

ØKOLOGISK VALIDITET OG EKSPERIMENTEL SPÆDBARNSFORSKNING¹

Peter Krøjgaard

Økologisk validitet er et centralt begreb inden for den økologiske psykologi. I eksperimentel spædbarnsforskning, som er forfatterens eget forskningsfelt, nævnes begrebet til gengæld relativt sjældent. Nærværende artikel har til formål at undersøge, hvorvidt begrebet økologisk validitet kan bidrage til den eksperimentelle spædbarnsforskning. Først præsenteres forskellige definitioner af begrebet. Dernæst sammenholdes og diskuteres begrebet i forhold til en række konkrete eksperimentelle undersøgelser. Det konkluderes, at økologisk validitet absolut er et vigtigt begreb, men at begrebet skal antage en specifik karakter, hvis det skal være anvendeligt. Samtidig søges det vist, at en sådan specificering imidlertid er vanskelig.

Efter dekader med behaviorismen som den dominerende almenpsykologiske retning på det internationale landkort blev det med den kognitive revolutions indtræden sidst i 1950'erne endelig igen betragtet som stueren psykologisk videnskabelig praksis at beskæftige sig med mentale fænomener generelt og kognitive fænomener i særdeleshed (Gardner, 1987; Neisser, 1967). I og med at man flyttede fokus fra adfærden til informationsprocesseringen i den »sorte boks« nedtonede man imidlertid samtidig betydningen af den livsverden, hvor menneskets erkendelse fandt sted (Gardner, 1987). Kognitionsforskningens (*Cognitive Science*, eng.) begrænsede interesse for konteksten for erkendelse var ikke blot et tilfældigt opstået biprodukt, men et resultat af en bevidst forskningsstrategi – den *metodiske solipsisme*. Tesen var, at *The Mind* i princippet kunne undersøges i sig selv uden synderlig hensyntagen til den verden, som det informationsprocesserende system skulle leve og overleve i (Stillings, Feinstein, Garfield, Rissland, Rosenbaum, Weisler, & Ward, 1987).

Økologisk psykologi har som kognitionsforskningen også menneskelig erkendelse som sit genstandsområde, men adskiller sig fra kognitionsforsk-

1 Jeg vil gerne takke Dorthe Berntsen og Johan Trettvik for indsigtfulde kommentarer og forslag til en tidligere udgave af det foreliggende manuskript.

ningen ved at påbegynde analysen af den menneskelige erkendelse med et ganske andet spørgsmål: Hvilke kompetencer har vi som mennesker brug for med henblik på at kunne leve og overleve i vores specifikke økologiske niche? Dette anderledes udgangspunkt synes at have mindst to overordnede konsekvenser: For det første, at økologisk psykologi i langt højere grad end kognitionsforskningen forpligter sig på en analyse af menneskets livsverden. For det andet, at økologisk psykologi fortrinsvis tager afsæt i kognitive fænomener og problemstillinger, som de optræder i menneskers dagligdag, frem for i laboratoriet.

Der findes bestemt deciderede økologiske *teorier*, hvoriblandt særligt J.J. Gibsons økologiske perceptionsteori (Gibson, 1966; 1972; 1979) formentlig er både den mest radikale og den mest konsekvent gennemførte. For andre forskere er økologisk psykologi ikke nødvendigvis et egentligt og endeligt opgør med kognitionsforskningen, men skal snarere anskues som en anden *tilgang* til menneskelig erkendelse, hvor det har høj prioritet, at de fænomener der undersøges, hvad enten det foregår i eller uden for laboratoriet, skal være direkte relevante for menneskers erkendelse i dagligdagen. Eksponenter for denne mere moderate tilgang finder vi både blandt udlandets forskere (f.eks. Neisser, 1976; 1985) og blandt vores egne (f.eks. Berntsen, 1996; 1998; *in press*; Larsen & Thompson, 1995; Larsen & Conway, 1997; Mammen, 1993).

