \leq -1,8 + 0,55 n$$

2. for forkasting av nullhypotesen

$$r \geq 2,3 + 0,55 n$$

Intervallet med de observerte verdiene er tegnet opp i figur (16). Tre prøveomganger hadde her vært nok til å godta nullhypotesen.

Beregningsmessig finner vi:

$$E_{0,7}(n) \approx 19$$

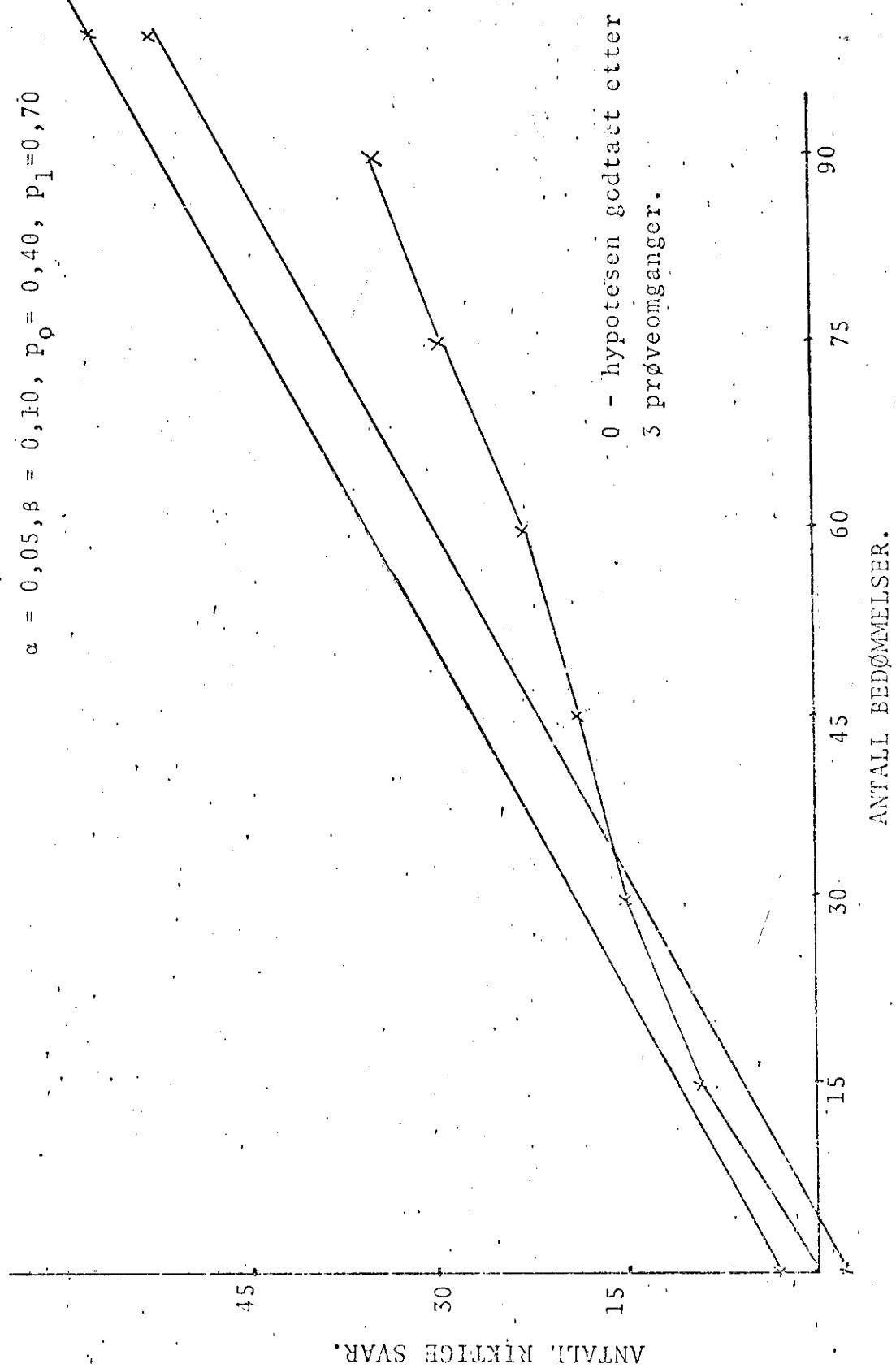
$$E_{0,4}(n) \approx 11$$

Figur 16.

SEKVENSIELL -TRIANGELTEST  
APPELSINSAFT MED 8 , RESP. 9 % SUKKER.  
RESULTAT PR. PRØVEOMGANG FOR 15 DOMMERE.

$$\alpha = 0,05, \beta = 0,10, p_0 = 0,40, p_1 = 0,70$$

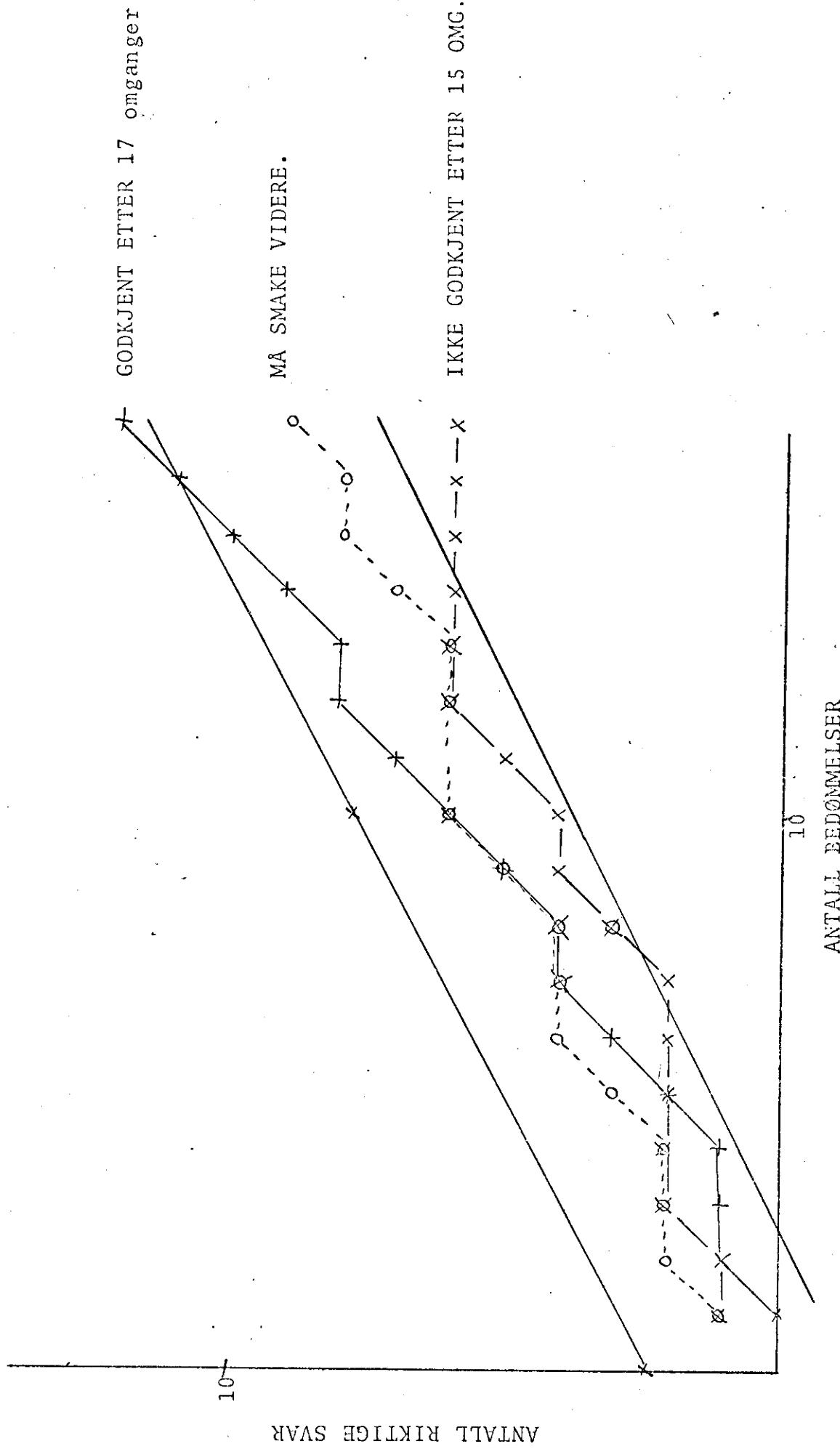
54



Figur 17.  
SEKVENTIELT TRIANGELTEST.

APPELSINSAFT MED 8, RESP. 9 % SUKKER.  
RESULTAT FOR 3 DOMMERER.

$$\alpha = 0,05, \beta = 0,10, p_0 = 0,40, q_1 = 0,70$$



De enkelte dommerne ble også testet ved sekvensiell analyse:

1. ble godkjent etter 8 omganger

1. ble " " 17 " (se figur 17.)

6 lå innenfor intervallet etter 17 omganger (se figur 17.)

1 ble underkjent etter 5 omganger

3 " " " 9 "

1 " " " 11 "

1 " " " 15 " (se figur 17.)

## 16. RANGORDNINGSPRØVER.

Mens en ved testene under A I. bare behandler maksimalt 2 prøver ad gangen og hvor innføring av flere prøver vil by på store problemer, er rangordnings-prøvene nettopp beregnet på undersøkelse av mange prøver.

