

Det skulle derfor synes klart at det ville være en stor fordel om en kunne erstatte det subjektive skjønn med mer *objektive* bestemmelser av de evner og personlighetstrekk som det i strafferettspleien blir tatt særlig hensyn til.

Her er det spørsmålet om bruken av tests kommer inn. I hvilken utstrekning kan vi overhode „måle“ evner og egenskaper? Og hvilken verdi har i virkeligheten den testing som alt er tatt i bruk i vår rettspleie?

De fleste vet sikkert hva en test er. Men det kan kanskje allikevel være nyttig å minne om det vesentlige og forsøke å presisere begrepet.

Ved vanlig eksperimentell forskning går en frem på den måten, at en systematisk *varierer* de forskjellige forsøksbetingelser for å undersøke, hvordan de forskjellige faktorer virker. Ved testing gjør vi det på en måte omvendt. Vi lager en bestemt situasjon eller en bestemt oppgave. Så holder vi alle forsøksbetingelser mest mulig konstante og undersøker hvordan *forskjellige* mennesker reagerer overfor *samme* situasjon eller *samme* oppgave. Jeg vil f. eks. undersøke evnen til å oppfatte og gjengi litt mer kompliserte inntrykk — det er jo noe som bl. a. har betydning når en skal vurdere vitneutsagn. Jeg lager da en såkalt utsagnstest: Jeg lar en stor mengde forsøkspersoner se *samme* bilde i *samme* avstand i løpet av *samme* tid. Så undersøker jeg hva de forskjellige forsøkspersoner kan gjengi riktig spontant etter en bestemt tid, og underkaster dem senere også et forhør ved hjelp av bestemte, på forhånd nøye fastsatte spørsmål. Det viser seg da at noen klarer å gjengi mange enkeltheder riktig, andre ferre, noen gir mange positivt gale svar, andre få. Har jeg gjort forsøket med tilstrekkelig mange, kan jeg regne ut gjennomsnittsverdier og karakterisere den enkelte i forhold til gjennomsnittet som et så og så godt eller et så og så dårlig vitne.

Det er naturligvis bestemte metodiske krav en må stille både til konstruksjonen av tests, standardiseringen og selve testingen. Testpsykologien har utviklet seg til en egen spesialitet som etterhvert har tatt i sin tjeneste stadig mer kompliserte statistiske metoder.

Men selve prinsippet for testmetoden er altså enkelt nok: En lager en rekke standardiserte situasjoner eller prøver. Ved å anvende dem på tilstrekkelig mange forsøkspersoner kan en konstruere en fordelingskurve for svarene, og finne det normale gjennomsnitt. Nye forsøkspersoner kan en så senere plasere som så og så meget over eller under gjennomsnittet.

På denne måten kan en få et mer eller mindre pålitelig mål for svært mange evner og egenskaper hos et menneske, men naturligvis slett ikke for alle. Og spørsmålet av særlig interesse for strafferettspleien er: Hvordan er det med „målingen“ av de evner og egenskaper som har vesentlig betydning for straffutmålingen?

Jeg skal ikke her ta opp hele dette spørsmål. Det er *intelligensmålingen* som er emnet for diskusjonen. Det er vel også bare den som hittill har hatt noen større praktisk betydning for vår rettspleie.

Når det blir spørsmål om å „måle“ intelligensen, støter vi straks på en tilsynelatende uoverstigelig vanskelighet: Hvis vi skal måle noe, burde vi jo helst vite, *hva* det er vi vil måle. Men hva er egentlig intelligens?

Vi har nok alle et visst vagt begrep om det. Vi sier at den og den har en høy eller en lav intelligens. Men på den annen side vet vi også at intelligens kan ytre seg på mange forskjellige måter, og vi oppdager snart at det slett ikke er lett å få noe pålitelig inntrykk av hvor intelligent et menneske er. Noen er svært flinke til å snakke for seg, de imponerer sine medmennesker og går for å være intelligente. Men hvis de blir stillet overfor oppgaver, som krever virkelig omtanke og kombinasjonsevne, viser det seg at deres evner slett ikke på langt nær svarer til det inntrykk deres omgivelser hadde fått. Omvendt kan andre i vanskelige situasjoner vise seg meget mer intelligente enn en trodde de var.

Kan testmetoden her virkelig hjelpe oss til en mer objektiv bestemmelse?