En central term i den økologiske tilgangs kritik af kognitionsforskningen er begrebet *økologisk validitet*. Coolican (1999) definerer begrebet således:

Extent to which investigation results can be generalised to other places and conditions, in particular from the artificial and/or controlled (e.g. laboratory) to the natural environment (Coolican, 1999, p. 107).

Økologisk validitet kan således rubriceres som en underkategori af *ekstern validitet*, der vedrører resultatets generaliserbarhed ud over forsøgssituationen, hvor den økologiske validitet særligt synes forbeholdt potentielle problemer hermed i forbindelse med laboratorieforsøg. Mit eget forskningsfelt omhandler spædbørns kognitive udvikling – et område, der i vid udstrækning gør brug af laboratorieforsøg. I artikellitteraturen inden for dette genstandsfelt er det imidlertid relativt sjældent, at man støder på begrebet økologisk validitet. Men heraf kan man ikke slutte, at begrebet er irrelevant for denne type forskning. Selvom begrebet om økologisk validitet sjældent nævnes, så kunne det jo være, at begrebet alligevel havde noget at tilbyde den eksperimentelle spædbarnsforskning. I denne artikel skal vi derfor se nærmere på, hvorvidt begrebet økologisk validitet bidrager til den eksperimentelle spædbarnsforskning. Jeg vil forsøge at argumentere for, at begrebet bestemt er anvendeligt; at begrebet faktisk ofte anvendes implicit, dvs. uden eksplicit at henvise til termen; at en dogmatisk anvendelse af begrebet kan være decideret hæmmende for vidensproduktionen; at begre-

bet bør specificeres, men at en sådan specificering samtidig er en vanskelig opgave.

Artiklen er disponeret som følger: Først skal vi se på nogle forskellige afgrænsninger af begrebet økologisk validitet, og hvilke konsekvenser dette har for begrebets anvendelighed. Dernæst vil jeg præsentere en række eksperimentelle undersøgelser og sideløbende sammenholde disse med en række specifikke kriterier for økologisk validitet, som anføres i første del. Sidst følger en sammenfatning.

Forskellige afgrænsninger af begrebet *økologisk validitet*

Begrebet økologisk validitet stammer fra Egon Brunswik (1956), som anvendte begrebet som et led i en kritik vedrørende manglende fokus på den forsøgssituation, som subjektet befinder sig i. Med tiden er begrebet imidlertid blevet anvendt i en lidt anden betydning med større fokus på generaliserbarhedsaspektet, særligt i forhold til laboratorieforsøg. Det gælder f.eks. Coolicans (1999, p. 107) definition ovenfor, ligesom det gælder Corsinis definition, som følger her:

The concept of ecological validity pertains mainly to the areas of comparative psychology and animal learning, but it can apply to any research activity as to *whether its findings have any relevance to real-life situations* or habitats for the class of subjects involved, including humans (Corsini, 1994, p. 457, min kursivering).

Der synes således at herske en vis konsensus omkring den *generelle* forståelse af begrebet økologisk validitet i den forstand, at alle udlægninger af begrebet placerer kravet om generalisering fra laboratoriet til forsøgspersonernes livsverden centralt.

Det er formentlig de færreste, der vil påstå, at kravet om økologisk validitet ikke skulle være relevant og vigtigt. Det er imidlertid ulig meget nemmere at påpege et problem, end det er at løse ét. Og det er spørgsmålet, hvor anvendeligt begrebet er, så længe det alene afgrænses relativt abstrakt. Neisser (1976) argumenterer således for, at hvis begrebet økologisk validitet skal kunne bidrage til udviklingen af bedre undersøgelser, så skal kritikken af laboratorieforsøg bevæge sig væk fra det generelle og abstrakte plan og være mere specifik. Han skriver f.eks.:

Demands for ecological validity are only intelligible if they are specific. They must point to particular aspects of ordinary situations that are ignored by current experimental methods, and there must be good reason to suppose that those aspects are important (Neisser, 1976, p. 34).