Nå er det selvsagt en begrensning i antall prøver en dommer er i stand til å rangere, alt avhengig av prøvenes art, av forskjel-lens størrelse, etc. En samvittighetsfull rangering fordrer også at en og samme prøve blir smakt på flere ganger, noe som også vil virke begrensende på antallet prøver.

I en analytisk metode vil dommerne rangere prøvene etter styrken på en bestemt egenskap (syhet, surhet, saltet, o.s.v.), og oppdragsgiverens formål kan være å teste virkningen av forskjellige koncentrasjoner av tilsetninger, forskjellige fremstillings-måter o.l.

I slike tilfeller får vi direkte rangeringstest. Vi kan også få indirekte rangeringstest ved å benytte måleverdier eller verdier funnet ved scorings test e.l.

### 15.1. Direkte rangeringstest.

#### To serier.

Forutsetningen er her at n prøver rangeres etter en bestemt egenskap av to dommere eller etter to egenskaper av en dommer.

Prøve 1, 2 .....n

1. bedømmelse  $r_{11} r_{12} \dots r_{1n}$

2. "  $r_{21} r_{22} \dots r_{2n}$

Kendall's (11) rangeringskoeffisient,  $\tau$ , defineres ved

$$\tau = \frac{S}{\binom{n}{2}} = \frac{S}{\frac{1}{2}(n-1)}$$

For hvert observasjonspar setter en poeng, p, som er + 1 hvis begge observasjonene har samme innbyrdes rangering  $\neq$  1 ellers:

$S = \sum_p$  for det  $\binom{n}{2}$  mulige parrett

S kan da variere mellom  $\pm$  og  $\binom{n}{2}$  og  $\tau$  derfor mellom  $\pm 1$  og  $\pm 1$ .

Signifikansgrensen for S kan finnes i tabeller. Hvis  $n > 10$  blir S tilnærmet normalfordelt med middel 0 og varians  $n(n-1)(2n+5)/18$

$$\mu = \frac{S - 1}{\sqrt{n(n-1)(2n+5)/18}}$$

og kritisk verdier for  $\mu$  kan finnes i tabeller over normalfordelingen.

Ved beregningen av S (eller  $\tau$ ) er det enklest å sette den ene serien i stigende orden, slik som i eksemplet:

E	C	A	I	F	D	G	J	B	H
1	2	3	4	5	6	7	8	9	10
2	7	4	9	6	1	5	3	10	8

En betrakter nå bare den nederste linja.

$$S = P - Q$$

$$P = 8 + 3 + 5 + 1 + 2 + 4 + 2 + 2 + 0 = 27$$

$$Q = 1 + 5 + 2 + 5 + 3 + 0 + 1 + 0 + 1 = 18$$

$$S = 9$$

Tallene er fremkommet ved at en i P-rekka har satt opp antall tall høyere enn tallet selv til høyere for hvert tall og i Q-rekka antall tall mindre enn tallet selv til høyere for hvert tall.

Kritiske verdier for S og  $\tau$  på 5 % nivået for  $\leq 10$  er satt opp i tabell (15.1.).

#### Spearman's rangkorelasjonskoeffisient.

$\rho$ , defineres ved:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

hvor  $d_i = (r_{1,i} - r_{2,i})$ : differensen mellom rangeringsnummer for objekt i i 1. og 2. rangering.

Vi ser lett at hvis alle par rangeres på samme måten, er  $d_i = 0$  og  $\rho = 1$ .

Hvis vi har maksimal spredning i rangeringene d.v.s.:

1. rangering  $x_1 \ x_2 \ \dots \ x_n$

2. rangering  $x_n \ x_{n-1} \ \dots \ x_1$

så er:

$$d_1 = 1 - n = -(n-1)$$

$$d_2 = 2 - (n-1) = -(n-3)$$

$$d_{n-1} = (n-1)-2 = n-3$$

$$d_n = n-1$$

Den bakerste loddrette kolumnen er da ei rekke, med  $\sum d_i = 0$ , og  $\sum d_i^2 = \frac{n(n^2-1)}{3}$  som er den største verdien  $\sum_{l=1}^n d_i^2$  kan få, og da er  $\rho = -1$ .

Derfor er alltid  $-1 \leq \rho \leq 1$

For stor n er  $\rho$  tilnærmet normalfordelt. For mindre n har vi satt opp de kritiske verdiene av  $\rho$  på 5 %-nivået i tabell (15.1).

### To eller flere serier.

Vi antar at n objekter rangeres etter en egenskap av m dommere. Dette er et meget vanlig opplegg.

Objekt	1	2	.....	n
1.dommer	$r_{11}$	$r_{12}$	.....	$r_{1n}$
i.dommer	$r_{i1}$	$r_{i2}$	.....	$r_{in}$
m.dommer	$r_{m1}$	$r_{m2}$	.....	$r_{mn}$

Hvis alle rangeringene er identiske blir summen av rangeringene for objektene:  $m, 2m, \dots, nm$ . Summen av alle blir:  $m(1+2+\dots+n) = nm(n+1)/2$  og middelet  $\frac{nm(n+1)}{2n} = m(n+1)/2$

Avvikelsen fra middeltallet blir:

$$m - m(n+1)/2 = \frac{2m - mn - n}{2} = \frac{1}{2}m(n-1)$$

$$2m - m(n+1)/2 = \frac{4m - mn - n}{2} = \frac{1}{2}m(n-3)$$

$$mn - m(n+1)/2 = \frac{2mn - mn - m}{2} = +\frac{1}{2}m(n-1)$$

Kvadratsummen av avvikelsene blir da:

$$m^2(n^3 - n)/12 \text{ som er den største verdien denne kan få.}$$

KENDALL (11) definerte en "coefficient" of concordance", W.

$$W = \frac{S}{\frac{m^2(n^3-n)}{12}} = \frac{12S}{m^2(n^3-n)}$$

hvor S er den verdien vi finner for kvadratsummen for de observerte avvikere. W vil ligge mellom 0 og 1.

For lave verdier av m og n er det utarbeidet tabeller for testing.

For  $n > 5$  kan en tilnærmet sette

$$F = \frac{(m-1)W}{1-W}$$

med frihetsgradene  $(n-1-2/m)$  og  $(m-1)/(n-1-2/m)$ .

For  $n > 7$  kan en også bruke  $\chi^2$  fordelingen:

$$\chi^2 = \frac{m(n-1)W}{mn(n+1)}$$

Dette er altså et "overall" test som kan vise om det er forskjell mellom prøvene i hele rangerings-serien. Spørsmålet om hvilke prøve(r) som avviker fra de øvrige blir ikke bestemt. Nå er det imidlertid også utviklet test for dette som kan brukes under forutsetning av at hele serien viser forskjell.

Signifikansgrensen er her satt opp i tabell-form (tabell 18.2.) for den rangeringssummen som en må over eller under ved et vist antall rangeringer og dommere, for å kunne si om prøven er dårligere eller bedre enn ventet utfra hypotesen om en tilfeldig rangering.

Som et eksempel skal vi ta et rangeringstest for pølser hvor 18 dommere rangerte 7 forskjellige pølsetyper 2 ganger.

Vi fikk følgende resultat for rangeringen:  
Prøve nr.:

Rangering:	4	3	7	2	6	1	5
1	37	62	65	68	73	99	100
2	27	61	66	65	75	112	98

Som en ser er rangeringssummene ganske like for hver prøve i de to seriene. For rekkene blir middelet:

$$r = \frac{18 \cdot (7-1)}{2} = 72 \text{ og for rangering nr. 2 er:}$$

$$S = \sum (r_i - r)^2 = 4516$$

$$W = \frac{4516 \cdot 12}{324 \cdot 336} = 0,4978$$

$$F = \frac{17 \cdot 0,4978}{17 \cdot 5022} \approx 16,86$$

med  $7-1-0,1 = 5,9 \approx 6$

og  $17/5,9 = 2,9 \approx 3$  frihetsgrader.

For lave verdier av m og n er det utarbeidet tabeller for testing.

For  $n > 5$  kan en tilnærmet sette

$$F = \frac{(m-1)W}{1-W}$$

med frihetsgradene  $(n-1-2/m)$  og  $(m-1)/(n-1-2/m)$ .

For  $n > 7$  kan en også bruke  $\chi^2$  fordelingen:

$$\chi^2 = \frac{m(n-1)W}{mn(n+1)}$$

Dette er altså et "overall" test som kan vise om det er forskjell mellom prøvene i hele rangerings-serien. Spørsmålet om hvilke prøve(r) som avviker fra de øvrige blir ikke bestemt. Nå er det imidlertid også utviklet test for dette som kan brukes under forutsetning av at hele serien viser forskjell.

Signifikansgrensen er her satt opp i tabell-form (tabell 16.2.) for den rangeringssummen som en må over eller under ved et vist antall rangeringer og dommere, for å kunne si om prøven er dårligere eller bedre enn ventet utfra hypotesen om en tilfeldig rangering.

Som et eksempel skal vi ta et rangeringstest for pølser hvor 18 dommere rangerte 7 forskjellige pølsetyper 2 ganger.

