**2.4 Om kritiske vurderinger af psykologiske undersøgelser – validitet, reliabilitet og forstyrrende faktorer**

I den første del af kapitlet har vi gennemgået, hvordan man kan udføre psykologisk forskning, og hvilke fordele og ulemper der er ved de forskellige metoder. Der er dog en række ting, som kan svække kvaliteten
af psykologiske undersøgelser, og det skal vi kigge på i det følgende. Det er nemlig ofte nødvendigt at foretage en kritisk vurdering af værdien af en undersøgelse, så man ved, om man overhovedet kan bruge undersøgelsen til noget, dvs. om man kan bruge dens resultater i en analyse af eller diskussion i forhold til en psykologisk problemstilling. For at kunne foretage en kritisk vurdering er det godt at kende til begreberne validitet, reliabilitet og forstyrrende faktorer/variable (fejlkilder m.v.).

# 2.4.1 Validitet

En undersøgelses validitet betyder dens gyldighed, dvs. om den faktisk viser det, som den er designet til at vise. Validitet er ikke et spørgsmål om, hvorvidt en undersøgelse har eller ikke har validitet, men hvor høj en grad af validitet der er tale om. Man skelner mellem to former for validitet:

* **Intern validitet**, som handler om, hvorvidt en undersøgelse er designet på en måde, så den afdækker det, som forskerne gerne vil afdække. Der kan bl.a. være tale om, at designet sigter skævt. En spørgeskemaundersøgelse om unges alkoholforbrug har f.eks. en meget lav grad af validitet, hvis alle spørgsmålene kun handler om lektiepresset i gymnasiet. Så er det meget svært at konkludere noget om alkoholforbrug. Der kan også være tale om en lav grad af validitet i et eksperimentelt design, hvor det er nødvendigt, at deltagerne ikke må vide, hvad forsøget går ud på, men hvor eksperimentsituationen faktisk meget let bliver gennemskuet af deltagerne, så de handler ud fra dette.

### Eksempel på undersøgelse, hvor der kan have været problemer med den interne validitet

Den amerikanske psykolog Stanley Milgram udførte i 1960’erne nogle berømte og berygtede forsøg om lydighed over for autoriteter, hvor han testede, om forsøgsdeltagere ville adlyde eksperimentlederen som en autoritetsperson og tildele en anden deltager elektriske stød, hvis denne deltager svarede forkert i en række huskeopgaver. Forsøget var manipuleret på den måde, at der ikke blev givet rigtige stød til den anden deltager (som var skuespiller), der befandt sig i et andet rum end den rigtige forsøgsdeltager. Den rigtige forsøgsdeltager var placeret foran en imponerende stødtildelingsmaskine og skulle for hvert forkert svar øge strømstyrken med 15 volt indtil en maksimal styrke på helbredstruende 450 volt – medmindre han nægtede at fortsætte. Arkivstudier har dog vist, at en af Milgrams assistenter analyserede sig frem til, at jo mere en deltager troede på, at den anden person ikke fik rigtige stød, desto mere tilbøjelig var vedkommende til at gå hele vejen til 450 volt. En stor portion af deltagerne i forsøgene har altså tilsyneladende gennemskuet, at der ikke blev tildelt rigtige stød – hvilket flere deltagere også gav udtryk for undervejs i forsøgene.

* **Ekstern validitet (også kaldet økologisk validitet)**, som handler om, hvorvidt man kan overføre resultater og konklusioner fra en undersøgelse til virkelighedens verden, dvs. virkeligheden uden for undersøgelsens ’verden’. Det er specielt i laboratorieeksperimenter med en meget høj grad af kontrol over variable, at den eksterne validitet er lav. Eksperimentsituationen (både omgivelserne og deltagernes opgaver) kan være så kunstig og virkelighedsfjern, at det kan være svært at konkludere noget om, hvordan mennesker uden for laboratoriet opfører sig.

# 2.4.2 Deltagerudvælgelse

For at sikre en høj grad af validitet i en undersøgelse skal man sørge for, at man har udvalgt en passende deltagergruppe (et såkaldt sample). I nogle undersøgelser (typisk de kvantitative) er man interesseret  i at lære noget om store grupper eller hele befolkninger, og i andre undersøgelser (typisk de kvalitative) er man blot interesseret i at lære noget om en lille gruppe.

En meget passende deltagergruppe er en gruppe, som udgør et repræsentativt udsnit af den større gruppe (dvs. den population), som man gerne vil vide noget om. Når et udsnit er repræsentativt, så afspejler det den større gruppe, dvs. det indeholder den samme sammensætning af forskellige folk. Det kan være svært at lave et udsnit, der repræsenterer alle, for i princippet skal udsnittet afspejle både kønsfordelingen, aldersfordelingen, den kulturelle sammensætning, den geografiske spredning m.v. i populationen. Dertil kommer, at man ikke nødvendigvis bare kan få fat i de deltagere, som man gerne vil have. Hvis deltagerne skal svare på f.eks. en annonce om at medvirke i et eksperiment, så er det muligvis nogle bestemte typer af folk, som vælger at besvare annoncen. Mange psykologiske undersøgelse er uheldigvis præget af, at man har udvalgt en meget ensartet og snæver undersøgelsesgruppe (f.eks. universitetsstuderende), som sjældent afspejler den bredere population, som man gerne vil lære noget om.

# 2.4.3 Reliabilitet

En undersøgelses reliabilitet betyder dens pålidelighed, dvs. om undersøgelsen kan gentages (eventuelt af andre forskere) og give de samme resultater. Reliabilitet måles ligesom validitet i grader. Umiddelbart kunne man tro, at undersøgelser selvfølgelig giver de samme resultater, hvis de gentages, men der er en række forstyrrende faktorer (ofte kaldet fejlkilder), som kan påvirke en undersøgelses pålidelighed. Dette vil blive uddybet i næste afsnit, hvor vi vil kigge på forskellige typer af forstyrrende faktorer/variable (også kaldet bias).