Neisser advarer altså mod at anvende begrebet økologisk validitet som en generel og abstrakt standardindvending mod eksperimentelle undersøgelser og hævder, at begrebet økologisk validitet skal konkretiseres, hvis det skal være anvendeligt i praksis. En sådan konkretisering kan selvsagt foretages på flere måder. Én mulighed består i at eksemplificere forskriften negativt i form af en specificering af, hvad der kendetegner laboratorieforsøg, der er økologisk *invalide*. Neisser skriver således:

Contemporary studies of cognitive processes usually use stimulus material that is abstract, discontinuous, and only marginally real. It is almost as if ecological *invalidity* were the deliberate feature of the experimental design (Neisser, 1976, p. 34, forfatterens kursivering).

Det er altså anvendelsen af et *kunstigt stimulus-materiale*, der ifølge Neisser er den primære årsag til manglende økologisk validitet i datidens kognitionsforskning. Ved denne konkretisering bliver Neissers anvendelse af begrebet økologisk validitet måske mere snæver, men også mere specifik, hvorved begrebet nærmer sig en egentlig operationalisering. Som vi skal se, så er netop operationalisering en egenskab, som vi bør sætte højt.

Gibsons (1979) kritik af den traditionelle laboratorietilgang synes at have et andet fokus end dét, vi finder hos Neisser (1976), selvom der også var fællestræk. Ligesom Neisser forholdt Gibson sig også kritisk til kunstigt stimulus-materiale, hvilket f.eks. manifesterede sig i hans afstandtagen fra anvendelsen af billeder og fotografier som det primære forsøgsmateriale i mange traditionelle visuelle perceptionsforsøg (Gibson, 1979). Men Gibsons væsentligste indvending mod datidens eksperimentelle undersøgelser af visuel perception synes imidlertid at have været en anden, nemlig at subjektet i forsøgssituationen typisk var afskåret fra de *handlemuligheder* og *responser* i form af eksplorativ adfærd, som han eller hun normalt havde til sin rådighed uden for laboratoriet, og som for Gibson (1979) var uløseligt forbundet med dét at percipere.

Flere har gennem tiden angivet forskellige mulige konkretiseringer af begrebet, og Schmuckler har forsøgt at systematisere disse konkretiseringer og rubricerer dem inden for følgende tre områder:²

- [1.] the types of stimuli used (they should reflect the subject's naturalistic setting),
- [2.] the nature of the experimental setting (it should be as similar as possible to the subject's everyday environment), and

2 Schmucklers review er *in press* i samme nummer af det nye tidsskrift *Infancy*, som Lewkowicz (*in press*) også udkommer i. Det er imidlertid kun lykkedes mig at få fat i Lewkowicz' artikel, hvorfor jeg ikke kan henvise til Schmucklers artikel, men må nøjes med en andenhånds reference via Lewkowicz.

- [3.] the type of response (it should capitalize on the subject's available response repertoire) (Schmuckler, *in press*, efter Lewkowicz, *in press*, p. 1, mine tilføjelser i kantet parentes).

Systematikken rummer altså ud over Neissers krav til anvendelsen af naturlige stimuli (jf. punkt 1. ovenfor) og Gibsons insisteren på bibeholdelsen af subjektets naturlige handlemuligheder (jf. punkt 3. ovenfor) et krav om, at forsøgssituationen i videst muligt omfang bør afspejle subjektets almindelige livsverden.

Nærværende artikel fokuserer på forholdet mellem økologisk validitet og eksperimentel spædbarnsforskning. Og den enkeltstående person, der nok har haft størst indflydelse på udviklingspsykologien med sin kritik af traditionelle udviklingspsykologiske laboratorieforsøg, er formentlig Bronfenbrenner (1977; 1979; Bronfenbrenner & Morris, 1998). Det var som bekendt Bronfenbrenner, der med slet skjult reference til Ainsworths Strange Situation metode skrev:

[...] much of developmental psychology, as it now exists, is *the science of the strange behavior of children in strange situations with strange adults for the briefest possible periods of time* (Bronfenbrenner, 1979, pp. 18-19, forfatterens kursivering).