Vi fikk følgende resultat for rangeringen:

Prøve nr.:

Rangering:	4	3	7	2	6	1	5
1	37	62	65	68	73	99	100
2	27	61	66	65	75	112	98

Som en ser er rangeringssummene ganske like for hver prøve i de to seriene. For rekkenes blir middelet:

$$r = \frac{18 \cdot (7-1)}{2} = 72 \text{ og for rangering nr. 2 er:}$$

$$S = \sum (r_i - \bar{r})^2 = 4516$$

$$W = \frac{4516 \cdot 12}{324 \cdot 336} = 0,4978$$

$$F = \frac{17 \cdot 0,4978}{5022} = 16,86$$

med  $7-1-0,1 = 5,9 \approx 6$

og  $17/5,9 = 2,9 \approx 3$  frihetsgrader.

F-verdien er derfor signifikant på 1 % nivået.

Går vi inn i tabell (l6'.2) finner vi for begge rangeringsrekkenes at pølse nr. 4 har blitt rangert bedre enn en skulle vente og 1 og 5 dårligere utfra hypotesen om tilfeldig rangering. Forkastningsnivået er også her 1 %.

Selv om altså rangeringen for alle dommerne ga signifikant utslag, var det bare to dommere av de 18 som rangerte prøvene i de to seriene på en slik måte at det var mindre enn 5 % sannsynlighet for at rangeringen var tilfeldig.

En av disse dommerne fikk dette resultatet:

Rangering.	Prøve nr.						
Omgang.	4	2	3	7	6	5	1
1	1	2	3	4	5	6	7
2	1	3	2	4	6	5	7
Differense:	0	-1	1	0	-1	1	0

Prøvene er i tabellen ordnet slik som dommeren rangerte dem i 1. prøveomgang.

Her er da:

$$P = 6+4+4+3+1+1 = 19$$

$$Q = 0+1+0+0+1+0 = \underline{\underline{2}}$$

$$\underline{\underline{S = 17}}$$

Kendall's rangeringskoeffisient blir:

$$\pi = \frac{17}{\frac{1}{2} \cdot 7 \cdot 6} = \frac{17}{21} = \underline{\underline{0,81}}$$

Spearman's rangkoorelasjonskoeffisient:

$$d_1 = 0, d_2 = -1, d_3 = 1, d_7 = 0, d_5 = -1, d_6 = 1 \text{ og } d_7 = 0$$

$$\sum d_i^2 = 4$$

$$\rho = 1 - \frac{6.4}{7.48} = 0,93$$

Både  $S$ ,  $\tau$  og  $\rho$  er altså i hvert fall signifikante på 5 % nivået.

### 15.2. Indirekte rangeringsserier.

#### a. 2 stikkprøver.

Vi antar at vi har tatt ut  $n_1$  og  $n_2$  individuelle prøver fra to stikkprøver og analysert disse med hensyn til en variabel. Testoppgaven er da å teste om en kan si at to prøvene sannsynligvis stammer fra en og samme populasjon.

#### WILCOXON's (30) test.

Det totale ( $n = (n_1 + n_2)$ ) antallet observasjoner ordnes etter måleverdier og hver gis rangeringsnummer fra 1 til  $n$ . For hver av observasjonsseriene  $n_1$  og  $n_2$  beregnes rangeringssummene  $\Sigma r_1$  og  $\Sigma r_2$ . Vi lar så  $T$  være den minste av disse. Den blir da lik eller større enn den minste av  $n_1(n_1+1)/2$  eller  $n_2(n_2+1)/2$ .  $\Sigma r_1 + \Sigma r_2 = n(n+1)/2$ .

For en del kombinasjoner av  $n_1$  og  $n_2$  finnes det tabeller over kritiske verdier for  $T$ .

Hvis  $m$  er antall observasjoner som  $T$  baseres på ( $n_1$  eller  $n_2$ ) kan en utfra normalfordelingen finne de kritiske verdier for:

$$u = \frac{2T - m(n+1) \pm 1}{m(m+1)(n-m)/3}$$

hvor en velger tegnet foran 1 slik at  $n$  blir minst mulig.

Tabell (16.1).

## Signifikansgrensen.

Sammenlikning av to rangerings-serier.

n prøver er rangert etter en bestemt egenskap av to dommere eller to ganger av en dommer.

Kendalls rangeringskoeffisient:

Kritiske verdier for S og  $\tau$  på 5 %-nivået

$$\tau = \frac{S}{\frac{1}{2}n(n-1)}$$

n	5	6	7	8	9	10
s	10	13	15	18	20	23
$\tau$	1,00	0,87	0,71	0,64	0,56	0,51

Spearman's rangkorrelasjonskoeffisient

Kritiske verdier av  $\rho$  på 5 %-nivået.

$$\rho = 1 - \frac{\sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

n	5	6	7	8	9	10
$\rho$	0,96	0,90	0,83	0,76	0,71	0,68

Tabell 15.2.

## Rangeringstest.

Pangorden-totaler, som er nødvendig for å vise signifikante forskjeller på 5 og 1 % nivået (øverste og nederste linje i hvert tallpar).

Antall gjentak.	Antall prøver, som er sammenliknet.					
	3	4	5	6	7	8
4	5-11 -	5-15 -	6-18 5-19	6-22 5-23	7-25 5-27	7-29 6-50
5	6-14 -	7-18 6-19	8-22 7-23	9-26 7-28	9-31 8-32	10-35 8-37
6	8-16 7-17	9-21 8-22	10-26 9-27	11-31 9-33	12-36 10-38	13-41 11-43
7	10-18 8-20	11-24 10-25	12-30 11-31	14-35 12-37	15-41 13-43	17-46 14-49
8	11-21 10-22	13-27 11-29	15-33 13-35	17-39 14-42	18-46 16-48	20-52 17-55
9	13-23 12-24	15-30 13-32	17-37 15-39	19-44 17-46	21-51 19-53	24-57 21-60
10	14-26 13-27	17-33 15-35	20-40 18-42	22-48 20-50	25-55 22-53	27-63 24-66
12	18-30 16-32	21-39 19-41	24-48 22-50	28-56 25-59	31-65 28-68	34-74 31-77
14	21-35 20-36	25-45 23-47	29-55 27-57	33-65 30-68	37-75 34-78	41-85 37-89
16	24-40 23-41	29-51 27-53	34-62 32-64	39-73 36-76	43-85 40-88	48-95 44-100
18	28-44 26-46	33-57 31-59	39-69 36-72	44-82 41-85	49-95 46-98	55-107 51-111
20	31-49 29-51	37-63 35-65	43-77 41-79	49-91 46-94	55-105 52-108	61-119 58-122

MANN OG WHITNEY'S (13) test likner en del på det forrige, og det er matematisk sammenheng mellom metodene.

Her ordnes også det totale antallet observasjoner ( $n = (n_1 + n_2)$ ), og de individuelle observasjonene betegnes henholdsvis  $x_1$  og  $x_2$  etter den prøven de tilhører. For hver  $x_1$ , telles antall  $x_2$  som er mindre enn denne og antallet summeres for alle  $x_1$ . Tilsvarende gjør vi for  $x_2$ . Setter vi U som det minste av disse summene, kan vi finne kritiske verdier for denne tabeller for alle kombinasjoner av  $n_1 \leq 8$  og  $n_2 \leq 8$ .

I stedet for WILCOXON's eller MANN og WHITNEY's test kan en bruke vanlig STUDENT's t-test.

b. To eller flere stikkprøver.

KRUSKAL OG WALLI's (16) test.

Vi antar at vi har K prøver med  $n_1, \dots, n_i, \dots, n_k$  målinger av en bestemt variabel.

Det totale antall observasjoner  $n = \sum n_i$  ordnes i størrelsesorden etter målingene og målingene får rangeringsnummer 1 til n. For hver av de opprinnelige prøvene beregner vi summen av rangeringsnummerene.

Det defineres da en testvariabel H:

$$H = \frac{\frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(n+1)}{1 - \frac{\sum T}{n(n^2 - 1)}}$$

hvor:

$R_i$  = summen av rangeringsnummer for prøve nr.  $K_i$

$n_i$  = antall observasjoner av prøve nr.  $K_i$

$n = \sum n_i$

$K$  = antall prøver

$T = t(t+1)(t-1)$

$t$  = antall observasjoner i en gruppe av "tied observations"

For et stort antall prøver er  $H \chi^2$  - fordelt med  $(K-1)$  frihetsgrader, ellers finnes det tabeller over kritiske verdier for  $H$  ved små verdier av  $K$  og  $n_i$ .

## 17. SKILNADS TEST.

Denne testingstypen er særlig egnet til bruk i større konsumment-test. Det er da viktig at en dommer blir forelagt så få prøver som mulig (helst bare to) og at reglene for hvordan bedømmelsen skal gjøres er enkelt utformet.

SCHFFE (19) har utarbeidet et slikt parvis test som er enkelt, men som bygger på enkelte forutsetninger som kanskje ikke alltid er like holdbare.

Vi antar at vi har  $m$  artikler som skal bedømmes enten hedonisk eller også analytisk (f.eks. med hensyn til smaksstyrke o.l.) Det kan da lages  $M$  par av disse prøvene hvor da

$$M = \frac{1}{2}m(m-1)$$

Hvert par (la oss kalle dem  $i$  og  $j$ ) blir presentert for  $r$  dommere i rekkefølgen  $i, j$  og helst også  $r$  dommere i rekkefølgen  $j, i$ .