De psykologene som først tok fatt på oppgaven å måle intelligensen, forsøkte å finne *en enkelt evne* som skulle være utslagsgivende og konstruere tests som målte denne evnen. *Ebbinghaus* f. eks. mente at det karakteristiske for intelligensen er *kombinasjonsevnen* og konstruerte sin berømte utfyllingstest for å måle denne evnen. *Spearman* oppfattet til å begynne med *skjelne-evnen* som den sentrale intellektuelle evne og søkte å måle den.

Ettersom en fikk mer erfaring, ble det imidlertid stadig mer klart, at intelligensen ikke kan identifiseres med noen slik enkelt ferdighet. Men samtidig oppdaget en også noe annet som førte ut over vanskeligheten. Det viste seg at resultatene av en hel mengde *forskjellige* tests samsvarte eller korrelerte med intelligensen, slik denne viste seg i skoleytelser, lærerbedømmelse o. s. v.

Det avgjørende vendepunkt i intelligensmålingens historie mar-

keres som bekjent ved *Binets* intelligensprøver. Binets store oppdagelse var nettopp den, at barns intelligens kan måles ved hjelp av prøver av overordentlig mange *forskjellige* slags. I de forskjellige måtene vi klarer problemer på må det derfor gjøre seg gjeldende noe *felles*. Den som klarer bra problemsituasjoner av *en* art, vil i alminnelighet også stå høyt, hvor det gjelder løsning av andre problemer. Omfattende korrelasjonsundersøkelser gir tallmessig uttrykk for samsvaret mellom ytelsene i forskjellige problemsituasjoner.

Hva dette felles — det som vi kaller den *alminnelige intelligens* egentlig er — kan vi imidlertid ikke si. Men — som den amerikanske intelligensforsker *Wechsler* uttrykker det — på lignende måte som vi bestemmer elektrisiteten gjennom de mange forskjellige utslag den gir seg — termiske, kjemiske, magnetiske — søker vi å bestemme intelligensen ved hjelp av de mange forskjellige ting den setter oss istand til å gjøre — forbinde adskilte erfaringer, trekke slutninger, forstå meningen av ord, løse matematiske problemer, bygge broer, navigere skib o. s. v.

I de seneste årene har en forøvrig gjennom den såkalte *faktoranalyse* søkt å finne metoder til å løse spørsmålet om intelligensens struktur. Det ville imidlertid føre for langt å komme nærmere inn på her. Faktoranalysen er en høyt utviklet statistisk spesialmetode.

Grunnlaget for antagelsen av en målbar *alminnelig intelligens* er altså på den ene side interkorrelasjonene mellom resultatene av mange forskjellige slags prøver og på den annen side samsvaret mellom prøveresultatene og intelligenskriterier som skoledyktighet og lignende.

Det første praktiske brukbare system for intelligensmåling ble utarbeidet av *Binet* i samarbeide med legen *Simon*. Disse Binet-Simons prøver er senere blitt revidert og tilpasset for forholdene i andre land. Mest kjent er de to amerikanske revisjonene av *Terman*. Her hjemme har *Lofthus* søkt å standardisere prøvene for norske forhold. Det er de såkalte Oslo-prøver.

Binets prøver er beregnet for *barn*. For hvert alderstrin er det utarbeidet en rekke oppgaver av en slik vanskelighetsgrad at de *gjennomsnittlig* blir klart av barn på dette alderstrin, men ikke av gjennomsnittsbarn som er yngre. En får altså 5-års prøver, 6-års prøver, 7-års prøver o. s. v. For de lave alderstrin er naturligvis prøvene meget enkle, men blir vanskeligere og vanskeligere ettersom en kommer oppover i årsklassene.

Foruten alderstrins-prøvene etter Binets system, er det senere også utarbeidet mengder av andre tests til å måle den alminnelige intelligensen. Når det gjelder å måle voksnes intelligens, har en mest brukt *gruppe-prøver*. Det store gjennombrudd for gruppe-testene kom under den første verdenskrig. Da utarbeidet amerikanske psykologer de berømte armé-prøver. Henimot 2 millioner mann ble prøvet med disse testene, som fikk en vesentlig betydning for „organisationen av den amerikanske armé. Ved hjelp av dem ble tatt ut dem som etter sitt intelligensnivå egnet seg til offiserskurser og spesialvåpen. På den annen side ble de skilt ut som var for dårlig utstyrt til overhode å gjøre brukbar soldattjeneste. Etter krigen vokste interessen for gruppe-prøver enormt og førte til et veldig arbeid med konstruksjon og standardisering av nye tests. Gruppe-prøvene er sammensatt av oppgaver av mange forskjellige arter, utfyllingsoppgaver, regneoppgaver, oppfatning av analogier og motsetninger o. s. v. I de senere år er det forøvrig også blitt utarbeidet en brukbar individualprøve for voksne, nemlig *Wechsler's* meget benyttede test.