Bronfenbrenner insisterede på, at en kontekstbeskrivelse var nødvendig, hvis man ville forstå barnets udvikling. Han var således en af de første, der med sin *økologiske model* præsenterede en egentlig ramme til systematisk beskrivelse af barnets miljø. Sammen med f.eks. McCall (1977) var han også blandt de første, der forholdt sig kritisk til store dele af datidens fremherskende eksperimentelle forskning (Fabes, Martin, Hanish, & Updegraff, 2000). Sidst i 1970'erne var forholdet mellem økologisk validitet og laboratorieforsøg allerede blevet diskuteret flittigt, men selvom Bronfenbrenner var en af de fremmeste kritikere af megen af datidens eksperimentelle forskning, var han faktisk ikke modstander af laboratorieforsøg. Nogle samtidige forskere havde udlagt begrebet økologisk validitet således, at en undersøgelse for at være økologisk valid *skulle* foretages i naturlige omgivelser med dagligdags objekter og aktiviteter (jf. punkt 2 i Schmucklers systematisering). Herom skrev Bronfenbrenner:

Although originally attracted to this notion, I have upon reflection come to view it not only as too simplistic but as scientifically unsound on several counts (Bronfenbrenner, 1979, p. 28).

En radikal anvendelse af begrebet, hvor undersøgelser per definition *skal* foregå i naturlige omgivelser, anser Bronfenbrenner altså ikke for befor-

drende for vidensproduktionen. I forlængelse heraf definerede Bronfenbrenner selv økologisk validitet således:

Ecological validity refers to the extent to which the environment *experienced by the subjects* in a scientific investigation has the properties it is supposed or assumed to have by the investigator (Bronfenbrenner, 1979, p. 29, min kursivering).

Sammenligner vi Bronfenbrenners definition med de tidligere fremførte af Coolican (1999) og Corsini (1994) finder vi både ligheder og forskelle. Ligheden mellem disse definitioner af økologisk validitet består i, at de alle placerer generaliseringsaspektet helt centralt. Bronfenbrenners definition adskiller sig imidlertid fra de øvrige på to punkter. For det første ved at være mere *specifik* med hensyn til, hvilke kriterier, der skal være opfyldt for, at denne generalisering fra forsøgssituationer til forsøgspersonens dagligdag er mulig. Herved kan Bronfenbrenners definition anskues som en tilnærmelse mellem en abstrakt afgrænsning og en mere specifik anvendelse af begrebet. For det andet ved eksplicit at placere en del af *vurderingen* af, hvorvidt forsøgssituationen opleves, som den var tilsigtet af forsøgslederen, hos *forsøgspersonen*, frem for at overlade vurderingsansvaret til forsøgslederen alene (jf. Bronfenbrenners definition ovenfor). Netop det sidste punkt er efter min vurdering et helt afgørende aspekt af begrebets afgrænsning, men det vender vi tilbage til.

Efter at have set på forskellige afgrænsninger af begrebet økologisk validitet retter vi nu fokus mod, hvorledes begrebet manifesterer sig i forskellige eksperimentelle undersøgelser inden for spædbarnsforskningen. Med dette formål for øje vil jeg dels anvende Schmucklers systematisering og dels Bronfenbrenners definition. Sidstnævnte er medtaget som analytisk redskab, da den, som vi skal se, synes særligt relevant netop i forhold til spædbarnsforskning.

Eksempler fra den eksperimentelle spædbarnsforskning

Jeg har bestræbt mig på at præsentere de første af eksemplerne, så de følger Schmucklers systematik. Derefter skal vi se eksempler på, hvorledes Bronfenbrenners anvendelse af begrebet også er relevant. Som det vil fremgå, kan overlap mellem de enkelte fænomener dog ikke helt undgås.