Prøvene presenteres alltid i kodet rekkefølge A-B (eller 1-2) og hvis det blir antatt at rekkefølgen av bedømmingen har betydning, skal prøve A bedømmes før B, og en skal ikke gå tilbake til A igjen. Vi får da en rekke bedømmelser som vi kan betegne  $x_{ij}^k$  som betyr bedømmelse i rekkefølge  $i-j$  av dommer  $k$ ,  $i$  er da alltid forskjellig fra  $j$ .

For at observasjonene skal være uavhengig fordrer teorien egentlig at en dommer bare skal bedømme et par. Dette betyr at vi må ha  $2 r M$  dommere, som kan bli et meget høyt tall. SCHEFFE hevder imidlertid at en kan la en dommer smake på flere par, men da slik at han smaker like mange ganger på hver prøve.

Dommeren skal nå vurdere prøvene A-B mot hverandre ved hjelp av en 7 delt skala hvor det settes et kryss ut for det utsagnet han mener er riktig:

- Jeg liker A mye bedre enn B (3).
- Jeg liker A bedre enn B (2).
- Jeg liker A svakt bedre enn B (1).
- Jeg synes prøvene er like (0).
- Jeg liker B svakt bedre enn A (-1).
- Jeg liker B bedre enn A (-2).
- Jeg liker B mye bedre enn A (-3).

Poengene for de enkelte svarene er satt i parentes. Det må imidlertid understrekkes at om en prøve her får minuspoeng ikke alltid betyr at prøva er av dårlig kvalitet. Poenget som en dommer gir ved sitt utsagn kan tenkes å være sammensatt av to komponenter, (a) en som representerer dommerens gjennomsnittlige smak og (b) en som representerer dommerens avvik fra sitt eget gjennomsnitt. Komponent a er en tilfeldig variabel, og med verdien  $\mu_{ij}$  vil vi da mene gjennomsnittet for denne komponenten i utvalget som dommeren er en del av. Forutsetningen om at  $x_{ijk}$  har lik varians er tvilsom, og at  $x_{ijk}$  er normalfordelt er lite sannsynlig. Det siste betyr imidlertid mindre for varians-analysen.

Preferanse for i over j betegnes  $\mu_{ij}$  og for j over i er da  $\mu_{ji}$ . Gjennomsnittet av preferansen for i over j er da:

$$\tau_{ij} = \frac{1}{2} (\mu_{ij} + \mu_{ji})$$

som estimat bruker vi:

$$\hat{\mu}_{ij} = \sum_{k=1}^r x_{ijk}/r$$

og

$$\hat{\tau}_{ij} = \frac{1}{2}(\hat{\mu}_{ij} - \hat{\mu}_{ji})$$

Nå vil som regel  $\mu_{ji}$  ikke være lik  $-\hat{\mu}_{ij}$ . Forskjellen her kan skyldes serveringsrekkefølgen og den gjennomsnittlige differensen for paret  $ij$  er da:

$$\delta_{ij} = \frac{1}{2}(\mu_{ij} + \mu_{ji})$$

estimert ved

$$\hat{\delta}_{ij} = \frac{1}{2}(\hat{\mu}_{ij} + \hat{\mu}_{ji})$$

Nå er

$$\hat{\mu}_{ji} = -\hat{\mu}_{ij} \text{ og } \delta_{ji} = \delta_{ij}$$

og forventingen for observasjonen  $x_{ijk}$ :

$$E(x_{ijk}) = \hat{\mu}_{ij} + \delta_{ij}$$

Hvis serveringsrekkefølgen er av interesse kan en beregne den gjennomsnittlige rekkefølgeeffekten for alle parene:

$$\hat{\delta} = \frac{1}{M} \sum_{i < j} \hat{\delta}_{ij}$$

Parameteren  $\hat{\delta}$  er et mål for den gjennomsnittlige fordelen for artikkell i å bli bedømt i rekkefølge  $(i,j)$  heller enn i rekkefølge  $(j,i)$  for de  $M$  parene.

Vi tenker oss nå at det finnes parametre  $\alpha_1, \alpha_2, \dots, \alpha_m$  som karakteriserer de  $m$  prøvene vi har slik at den gjennomsnittlige preferanse  $\hat{\mu}_{ij}$  er differensen mellom de tilsvarende parametrene

$$\hat{\mu}_{ij} = \alpha_i - \alpha_j$$

Denne oppdelingen forutsetter substraktivitet.

Avviket fra substraktivitet kan i det generelle tilfeldet settes lik  $\gamma_{ij}$  hvor

$$\pi_{ij} = \alpha_i - \alpha_j + \gamma_{ij}$$

$\gamma_{ij}$  kan estimeres ved

$$\hat{\gamma}_{ij} = \hat{\pi}_{ij} - \hat{\alpha}_i + \hat{\alpha}_j$$

$\alpha_i$  er nå prøve i's gjennomsnittlige preferanse over de øvrige prøvene, og kan estimeres ved

$$\hat{\alpha}_i = \frac{(\sum \pi_{ij})}{m}$$

og

$$\sum_{i=1}^m \hat{\alpha}_i = 0$$

### Sammendrag av formlene.

$$\hat{\mu}_{ij} = \frac{r}{\sum_{k=1}^r x_{ijk}}$$

$$\hat{\pi}_{ij} = \frac{1}{2}(\hat{\mu}_{ij} - \hat{\mu}_{ji})$$

$$\hat{\delta}_{ij} = \frac{1}{2}(\hat{\mu}_{ij} + \hat{\mu}_{ji})$$

$$\hat{\alpha}_i = \frac{\sum_{j=1}^m \hat{\alpha}_{ij}}{m}$$

$$\hat{\gamma}_{ij} = \hat{\alpha}_{ij} - \hat{\alpha}_i + \hat{\alpha}_j$$

Følgende kvadratsummer beregnes for variansanalyse:

$$\text{Total: } S_t = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^r (\hat{x}_{ijk})^2$$

$$\text{Feil: } S_e = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^r (\hat{x}_{ijk} - \hat{\mu}_{ij})^2$$

Gjennomsnittseffekten:

$$S_\mu = r \sum_{i=1}^m \sum_{j=1}^m (\hat{\mu}_{ij})^2$$

Rekkefølge-effekt:

$$S_\delta = 2r \sum_{i < j} (\hat{\delta}_{ij})^2$$

Gjennomsnittlig preferanse:

$$S_\eta = 2r \sum_{i < j} (\hat{\eta}_{ij})^2$$

Avvik fra substraktivitet:

$$S_\gamma = 2r \sum_{i < j} (\hat{\gamma}_{ij})^2$$

Hovedeffekt:

$$S_\alpha = 2rm \sum_{i=1}^m (\hat{\alpha}_i)^2$$

## 18. GRADERING ELLER SCORINGSTEST.

Dette er meget vanskelige prøver i det dommerne her skal sette poeng, f.eks. etter hvordan markedet oppfatter den prøven de bedømmer. Ofte kan det være nødvendig at dommeren prøver å undertrykke sin egen oppfatning av prøven. Selv om f.eks. de vanlige bedømmelsene av meieriproducter skal være analytiske, skjer det nok ofte at poengene blir satt hedonisk ): hvor godt dommeren selv liker prøven.

Ved graderingstest trenger vi nødvendigvis en skala for klassifiser. Vanligvis skiller vi mellom fire typer:

1. Nominalskala uten sammenheng: 1 = ost, 2 = smør, 3 = flesk..
2. Ordinalskala eller rangeringsskala gir oss bare mulighet til å si om en vare er sòtere enn en annen, bedre enn en annen, etc. Avstanden mellom prøvene 1 ..... n vet vi imidlertid ikke noe om.
3. I intervallskala er avstanden mellom to elementer lik, men skalaen har intet nullpunkt og ingen enhet. Vi kan derfor bare regne med differenser.
4. En kvoteskala har både nullpunkt og enhet og alle regneoperasjoner kan utføres med elementene.

Ved våre vanlige bedømmelser hvor skalaen rekker fra 1-5 regner vi med å ha en kvoteskala. Helt riktig er dette ikke fordi skalaen er begrenset mens kvaliteten av produktet neppe er det, slik at hvert tall egentlig gir uttrykk for en kvalitetsgruppe. Det kan også være tvilsomt om differensene mellom klassene er like store. Ved bedømmelse av meieriproducter blir karakter 3 gitt for prøver som dommeren anser for å kunne selges, men som har en eller annen feil som skal anmerkes. 4 gis for en god vare og 5 for et utmerket produkt. Det er trolig vanskeligere å skjelne mellom 4 og 5 enn mellom 3 og 4, noe som kanskje virker til å skape lengre intervall mellom 3 og 4 enn mellom 4 og 5. Nå vil noen dommere på den andre siden sjeldent bruke karakter 5 fordi de mener den skal brukes til det allér ypperste. Dette skulle gjøre intervallet mellom 4 og 5 bredere. Denne begrensning i bruk av skalaer er

kanskje vanligere i den andre enden av skalaen hvor en nødig bruker laveste karakter. En annen skjevhet i bruk av skalaområdet kan oppstå når poengene brukes ved prisfastsettelsen av produktet til produsent.

En dommer vil muligens være noe forsiktigere med å bruke 2 i stedet for 3 når han vet at det får betydelige konsekvenser for produsentene.