Det vel mest brukte mål for intelligensen er den såkalte *intelligenskvotient*.

Et barn på 8 år som klarer 9 års trinets prøver sies å ha en levealder på 8 år og en såkalt intelligensalder, IA, på 9 år. Ved intelligenskvotienten, IQ, menes IA dividert med levealderen, LA, multiplisert med 100. For beregning av voksnes IQ regner en forøvrig gjerne med en levealder på 16.

Når bestemmelsen av et barns IQ har fått så stor betydning, er det først og fremst fordi denne intelligenskvotient har vist seg å være en relativt *konstant størrelse*. Derved blir det mulig med en høy grad av sannsynlighet å *forutsi* intelligensnivået på senere alderstrin.

Jeg må forresten tilføye at enkelte av de nyeste undersøkelser tyder på at vesentlige miljøforandringer kan fremkalle forholdsvis store forandringer i IQ hos *ganske små* barn, før 6 års alderen. Men iallfall når barn er kommet i skolealderen, kan vi i praksis regne med en relativt konstant IQ. En kan vente at den forskjell i IQ en finner ved senere testing ikke overskrider 5 points. Men i praksis er det ingen vesentlig forskjell mellom en IQ på 85 og 90 eller på 120 og 125. Den populære tro at et dumt barn kan vokse av seg sin dumhet og at et oppvakt barn godt kan bli en middels eller dum voksen er derfor avgjort gal. Det uintelligente barn blir praktisk talt uten unntagelse en uintelligent voksen, og omvendt. En kan ikke oppdra barn til å bli intelligente.

Av reglen om IQ's tilnærmede konstans følger altså at en som befinner seg langt nede på intelligensskalaen ikke — selv ved den mest intense trening i det intellektuelt mest stimulerende miljø — kan komme høyt opp. Hans IQ kan forskyves noe, men ikke meget. *Anlegget* setter bestemte *grenser* som ikke kan overskrides. Et menneske med IQ under 75 kan nok kanskje lære å lese og utføre den enkleste regning og forskjellig praktisk arbeid. Men lite ville det nytte ham å forsøke seg på realskoleeksamen eller artium. Og en kan ikke forlange av ham at han skal kunne bedømme en komplisert situasjon og forstå konsekvensene av sine handlinger på samme måte som en med en IQ på 110 eller 130.

Intelligensmålingen har naturligvis — som alt nytt — vært gjenstand for megen kritikk. Svært meget av denne kritikk behøver en imidlertid ikke å ta alvorlig. Det har nå i over 40 år vært drevet et intenst og omfattende forskningsarbeid på intelligensmålingens område. Og jeg tror trygt en kan si at metoden har vist sin store verdi både i teori og praksis. Intelligensmålingen er i virkeligheten en av den moderne psykologis store landvinninger. En annen sak er det at intelligensprøvene ofte har vært brukt altfor ukritisk. Ettersom en har fått større erfaring, er en imidlertid blitt stadig mer klar over de metodiske vanskeligheter og de mange feilkilder. Det bør derfor innskjerpes: *Intelligensmålingen som metode er en sak for ekspertene.*

Nettopp i denne sammenheng vil jeg da gjerne ha sagt noe om den anvendelse intelligensmålingen har fått i strafferettspleien. Binets prøver og de forskjellige revisjoner av dem er blitt uteksperimentert på og passer for *skolebarn*. Bruker en slike prøver på *voksne*, er det fare for store feil. Men det er nettopp det en i stor utstrekning har gjort. Det er derfor grunn til å betrakte den anvendelse rettspsykiaterne har gjort av intelligensmålingen med atskillig skepsis.

Jeg har hørt jurister og fornuftige legmenn gi uttrykk for slik skepsis: Hva i all verden bryr en gammel innbruddstyv seg om forskjellen mellom president og konge? Eller om å lage en fornuftig setning av ordene: Herre Hund En Forsvarer God Sin Tappert? Nei, sett ham overfor den oppgave å arrangere et vanskelig innbrudd, da vil det vise seg hvor intelligent han er!