Problemet med kunstige stimuli

I en eksperimentel undersøgelse af 10 måneder gamle børns opfattelse af objekters numeriske identitet (Krøjgaard, 2000) havde jeg brug for et såkaldt *baseline* mål for, hvor meget mere interessante 2 objekter var for børnene sammenlignet med ét objekt. Til det formål blev børnene placeret

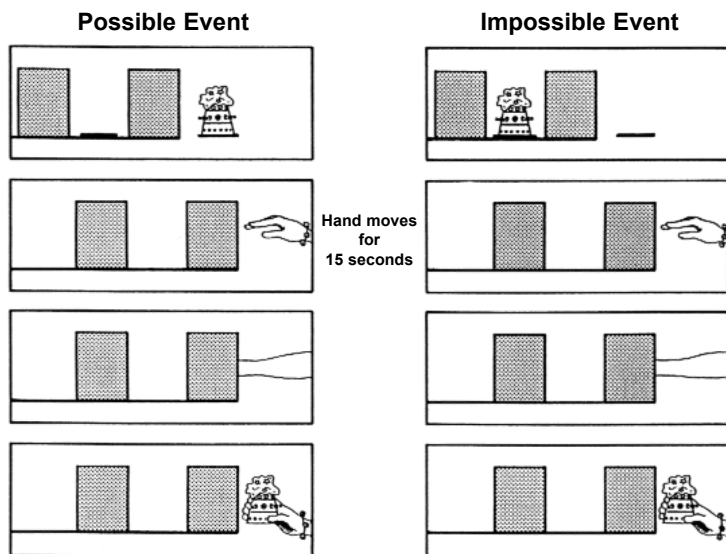
foran en dukketeater-lignende opstilling, hvor de i balanceret rækkefølge blev præsenteret for henholdsvis ét og to objekter. Børnenes interesse for de respektive stimuli blev vurderet ved hjælp en den såkaldte blikfokuspræference-metode (se f.eks. Spelke, 1985; Krøjgaard, 1998), hvor man sammenligner tidsestimater fra, hvor længe barnet kigger på de relevante stimuli. Rationalet er, at jo længere fokus varer, jo større er interessen. Jeg begyndte med at præsentere børnene for enten en eller to farvede træklodser i terning-, pyramide-, eller cylinderform. Det viste sig imidlertid hurtigt, at disse stimuli ganske enkelt var alt for kedelige for børnene, der kiggede alle mulige andre steder hen, end lige netop på objekterne. Først da træklodserne blev erstattet med sødt-udseende bamser, blev det muligt at fastholde børnenes fokus tilstrækkeligt længe til, at et baselinemål kunne tilvejebringes (Krøjgaard, 1999). Eksemplet er måske ikke det mest ophidsende, men det demonstrerer, hvad man kan risikere, hvis man anvender stimuli, der ikke har nogen relevans eller interesse for børnene.

Er det rimeligt at hævde, at anvendelsen af kunstige stimuli *per automatik* diskvalificerer laboratorieforsøg? Det mener jeg ikke er tilfældet. Lad os se på endnu et eksempel: I kølvandet på Piagets (1954) pionérarbejde omkring objektpermanens fulgte en lang række undersøgelser, der blandt andet søgte efter mulige forklaringer på den såkaldte Stadie-IV-fejl eller A-ikke-B fejl. Stadie-IV-testen består af en slags gemme-leg, hvor det 8-9 måneder gamle barn skal finde et objekt, der typisk gemmes i én af to containere. Piaget gjorde den iagttagelse, at når barnet efter gentagne gange at have genfundet objektet ved gemmested A, hvor det kort forinden var blevet skjult, nu var vidne til, at objektet blev gemt ved B, så ledte barnet ofte *ikke* efter det nyligt gemte objekt ved gemmested B, men derimod ved gemmested A (Piaget, 1954). Det har siden vist sig, at barnets hukommelse har en vis indflydelse på forekomsten af disse fejl. Stadie-IV-fejlen fremprovokeres nemlig typisk, hvis det 8-9 måneder gamle barn skal vente omkring 3 sekunder, inden det får lov til at søge efter objektet, medens fejl-frekvensen daler, hvis det tidsrum, som barnets skal vente, reduceres eller fjernes helt (Gratch, Appel, Evans, LeCompte, & Wright, 1974; Harris, 1973; for et review se Wellman, Cross & Bartsch, 1987).