# 2.4.4 Forstyrrende faktorer (fejlkilder m.v.)

* **Forsøgseffekt** (eng. observer effect): Fordi deltagerne i en undersøgelse ved, at de bliver undersøgt, så kan der være en tendens til, at de præsterer bedre eller på en anden måde, end de ville gøre, hvis de ikke blev observeret. Forsøgseffekten kan være særlig udtalt i eksperimenter og åbne observationer.
* **Selvrepræsentationseffekt**: De fleste mennesker vil gerne fremstå som bl.a. ordentlige, høflige, pålidelige og generelt gode mennesker over for andre. Dette ønske om at præsentere sig selv på en god
og socialt acceptabel måde kan også påvirke deltageres adfærd i undersøgelser. Selvrepræsentationseffekten kan finde sted i både observationer (åbne), eksperimenter, spørgeskemaer og interviews. I specielt spørgeskemaer og interviews kan effekten medføre, at deltagerne lyver, fordi de er bange for, at det sande svar vil stille dem i et dårligt lys.
* **Kontroleffekt** (kaldes også fordringseffekt – eng. demand characteristics): Vi mennesker er meget nysgerrige, og vi vil meget gerne vide, hvad det er, der sker med os. Når folk deltager i en undersøgelse, så ved de, at forskerne har en hypotese, og denne intention med undersøgelsen kan deltagerne ikke lade være med at prøve at gætte. Det er underordnet, om deltagerne rent faktisk gætter rigtigt, for pointen med kontroleffekten er, at de retter deres adfærd ind, så den passer med deres gæt. På den måde kan det give unaturlige resultater – specielt hvis deltagerne har fået penge for at deltage og derfor gerne vil være gode deltagere, der hjælper forskerne. For at modvirke denne effekt bruger man ofte en hvid løgn om, hvad undersøgelsen handler om (såkaldt enkeltblinding), så man forhåbentlig lidt bedre kan styre deltagernes gætterier. Kontroleffekten kan være særlig udtalt i eksperimenter, men kan også finde sted i åbne observationer og spørgeskemaer.

### Eksempel på demonstration af kontroleffekten

Martin Orne og hans kolleger demonstrerede i 1960’erne, hvor stærk kontroleffekten kan være. De gav bl.a. forsøgsdeltagere en bunke med 2000 ark papir og satte dem til selv at skrive 224 plusstykker pr. ark papir, som de efterfølgende skulle regne. Når de var færdige med et ark, skulle de blot fortsætte på det næste ark og blive ved, indtil de blev bedt om at stoppe. Forskerne havde taget deltagernes armbåndsure, så deltagerne ikke var klar over, hvor længe de havde været i gang. De fleste deltagere fortsatte med den meningsløse opgave med plusstykkerne i over fem timer (!) uden at brokke sig. De fortsatte altså i unaturligt lang tid, blot fordi de vidste, at de deltog i et psykologisk eksperiment og sandsynligvis tænkte, at det handlede om udholdenhed eller noget i den stil.

* **Forsøgsledereffekt** (eng.researcher bias): Denne effekt kaldes også Rosenthal-effekten efter den amerikanske psykolog Robert Rosenthal, der har forsket meget i den. Effekten handler om, at forskeres hypoteser om deres undersøgelser (dvs. deres forventninger til, hvad de finder) kan komme til at påvirke deltagerne i undersøgelsen på en sådan måde, at forskeren får de resultater, som han eller hun gerne vil have. Det er ofte på ubevidste og meget subtile måder, at forskere får hjulpet deltagere hen i den ønskede retning – det er sjældent for at fuske med resultaterne. For at modvirke denne effekt bruger man i specielt eksperimenter såkaldt dobbeltblinding, hvor hverken deltagerne eller den forsker, som har med deltagerne at gøre, ved, hvad hypotesen er, eller om deltagerne er i enten kontrol- eller eksperimentgruppen. Forsøgsledereffekten kan være særlig udtalt i eksperimenter, men den kan også finde sted i åbne observationer, spørgeskemaer og interviews.

### Eksempel på undersøgelse, hvor forsøgsledereffekten lader til at have spillet en rolle

I 1950’erne udførte den tyrkisk-amerikanske psykolog Muzafer Sherif og hans kolleger en række klassiske forsøg, hvor 11-årige amerikanske drenge fra den hvide middelklasse blev opdelt i to grupper på en sommerlejr i et stort naturområde (bl.a. i Robbers Cave State Park i Oklahoma). De to grupper skulle konkurrere mod hinanden i forskellige dyster, og det viste sig, at de to grupper hurtigt udviklede en meget stærk negativ holdning til hinanden, og det kom ligefrem til voldelige sammenstød såsom slåskampe og ødelæggelse af hinandens hytter. Arkivstudier af bl.a. forskernes noter fra undersøgelserne har dog vist, at forskerne i høj grad opmuntrede drengene til at være fjendtlige over for hinanden. Et af Sherif og kollegers forsøg (i 1953) endte også med, at drengene i samlet flok gjorde oprør mod forskerne, fordi de ikke ønskede at være uvenner, sådan som forskerne havde forsøgt at præge dem til. Forskerne havde bl.a. ødelagt den ene gruppes telt i håbet om, at gruppen ville give den anden gruppe skylden.

Kilde: Magnus Riisager: *Psyk C – Grundbog til psykologi på C-niveau.* Systime