Det er noe helt riktig i en slik innvending. *Skal en intelligensmåling ha mening, må testmaterialet passe for dem som blir målt, og også vekke tilstrekkelig interesse.* Å be en slitt husmor finne rim till dag, sill, vår (en av Oslo-prøvene på 12-års trinnet) eller en

underoffiser å lage en setning med ordene gutt, ball, elv, vil ikke vekke særlig interesse eller respekt for prøvene¹).

At det virkelig dreier seg om en alvorlig feilkilde, når en uten videre bruker Binet-revisjonene på voksne, fremgår klart nok av resultatene av de amerikanske armé-prøver. Ved å måle en utvalgt gruppe soldater både med Army- α og Stanford-Binet, fant en hvilken intelligensalder der skulle svare til de forskjellige points i armé-prøvene. Så undersøkte en fordelingen av intelligensaldre innenfor hele den målte armé. Resultatet en fant var nokså forbløffende: Omtrent 50 % av mennene i den amerikanske arméen hadde en intelligensalder lavere enn 13 år, og henimot 40 % av blomsten av Amerikas ungdom måtte klassifiseres som mer eller mindre „åndssvake“ („feeble-minded“), d. v. s. hadde en IQ under 75.

Her er åpenbart noe som ikke kan stemme. Og feilen ligger nettopp i den måten samsvaret mellom IA og points i armé-prøven ble bestemt på. Man målte voksne, som forlenget var ute av skolen, med prøver standardisert for skoleungdom og mente at samme ytelse måtte bety samme intelligensalder.

Det er svært viktig å være oppmerksom på denne feilkilden, nettopp når det gjelder anvendelsen av intelligensprøver i rettspleien. I sin „Rettspsykiatri for jurister og leger“ betegner prof. *Langfeldt* dem med IQ mellom 55 og 75 som debile, altså lidende av evnesvakhet i sterkere grad enn de såkalte sinker. Så vidt jeg vet er imidlertid den norske revisjon av Stanford-Binetprøvene nettopp den intelligensprøve som vår retts-medisinske kommisjon har godkjent. Jeg er redd for at skulle vi måle hele den voksne norske befolkning med disse prøvene, ville det vise seg at betraktelig mer enn en tredjepart av oss var „åndssvake“!

Et slikt resultat ville kanskje ikke si så meget om det norske folks lave intelligens. Men det maner unektelig til større forsiktighet enn rettspsykiaterne hittil har vist.

Selv er jeg sterkt overbevist om den store verdien intelligensprøver kan ha både i strafferettspleien og på andre områder. Etter alt å dømme vil testmetodene i fremtiden få en overordentlig stor samfunnsmessig betydning. Men de må brukes med forsiktighet og kritikk. Forsvarlig intelligensmåling krever i virkeligheten grundig psy-

¹) Smlgn. *D. Wechsler*: The measurement of adult intelligence, s. 17.

kologisk skolering, både en bred utdanning i generell psykologi og psykometrisk spesialutdanning.

Intelligenstestene er ikke målestaver, som kan brukes overalt og forskyves fra en gruppe mennesker til en annen uten forandring. Tests laget for et skolebarnsmiljø kan ikke brukes på voksne, som forlengst er vokset ut over hele skolebarnsinnstillingen. Det var alt Binet selv fullt klar over; testene må — innskjerpet han — bruke materiale som passer for deres miljø som skal testes, og de har bare validitet for grupper, som med hensyn til testmaterialet har vesentlig samme miljø. Binets kloke advarsel ble imidlertid senere glemt av så altfor mange.

I strafferettssaker blir intelligensmåling idag her i landet betraktet som et *rettsmedisinsk* spørsmål, og den rettsmedisinske kommisjon blir betraktet som øverste sakkyndige instans. Nå er det ganske visst riktig at det *finns* leger som f. eks. dr. Dahlstrøm og dr. Lofthus som har gjort noe av et pionerarbeid på dette område. Men allikevel er det helt galt å betrakte legene *qua leger*, selv om de har psykiatrisk spesialistutdanning, som fagmenn i testpsykologi. Hele ordningen når det gjelder undersøkelse av siktedes såkalte mentale tilstand, er i virkeligheten foreldet. Når det oppnevnes to sakkyndige, burde det være en psykiater og en psykolog.

Psykologens spesialkunnskap trenges når det gjelder evnemålingen. Men også når det gjelder å bedømme den sjelelige helhetstilstand, er hans medarbeiderskap ønskelig på grunn av hans bredere utdanning i generell psykologi og i sosialpsykologi.

Harald Schjelderup.