Det var imidlertid spørgsmålet, om det *alene* var barnets hukommelsesmæssige begrænsninger, der resulterede i de systematiske søgefejl, eller om søgefejlene snarere var et produkt af hukommelsesmæssige begrænsninger *i kombination* med andre faktorer. I et forsøg på at afklare denne problematik simplificerede Baillargeon & Graber (1988) opgaven, så vellykket respons alene afhang af hukommelsesmæssige kompetencer. Til det formål udviklede de et elegant design, hvor 7-8 måneder gamle børn ikke selv manuelt skulle søge efter det objekt, der blev gemt, men hvor man i stedet ved hjælp af blikfokuspræference-metoden registrerede børnenes reaktioner på et scenarie, hvor en *anden* fandt objektet under forskellige

betingelser. Børnene blev i balanceret rækkefølge præsenteret for en »mulig« og en »umulig« testbegivenhed (se figur 1).

Figur 1: Skematisk fremstilling af den mulige/forventede og den umulige/ikke-forventede testbegivenhed (fra: Baillargeon & Graber, 1988).



I begge begivenheder blev et besynderligt udseende objekt placeret på en platform, og umiddelbart herefter blev objektet skjult af en skærm. En handskebeklædt hånd vævede nu i femten sekunder over de to gemmesteder inden den fremdrog objektet – *enten* på den forventede lokalisering (den »mulige« testbegivenhed), *eller* på den ikke-forventede lokalisering (den »umulige« testbegivenhed). Man tog så tid på, hvor længe børnene kiggede på henholdsvis den forventede og den ikke-forventede testbegivenhed. Rationalet var, at hvis børnene kunne huske lokaliseringen af det skjulte objekt i de femten sekunder, som tildækningen varede, så ville børnene undre sig (dvs. kigge i længere tid), hvis objektet efterfølgende dukkede op på en *anden* lokalisering. Dette var præcis, hvad resultaterne fra undersøgelsen viste. Børnene kiggede i signifikant længere tid på den ikke-forventede testbegivenhed, hvor objektet blev hentet frem fra en anden lokalisering end der, hvor objektet oprindeligt var blevet gemt, sammenlignet med den forventede testbegivenhed (Baillargeon & Graber, 1988). Objekterne var skjult i ca. fem gange så lang tid, som der skal til for at »producere« en Stadio-IV-fejl med børn, der er en måned *ældre*, og alligevel havde børnene ikke problemer med at huske lokaliseringen af objek-

terne. Dette forsøg var derfor med til at vise, at de 8-9 måneder gamle børns problemer med den klassiske Stadie-IV-test næppe alene skulle tilskrives hukommelsesmæssige begrænsninger.

Hverken selve objektet, der blev gemt, eller den handskebesklædte hånd, der vævede over og siden fremdrog objekterne, kan siges at være objekter, som børnene typisk ville stifte bekendtskab med i deres dagligdag, snarere tværtimod. Dertil kommer, at selve blikfokuspræference-metoden, hvor man sammenligner varigheden af blikfokus over for henholdsvis forventede (eller mulige) og ikke-forventede (eller umulige) testbegivenheder, nærmest per definition strider mod forskriften om, at såvel stimuli som forsøgssituationen (jf. næste afsnit) så vidt muligt skal minde om forsøgspersonens dagligdag. Hvis man ukritisk tilslutter sig en restriktiv udlægning af forskrifterne vedrørende økologisk validitet af de anvendte stimuli og af forsøgssituationen (jf. anvisningen fra Brofenbrenners samtidige), ville ikke blot ovennævnte undersøgelse, men hele rækken af forsøg, der gør brug af børns reaktioner på ikke-forventede eller »umulige« testbegivenheder skulle kasseres. Spørgsmålet er imidlertid, om man ikke herved ville afskære sig fra at anvende en lang række frugtbare metodiske redskaber. Jeg vender tilbage til denne problematik i artiklens sidste afsnit. Vi skal nu se på det andet af de tre anførte problemer med laboratorieforsøg.