En stor fordel ved et graderingstest sammenliknet med de testene som er behandlet tidligere, er at en kan behandle langt flere prøver. Ved store forsøk, f.eks. finnes det ingen andre metoder som er brukbare uten store omkostninger. De verdiene vi får ved bedømmelsene kan også lett tilpasses i våre vanlige test-metoder som t-test eller variansanalyse.

Det vanlige resultatet en vil få ut av et t-test eller en variansanalyse kan gi svar på om det innen hele serien er forskjell eller ikke. Vi er imidlertid interessert i mer enn dét, fordi vi gjerne vil vite om noen av de enkelte prøvene vi bedømmer er bedre eller dårligere en noen av de andre.

Hvis bedømmelsen er en del av et forsøk, vil det være høyst ønskelig om en på forhånd velger ut de kontrastene en vil teste. Det må presiseres at kontrapastene må settes opp på forhånd og ikke etter at en har sett resultatene.

Som et eksempel kan vi ta en del av et forsøk med iskrem hvor en er interessert i å teste hvorvidt en råvare vi bruker, A, kan erstattes med billigere råvarer, B, C, D. Av disse vet vi at B er den som er suverent billigst, mens C er lettest å ha med å gjøre i produksjonen. Det vil da være naturlig å teste A mot B, C og D, B mot C og D, og C mot D. Disse sammenlikningene må vi sette opp på forhånd.

Et slikt forsøk vil kreve flere gjentak, men for enkelhets skyld tenker vi oss at en prøve av hver iskrembatch er trukket tilfeldig ut og bedømt av fem dommere.

## Iskremprøve:

Dommer	A	B	C	D	
1	3,5	2,5	4,0	3,5	13,5
2	4,0	2,0	4,0	4,0	14,0
3	3,5	3,0	4,5	3,0	14,0
4	4,0	2,0	2,5	3,0	11,5
5	4,0	2,5	4,0	3,0	13,5
Sum	19,0	12,0	19,0	16,5	66,5
$\bar{x}$	3,8	2,4	3,8	3,3	

På vanlig måte finner vi:

Variasjonsårsak	f.f.	kv.s.	m.k.	F
Dommere	4	1,0750		
Prøver	3	6,5375	2.1792	8,65
Feil	12	3.0250	0,2520	

Vi regner først ut den første kontrasten:

$$q_a = -3(19) + 1(12) + 1(19) + 1(16,5) = -9,5 \quad \sum C_{ia}^2 = 12$$

$$\sum C_{ia} = 0$$

$$q_b = 0(19) + 2(12) - 1(19) - 1(16,5) = -11,5 \quad \sum C_{ib}^2 = 6$$

$$\sum C_{ib} = 0$$

$$q_c = 0(19) + 0(12) + 1(19) - 1(16,5) = 2,5 \quad \sum C_{ic}^2 = 2$$

$$\sum C_{ic} = 0$$

(for ortogonalitet må  $\sum C_{ia} C_{ib} = \sum C_{ib} C_{ic} = \sum C_{ia} C_{ic} = 0$ )

$$F_a = (-9,5)^2 / [12(5)(0,2520)] = 5,97^*$$

$$F_b = (-11,5)^2 / [6(5)(0,2520)] = 17,49^{***}$$

$$F_c = (2,5)^2 / [2(5)(0,2520)] = 2,48$$

med 1 og 12 frihetsgrader.

Det kan her konkluderes med at dommerne har funnet forskjell i iskremkvaliteten når vi bruker vår ordinære råvare sammenliknet med de alternative råvarene, og i favør av den ordinære. Det er også klart at dommerne har ment at B var dårligere enn C og D. Derimot er det ikke påvist forskjell mellom C og D.

I dette forsøket viste vi ikke på forhånd at kontrastene A mot D var den mest interessante. For et nytt forsøk ville det være all grunn til å sette opp denne i stedet for C mot D ! (Men dette må altså gjøres på forhånd).

Hvis vi imidlertid ikke har noe grunnlag for å sette opp kontraster på forhånd, noe som selvfølgelig er vanskelig ved ordinære bedømmelser, må vi ta utgangspunktet i de verdiene vi finner. Dette har DUNCAN (9) gjort på følgende måte:

For eksemplet vårt kan det beregnes en standarfeil:

$$S_m = \sqrt{0,2520/5} = 0,2245$$

Utfra spesielle tabeller (tabell 18.1 avleses så faktorer etter antall sammenlikninger som skal gjøres, antall frihetsgrader for feilledet og signifikansgrense.

Disse faktorene multipliseres med  $S_m$ : og gir det DUNCAN kalte for den korteste signifikante avstand (Rp). Analysen kan nå føres opp slik:

Avlest faktor ( $\alpha = 0,05$ )	3,08	3,23	3,33
Sammenligninger	2	3	4
Rp	0,691	0,725	0,748

Resultat:

Prøve	B	D	C	A
Gj.snitt	2,40	3,30	3,60	3,80

Tabel 18.2. Studentisert variasjonsvidde q =  $(x_n - x_1)/s_v$ 

n v	ε = 0,05										ε = 0,005									
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	13,6	27,0	32,8	37,1	40,4	43,1	45,4	47,4	49,1	50,6	52,0	53,2	54,3	55,4	56,3	57,2	58,0	58,8	59,6	
2	6,09	8,3	9,8	10,9	11,7	12,4	13,0	13,5	14,0	14,4	14,7	15,1	15,4	15,7	15,9	16,1	16,4	16,6	16,8	
3	4,56	5,91	6,82	7,50	8,04	8,48	8,85	9,18	9,46	9,72	9,95	10,15	10,35	10,52	10,69	10,84	10,93	11,11	11,24	
4	3,93	5,04	5,76	6,29	6,71	7,05	7,35	7,60	7,83	8,03	8,21	8,37	8,52	8,66	8,79	8,91	9,03	9,13	9,23	
5	3,64	4,60	5,22	5,67	6,03	6,33	6,58	6,80	6,99	7,17	7,32	7,47	7,60	7,72	7,83	7,93	8,03	8,12	8,21	
6	3,46	4,34	4,90	5,31	5,63	5,89	6,12	6,32	6,49	6,65	6,79	6,92	7,03	7,14	7,24	7,34	7,43	7,51	7,59	
7	3,34	4,16	4,68	5,06	5,36	5,61	5,82	6,00	6,16	6,30	6,43	6,55	6,66	6,76	6,85	6,94	7,02	7,09	7,17	
8	3,26	4,04	4,53	4,89	5,17	5,40	5,60	5,77	5,92	6,05	6,18	6,29	6,39	6,48	6,57	6,65	6,73	6,80	6,87	
9	3,20	3,95	4,42	4,76	5,02	5,24	5,43	5,60	5,74	5,87	5,98	6,09	6,19	6,28	6,36	6,44	6,51	6,58	6,64	
10	3,15	3,88	4,33	4,65	4,91	5,12	5,30	5,46	5,60	5,72	5,83	5,95	6,03	6,11	6,20	6,27	6,34	6,40	6,47	
11	3,11	3,82	4,26	4,57	4,82	5,03	5,20	5,35	5,49	5,61	5,71	5,81	5,90	5,99	6,06	6,14	6,20	6,26	6,35	
12	3,08	3,77	4,20	4,51	4,75	4,95	5,12	5,27	5,40	5,51	5,62	5,71	5,80	5,88	5,95	6,03	6,09	6,15	6,21	
13	3,06	3,73	4,15	4,45	4,68	4,88	5,05	5,18	5,32	5,43	5,55	5,63	5,71	5,79	5,86	5,93	6,00	6,05	6,11	
14	3,05	3,70	4,11	4,41	4,64	4,85	4,99	5,13	5,25	5,36	5,46	5,55	5,64	5,72	5,79	5,85	5,92	5,97	6,03	
15	3,01	3,67	4,08	4,37	4,60	4,78	4,94	5,08	5,20	5,31	5,40	5,49	5,58	5,65	5,72	5,79	5,85	5,90	5,96	
16	3,00	3,65	4,05	4,35	4,56	4,74	4,90	5,03	5,15	5,26	5,35	5,44	5,52	5,59	5,66	5,72	5,79	5,84	5,90	
17	2,98	3,63	4,02	4,30	4,52	4,71	4,86	4,99	5,11	5,21	5,31	5,39	5,47	5,55	5,61	5,68	5,74	5,79	5,84	
18	2,97	3,61	4,00	4,28	4,49	4,67	4,82	4,96	5,07	5,17	5,27	5,35	5,43	5,50	5,57	5,63	5,69	5,74	5,79	
19	2,96	3,59	3,98	4,25	4,47	4,65	4,79	4,92	5,04	5,14	5,23	5,32	5,39	5,46	5,53	5,59	5,65	5,70	5,75	
20	2,95	3,58	3,96	4,23	4,45	4,62	4,77	4,90	5,01	5,11	5,20	5,28	5,36	5,43	5,49	5,55	5,61	5,66	5,71	
24	2,92	3,53	3,99	4,17	4,37	4,54	4,68	4,81	4,92	5,01	5,10	5,18	5,25	5,32	5,38	5,44	5,50	5,55	5,59	
30	2,89	3,49	3,84	4,10	4,30	4,46	4,60	4,72	4,83	4,92	5,00	5,08	5,15	5,21	5,27	5,33	5,38	5,45	5,48	
40	2,86	3,44	3,79	4,04	4,25	4,59	4,52	4,63	4,74	4,82	4,91	4,98	5,05	5,11	5,16	5,22	5,27	5,31	5,36	
60	2,85	3,40	3,74	4,16	4,31	4,44	4,55	4,65	4,75	4,81	4,88	4,94	5,00	5,06	5,11	5,16	5,20	5,24	5,28	
120	2,80	3,36	3,69	4,10	4,24	4,36	4,48	4,56	4,64	4,72	4,78	4,84	4,90	4,95	5,00	5,05	5,09	5,13	5,16	
240	2,76	3,21	3,52	4,02	4,17	4,30	4,42	4,53	4,63	4,70	4,78	4,85	4,92	4,98	5,05	5,11	5,16	5,20	5,24	

Vi sammenlikner først AB. Differensen i gjennomsnitt er 1,40 som er større enn 0,748. Altså er forskjellen signifikant. Vi sammenligner så AD. Differensen er her 0,50 som er mindre enn 0,725 og differensen er ikke signifikant. Ikke signifikante effekter understrekkes. Så sammenlignes DB hvor forskjellen er 0,90 som er større enn 0,691 og altså signifikant. Analysen blir derfor:  
Ingen signifikant forskjell mellom A, C og D.  
B er signifikant dårligere enn A, C og D.

Nå er det kommet alvorlige innvendinger mot Duncan's test. (BOARDMAN og NOFFITT) (4) har vist at feil av type I øker med omtrent 4 % for hver sammenlikning utover to ved en valgt 5 % signifikans-grense. GILL (13) anbefaler heller Tukey's Honestly Significant Difference test (HSD).

For eksemplet vårt kan vi da teste alle parvise kontraster:

$$A - B: (3,8 - 2,4)^+ q_{05 \ 4,12} \sqrt{0,2520/5} = 1,4 \pm 0,95$$

hvor  $q_{05 \ 4,12}$  finnes i tabeller (18.2) over studentisert variasjonsvidde.

$$A - D : 0,5 \pm 0,95$$

$$B - C : 1,4 \pm 0,95$$

$$B - D : 0,9 \pm 0,95$$

$$C - D : 0,5 \pm 0,95$$

Det er etter dette bare påviselig forskjell mellom A og B og mellom B og C. Vi ser at denne metoden som ventet, er striktere enn DUNCAN's metode, i det den ikke gir forskjell mellom B og D.

#### 19. PROFILERING.

Smaks-profilering er en kvalitativ og kvantitativ metode for å beskrive smaken i et bestemt produkt. Metoden ble utviklet av CAIRNCROSS og SJØSTRØM (4) i 1950 ved firmaet Arthur D. Little i U.S.A.

Tabel 18.1) Signifikante studentiserte differenser på 5 % nivået for DUNCAN's test for flere gjennomsnitt.

Antall gjennomsnitt som sammenliknes

Reinleddets frihetssgrad	2	3	4	5	6	7	8	9	10	11.2	12	14	16	18	20	50	100
1	18,00	18,00	18,00	18,00	18,00	18,00	18,00	18,00	18,00	18,00	18,00	18,00	18,00	18,00	18,00	18,00	18,00
2	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09
3	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50
4	3,93	4,01	4,02	4,02	4,02	4,02	4,02	4,02	4,02	4,02	4,02	4,02	4,02	4,02	4,02	4,02	4,02
5	3,64	3,74	3,79	3,83	3,83	3,83	3,83	3,83	3,83	3,83	3,83	3,83	3,83	3,83	3,83	3,83	3,83
6	3,46	3,58	3,64	3,68	3,68	3,68	3,68	3,68	3,68	3,68	3,68	3,68	3,68	3,68	3,68	3,68	3,68
7	3,35	3,47	3,54	3,58	3,60	3,61	3,61	3,61	3,61	3,61	3,61	3,61	3,61	3,61	3,61	3,61	3,61
8	3,26	3,39	3,47	3,52	3,55	3,56	3,56	3,56	3,56	3,56	3,56	3,56	3,56	3,56	3,56	3,56	3,56
9	3,20	3,34	3,41	3,47	3,50	3,52	3,52	3,52	3,52	3,52	3,52	3,52	3,52	3,52	3,52	3,52	3,52
10	3,15	3,30	3,37	3,43	3,46	3,47	3,47	3,47	3,47	3,47	3,47	3,47	3,47	3,47	3,47	3,47	3,47
11	3,12	3,27	3,35	3,39	3,43	3,44	3,45	3,46	3,46	3,46	3,46	3,46	3,46	3,46	3,46	3,46	3,46
12	3,08	3,23	3,33	3,36	3,40	3,42	3,44	3,44	3,44	3,44	3,46	3,46	3,46	3,46	3,46	3,46	3,46
13	3,06	3,21	3,30	3,35	3,38	3,41	3,42	3,44	3,44	3,45	3,46	3,46	3,46	3,46	3,46	3,46	3,46
14	3,03	3,18	3,27	3,33	3,37	3,37	3,39	3,41	3,42	3,42	3,44	3,45	3,46	3,46	3,46	3,46	3,46
15	3,01	3,16	3,25	3,31	3,36	3,38	3,40	3,42	3,42	3,43	3,44	3,45	3,45	3,45	3,45	3,45	3,45
16	3,00	3,15	3,23	3,30	3,34	3,37	3,39	3,41	3,41	3,43	3,44	3,45	3,45	3,45	3,45	3,45	3,45
17	2,98	3,13	3,22	3,28	3,33	3,36	3,38	3,40	3,40	3,42	3,44	3,45	3,45	3,45	3,45	3,45	3,45
18	2,97	3,12	3,21	3,24	3,27	3,32	3,35	3,37	3,39	3,41	3,43	3,45	3,45	3,45	3,45	3,45	3,45
19	2,96	3,11	3,19	3,21	3,26	3,31	3,35	3,37	3,39	3,41	3,44	3,44	3,44	3,44	3,44	3,44	3,44
20	2,95	3,10	3,18	3,25	3,30	3,34	3,36	3,38	3,38	3,40	3,43	3,44	3,44	3,44	3,44	3,44	3,44
22	2,93	3,08	3,17	3,24	3,29	3,32	3,35	3,37	3,39	3,42	3,44	3,45	3,45	3,45	3,45	3,45	3,45
24	2,92	3,07	3,15	3,22	3,28	3,31	3,34	3,37	3,38	3,41	3,44	3,45	3,45	3,45	3,45	3,45	3,45
26	2,91	3,06	3,14	3,21	3,27	3,30	3,34	3,36	3,38	3,41	3,43	3,45	3,45	3,45	3,45	3,45	3,45
28	2,90	3,04	3,13	3,20	3,26	3,30	3,33	3,35	3,37	3,40	3,43	3,45	3,45	3,45	3,45	3,45	3,45
50	2,89	3,04	3,12	3,20	3,25	3,29	3,32	3,35	3,37	3,40	3,43	3,44	3,44	3,44	3,44	3,44	3,44
40	2,86	3,01	3,10	3,17	3,22	3,27	3,30	3,33	3,35	3,39	3,42	3,44	3,44	3,44	3,44	3,44	3,44
60	2,83	2,98	3,08	3,14	3,20	3,24	3,28	3,31	3,33	3,37	3,40	3,43	3,45	3,45	3,45	3,45	3,45
100	2,80	2,95	3,05	3,12	3,18	3,22	3,26	3,29	3,32	3,36	3,40	3,42	3,45	3,45	3,45	3,45	3,45
24	2,77	2,97	3,02	3,15	3,19	3,23	3,27	3,30	3,33	3,37	3,40	3,43	3,46	3,46	3,46	3,46	3,46

## §2

Ved smaksprofileringen skal dommerne

1. bestemme de enkelte komponentene som kan bli luktet eller smakt,
2. bestemme intensiteter til de enkelte smaks- eller luktsfornemmelser,
3. bestemme rekkefølgen som de enkelte smaks- og luktestoffene blir observert i,
4. bestemme ettersmak og
5. gi uttrykk for helhetsinntrykket av varen.

Profilering krever en meget god dommerstab hvor det under treningen særlig legges vekt på evnen til å gjenkjenne forskjellige lukter og smaker. Et godt kjennskap til kjemi og til næringsmiddelteknologi er vesentlig. Et panel består derfor av godt trenede spesialister. Det har da også vist seg at det samme panelet har gitt et og samme produkt den samme profilen med års mellomrom, og at også mange forskjellige panel har kommet til samme profileringen av et næringsmiddel uavhengig av hverandre.

1. De enkelte komponentene i en prøve (næringsmiddel, kosemetikk, etc. blir karakterisert med beskrivende ord f.eks.: Eggsmak, kålsmak, stinkende o.s.v., men helst skal en bruke referanse til definerte kjemikalier eller kjente stoff når det er mulig. For det enkelte produkt blir det utarbeidet en nomenklatur for lukter, smaker og konsistensformer som er vanlige.
2. Intensiteten for hver komponent noteres vanligvis ved 5 graderinger.

0	= ikke tilstede
X	= så vidt merkbar
1 eller +	= svak
2 eller ++	= moderat
3 eller +++	= sterkt

Dommerne må trenes opp sammen slik at de mener det samme med sine anmerkninger. Et team med stor erfaring kan i helt spesielle tilfeller få lov til å bruke "halve karakterer".