Problemet med kunstige forsøgssituationer

Siden midten af 1970'erne har der været foretaget talrige undersøgelser af spædbørns præverbale, kommunikative færdigheder. Sådanne studier er både interessante i forbindelse med afdækningen af fundamentet for den egentlige sprogtilegnelse, og fordi de kaster lys over de tidligste ontogenetiske manifestationer af menneskebarnet som socialt og kommunikativt væsen. Vi skal se på et eksempel på en sådan undersøgelse:

Man kan bestemt blive klogere af at observere ansigt-til-ansigt kommunikation mellem f.eks. en voksen og et 2-3 måneder gammelt spædbarn direkte. Men når kommunikationen forløber glat, kan det undertiden være vanskeligt at afgøre, hvorvidt barnet vitterligt selv bidrager til kommunikationen, eller om det blot er den voksne, der tilpasser sig barnets ytringer. I bestræbelserne på at afdække, hvilke forventninger barnet kan tænkes at have til sådanne samspil, gør man derfor ofte brug af en strategi, hvor man bevidst *forstyrrer* samspillet, og så registrerer, hvorledes barnet reagerer på disse forstyrrelser. Colwyn Trevarthen og hans medarbejdere har foretaget en række undersøgelser af præverbalt, kommunikativt samspil mellem mor og spædbarn ud fra denne strategi. F.eks. undersøgte Murray & Trevarthen (1985) samspillet mellem 6-12 uger gamle spædbørn og deres mødre under forskellige betingelser. Mor og barn var placeret hver for sig foran hver deres TV-skærm. Ovenpå hver skærm var placeret et kamera, der filmede den person, der så på skærmen. Det filmede signal blev transmitteret til kommunikationspartnerens skærm og omvendt. Barnet kunne således på

sin skærm se sin mor, der tilsvarende på sin skærm kunne se barnet. Begge forsøg begyndte med, at mor og barn havde et fint ansigt-til-ansigt samspil trods de lidt specielle forhold, men så blev der introduceret nogle forstyrrelser: I et af eksperimenterne blev mødrene forud for selve forsøgene instrueret i at lave et apatisk, udtryksløst ansigt – et »sten-ansigt«. Når de havde kommunikeret med barnet i ca. et halvt minut, skulle mødrene uden varsel lave et »sten-ansigt«, og børnenes reaktioner herpå blev registreret. Børnene reagerede på sten-ansigtet med ubehag. Nogle forsøgte med gestik, smil og lyde at få mor til at genoptage kommunikationen, men da det ikke lykkedes, trak børnene sig typisk tilbage (Murray & Trevarthen, 1985, eksperiment 1).

Ovennævnte forsøg indikerer, at 6-12 måneder gamle børn forventer, at mor reagerer på deres ytringer under ansigt-til-ansigt samspil. For at kunne indfri barnets forventninger er det imidlertid ikke tilstrækkeligt, at mor reagerer; hun skal tilmed reagere på en sådan måde, at reaktionen tidsmæssigt er afstemt med barnets kommunikative adfærd. Dette fremgik af resultaterne fra det efterfølgende eksperiment, hvor kommunikationen igen efter et halvt minuts normalt samspil blev forstyrret ved, at barnet på et tidspunkt ikke længere kommunikerede med moren *on-line*, men blev præsenteret for en optagelse af moderens tidligere kommunikative samspil med barnet. Moderen var således i en vis forstand stadig »aktiv« i kommunikationen, men aktiviteten var ikke længere synkron med barnets ytringer på grund af den tidsmæssige forskydning. Børnene reagerede med at blive irriterede og ligeglade, ligesom forekomsten af antal positive emotioner daledede (Murray & Trevarthen, 1985, eksperiment 2).