3. Rekkefølgen av komponentene har stor betydning, idet en mindre god smak med en gang en smaker på et produkt eller en mindre god ettersmak betyr mer for smaken enn om den mindre gode smaken kommer mer sammen med de andre smakskomponentene. Det er f.eks. funnet at natriumglutamat kan virke inn på smaken til enkelte produkter fordi det forskyver rekkefølgen av smakskomponentene på en gunstig måte. Ved profilering angis rekkefølgen ganske enkelt ved at en skriver ned de enkelte smaks- og luktkomponentene i kronologisk rekkefølge.
4. Ettersmak eller etterlukt kan være meget karakteristisk for et produkt, som regel i ugunstig retning. En anmerking om dette går da oftest inn i en profilering.
5. Amplituden eller helhetsinntrykket av produktet er av stor viktighet. Ofte noteres en karakter for denne før en går løs på å merke seg de enkelte komponentene og deres intensiteter. Det nyttes vanligvis en firetrinns skala:
- 0 = meget lav
  - 1 = lav
  - 2 = middels
  - 3 = meget høy

Skalaen er ikke fiksert slik at et produkt som i og for seg er godt, kan få tegnet 0. La oss som eksempel ta pisket krem. Uten tilsetning kan den f.eks. få 0 til 1, med sukker kanskje 2 og med sukker og vanilje kanskje 3. Hvis vi imidlertid skal bestemme profilen for rein krem f.eks. med forskjellig fettinnhold kan kanskje den kremen som fikk 0 til 1 i den første forbindelsen, få 3 nå. Det er derfor viktig å etablere ei ramme for de enkelte produktene. I det første eksemplet vil da rammen omfatte krem med tilsetninger, i det andre rein krem.

#### Utdannelse for panel-medlemmer.

Vi har tidligere sett på hvilke krav som stilles til en dommer med hensyn til terskelverdier etc. Utgangspunktet må være at kandidatene til et panel har en normal evne til å oppfatte smak og lukt. Det er viktig at medlemmene sier ifra hvis de av en eller annen grunn er indisponert.

Ved Arthur D. Little laboratoriet er treningen bygd opp omrent slik for å utdanne paneler for industrien:

1. Kurs som gir den nødvendige bakgrunnen for profil teknikk, fysiske og psykiske aspekter ved smaking og lukting, teorier om fornemmelsen av smak og lukt, og sammenheng mellom kjemisk oppbygging av et stoff og dets lukt og smak. Dette kurset inneholder også demonstrasjon av øvede paneler.
2. Øvelser med diskusjoner.
3. Samarbeid med øvede paneler. Utdannelsen tar fra 6 mnd. til et år. Elevene på kursene er pekt ut av sine respektive arbeidsgivere, og ADL har til oppgave å gi oppdragsgiveren et brukbart panel. Øvelsene foregår derfor både ved laboratoriene til Arthur D. Little og i de laboratoriene hvor elevene er ansatt.

Panel-lederen har en meget viktig oppgave i profil-panelet. Denne personen skal organisere, veilede og dirigere panelet. Videre er han ansvarlig for prepareringen av prøvene, standardisering av bedømmelsesteknikken, og er leder for gruppas diskusjoner.

Det er vesentlig at han under diskusjonen ikke skal pådytte gruppen sine meninger, men i stedet få medlemmene av gruppa til å forstå viktigheten av hver enkelts observasjon.

Den egentlige profilering blir utført som en kombinert uavhengig observasjon fra hver enkelt i gruppa og en rundebordekonferanse.

Det er viktig at gruppelederen informerer medlemmene om hensikten med bedømmelsen og med eventuell fremgang i produksjonen hvis det er et nytt produkt det gjelder. Etter første diskusjon er det viktig at gruppemedlemmene individuelt studerer det produktet eller d produktgruppen de skal bedømme, med gjentatte diskusjoner slik at alle får det beste kjennskap til produktet.

I disse diskusjonene skal det da trekkes opp retningslinjer for bedømmelsen.

Det skal da først og fremst oppnås enighet om den nødvendige terminologien, utforming av skjema, hvilke faktorer som en særlig skal koncentrere seg om og hvilken ramme som skal settes

for produktets amplityde. Når det er oppnådd enighet på disse punktene, er panelet klar til den mer formelle delen av bedømmelsen. I den første formelle bedømmelsen profilerer medlemmene hver for seg de aktuelle prøvene. Ved diskusjonen etterpå blir det spesielt sett på hvor de enkelte medlemmene avviket fra hverandre. Avvikene kan f.eks. skyldes forskjell i prøvene. F.eks. vil en profilering av midtpartiet i ei steik kunne være forskjellig fra en profilering av ytterkantene. Noen ganger kan avvikene skyldes at en av medlemmene er indisponert. Hvis forskjellene er store, bør medlemmene igjen foreta individuell profilering med følgende diskusjon. Under diskusjonen settes det da opp en foreløpig profil.

Medlemmene av gruppa foretar så den endelige bedømmelsen. Denne følges av en diskusjon som fører fram til den endelige profil.

En profilering har vært sammenliknet med fremføringen av et teaterstykke. Først får aktørene tildelt sine roller og har leseprøve. Så skal rollene innstuderes med leseprøver innimellom, hvor instruktøren i diskusjon med aktørene skal legge vekt på det vesentlige i stykket for best mulig å få fram innholdet. Når rollene er grunndig innstudert holdes generalprøve og så endelig premieren (den endelige profilering).

Presentasjonen kan enten skje ved et skjema som i tabell (19.1) eller som figur 18, hvor intensiteten er angitt ved lengden på linjen og tidspunktet som på ei urskive. Seinere kan en gå mer i detalj som i figur 19, hvor bitter smak øl er bedømt i forhold til generell ølsmak. Ølprøven i det øverste profilet viser at dette er et problemøl som det må gjøres noe med !

Tabell 19.1. Eksempel på sammenstatt profil for en øltype (5).

Amplityde = 3

Aroma	Intensitet
Humle	2
Frukt	2
Sur	1,5
Gjær	X
Malt	1
Honning (fenyledikksyre)	1

Amplityde = 1.

Smak	Intensitet.
CO <sub>2</sub> -brusing	høy
Salt	1
Søt	1
Sur	2
Frukt (vin)	1
Bitter (metall)	3
Malt	X
Gjær	1

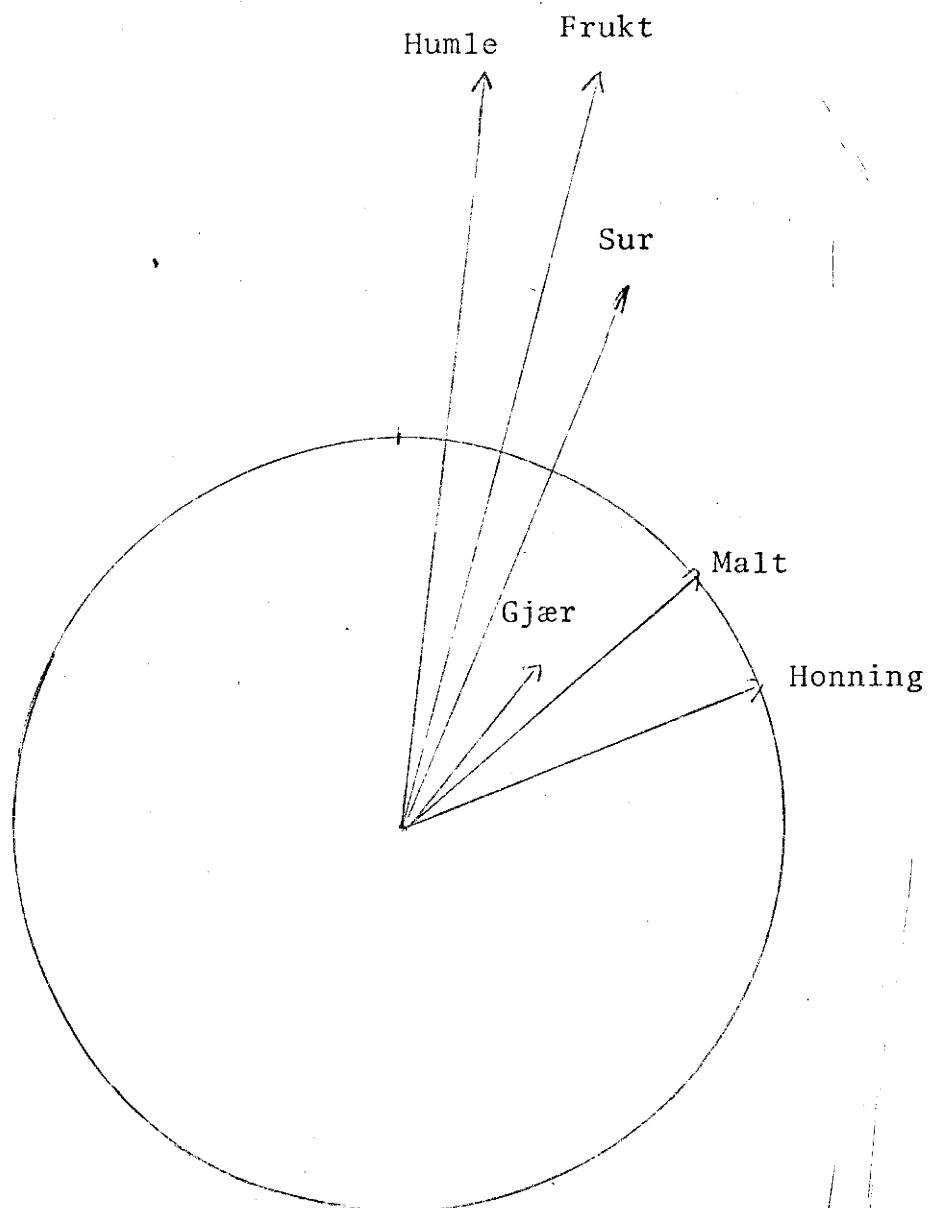
Andre: Snerpende.

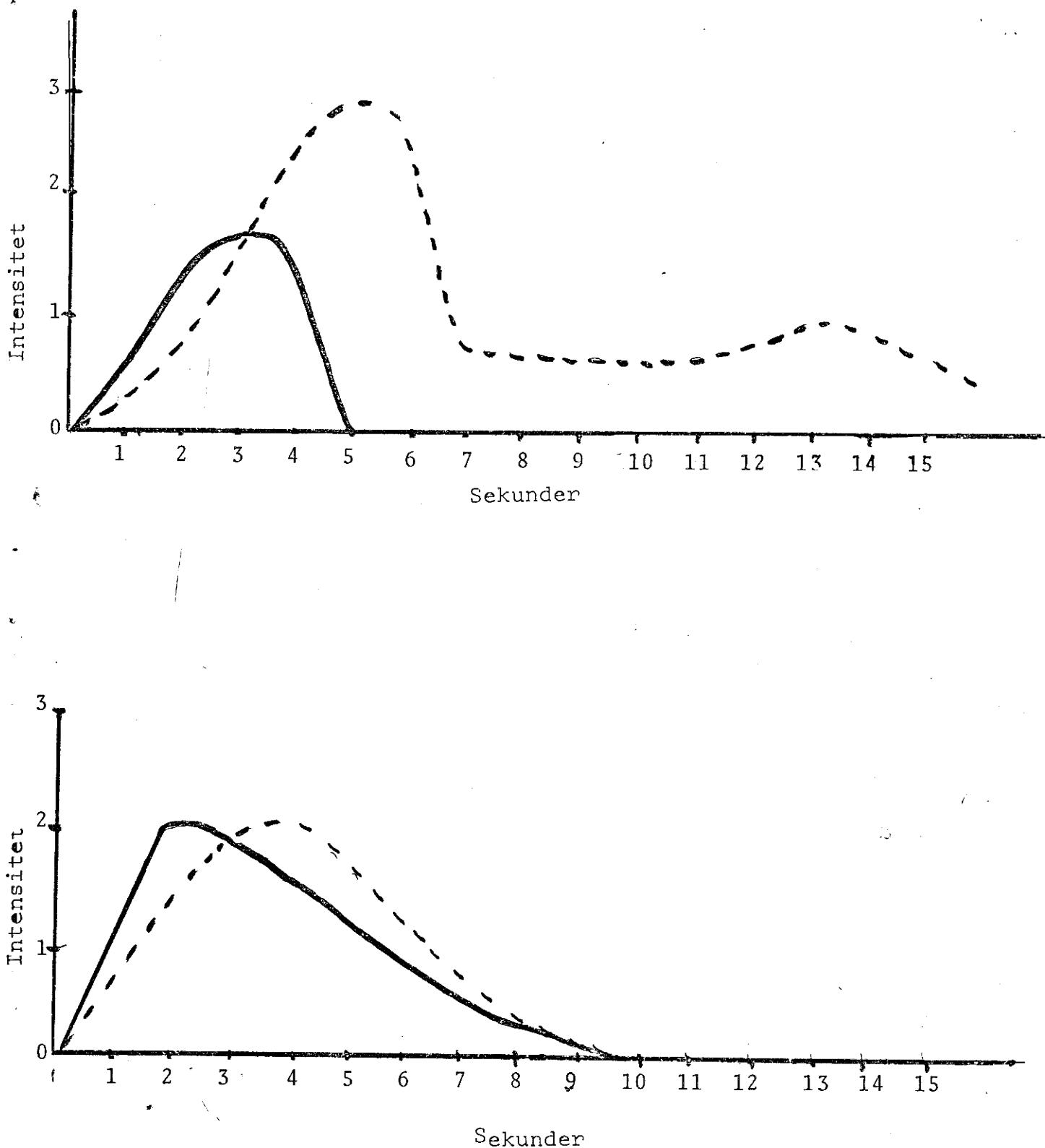
Ettersmak: Bitter,

snerpende og

tørr.

Figur 18. Aromaprofil av øl.





Figur 19. Bitter smak (---) og vanlig ølsmak (---

## 20. LITTERATUR.

1. Amoore, J. E., 1964: Die stereochemische Theorie des Geruchs. Umsch. Fortschr. Naturwiss. Med. Techn. 64:600.
2. Békésy , G. von, 1964: Sweetness produced electrically on the tongue and its relation to taste theories. J. Appl. Physiol. 19: 1105.
3. Békésy, G. von, 1966: Taste theories and the chemical stimulation of single papillae. J. Appl. Physiol. 21: 1.
4. Boardman, T.J. og Moffitt, D.R. 1971: Graphical monte carlo type I error rates for multiple comparison procedures. Biometrics 27: 738.
5. Cairncross, S.E. og Sjøstrøm, L.B. 1950: Flavor profiles-a new approach to flavor problems. Food Techn. 4: 308.
6. Caul, J.F. 1957: The profile method of flavor analysis. Advances in Food Research. 7:5.
7. Dastoli, F.R. og Price, S., 1966: Sweet- sensitive protein from bovine taste buds: isolation and assay. Science 154: 905.
8. Davies, J.T. 1953: Odor and morphology of molecules. Industrie Parfum. 8: 74.
9. Ducan, D.B. 1955: Multiple range and multiple F tests. Biometrics 11 : 1.
10. Døving, K. B. 1966: Analysis of odour similarities from electrophysiological data. Acta. Physiol. Scand. 68: 404.
11. Elsberg, C.A. og Levy, I. 1935: The sense of smell. I. A new and simple method of quantitative olfactometry. Bull. Neurol. Inst. N.Y. 4(1):5.
12. Elsberg, C.A., Levy, I. og Brewer, E.D. 1935: The sense of smell. II. A new principle of the classification of odors based upon their olfactory coefficients. Bull. Neurol. Inst. N. Y. 4(1): 20.

13. Gill, J. L., 1973: Current status of multiple comparisons of means in designed experiments. *J. of Dairy Sci.* 56(8): 973-7.
14. Kendall, M.G. 1948: Rank correlation methods. Charles Griffin & Co. ltd. London.
15. Kionka, H. og Strätz, E. 1922: Setzt der Geschmack eines Salzes sich zusammen aus dem Geschmack der einzelnen Ionen oder schmeckt man jedes Salz als Gesamt molekül? *Arch. exp. Path. Pharmacol.* 95: 241.
16. Kruskal, W. H. og Wallis, W. A. 1952: Use of ranks in one-criterion variance analysis. *J. Am. Stat. Ass.* 47(260): 583.
17. Mann, H. B. og Whitney, D.R. 1947: On a test whether one of two random variables is stochastically larger than the other. *Am. Math.* 18: 1.
18. Moncrieff, R. W. 1949: A new theory of odor. *Perfum. Ess. Oil. Rec.* 40: 279.
19. Ottoson, D. 1956: Analysis of the electrical activity of the olfactory epithelium. *Acta. Physiol. Scand.* 35, Suppl. 122: 1.
20. Ottoson, D. 1958: Studies on the relationship between olfactory stimulating effectiveness and physico-chemical properties of odour compounds. *Acta. Physiol. Scand.* 43: 167.
21. Ottoson, D. og Shepherd, G.M. 1967: Experiments and concepts in olfactory physiology. *Progr. Brain. Res.* 23: 83.
22. Randenbrock, R. E. 1968: Molecular theory of odor. *Nature.* 219: 503.
23. Scheffé, H. 1952: An analysis of variance for paired comparisons. *J. Am. Stat. Ass.* 47(259): 381.

24. Shallenberger, R. S. og Acree, T. E. 1967: Molecular theory of sweet taste. *Nature* 216: 480.
25. Sherman, P. 1970: Industrial rheology. Academic Press, London og N.Y.
26. Spearman, C. 1904: The proof and measurement of association between two things. *Am. Jour. Psych.* 15: 88.
27. Srinivas, M. 1955: Has the ear a role in registering flavour ? *Bull. Cent. Food. Techn. Res. Inst. Mysore.* 4: 136.
28. Szczesniak, A. S. 1963: Classification of textural characteristics. *J. Food Sci.* 28: 385.
29. Wald, A. 1947: Sequential analysis. John Wiley and Sons, N.Y.
30. Wilcoxon, F. 1947: Probability tables for individual comparisons by ranking methods. *Biometrics* 3(1): 119.
31. Wright, W. D. 1958: The measurement of colour. 2<sup>nd</sup> ed. Hilger & Watts Ltd. , London.