Forsøgene illustrerer, at børn blot 6-12 uger gamle synes at kunne indgå i samspil, hvor de opfører sig, som om de forventer, at man skiftes til at »tale« sammen, og at deres adfærd skal give anledning til kontingente responser fra kommunikationspartneren. Men forsøgssituationen, hvor kommunikationen finder sted via kamera og TV-skærm, minder ikke meget om, hvad man kan forvente, at børnene har været vant til i deres 6-12 uger gamle liv. Hvis man insisterer på, at sådanne undersøgelser bør foregå i naturlige omgivelser, synes det derfor umiddelbart nærliggende at betvivle validiteten af en sådan undersøgelse. Man skal imidlertid lægge mærke til, at ansigt-til-ansigt kommunikationen – trods de noget specielle rammer herfor – faktisk forløb uden problemer *inden* forstyrrelserne blev iværksat.

Vi skal kort se på endnu et eksempel på en videnskabsende undersøgelse, hvor forsøgssituationen adskiller sig væsentligt fra, hvad forsøgspersonerne må formodes at være vant til. Eksemplet er det nu klassiske *visual cliff*-perceptionsforsøg, hvor man undersøgte 6-14 måneder gamle børns evne til dybdeperception (Gibson & Walk, 1960). Opstillingen bestod af en massiv flade, hvorpå der var placeret et stykke refleksfrit glas hen over en »kløft«, så det så ud som om, at man ville falde ned, hvis man

bevægede sig ud over kløften. Mor blev så placeret, så kløften var mellem hende og barnet. Ved at kalde på barnet skulle mor nu prøve at få sit barn til at kravle ud over kløften og hen til sig. Selv når mor forsøgte at vise barnet, at glaspladen var solid ved at banke på den, var barnet tilbøjeligt til *ikke* at kravle over til sin mor (Gibson & Walk, 1960). Igen kan man næppe hævde, at denne forsøgssituation direkte skulle afspejle børnenes daglige liv. Det synes imidlertid samtidig vanskeligt at levere en konsistent afvisning af forsøgets resultater, ligesom det forekommer absurd at skulle betvivle *relevansen* af forsøgets resultater for børnenes motoriske udfoldelser. Efter min opfattelse synes det således ikke rimeligt at hævde, at forsøgssituationer, der adskiller sig fra forsøgspersonernes dagligdag, *per automatik* er problematiske.

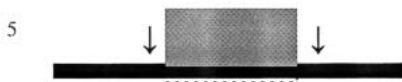
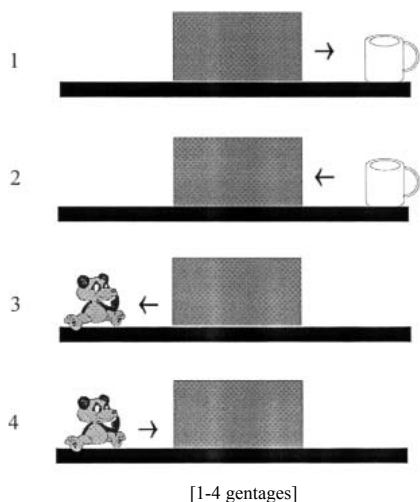
Vi skal nu se på det tredje og sidste af de af Schmuckler anførte foki for anvisninger af problemer i forbindelse med laboratorieforsøg.

Problemet med indskrænkningen af forsøgspersonernes responser

Forsøgspersoners begrænsede handlemuligheder i en eksperimentel undersøgelse kan undertiden være et problem, hvilket Gibson overbevisende har argumenteret for i forbindelse med sine undersøgelser med voksne. Dette synes imidlertid også at være et relevant problem i forhold til undersøgelser af børn. Til illustration heraf skal vi igen se på et eksempel fra min egen forskning: Undersøgelsen (Krøjgaard, *in prep.*) placerer sig i umiddelbar forlængelse af den tidligere omtalte undersøgelse (Krøjgaard, 2000) og handler også om 8 og 10 måneder gamle børns evne til at bestemme et objekts numeriske identitet. Undersøgelsen tager afsæt i en replikation af en undersøgelse af Xu & Carey (1996). For at tydeliggøre problematikken er det nødvendigt at præsentere selve designet relativt detaljeret, så jeg vil bede læseren væbne sig med tålmodighed: I en dukketeater-lignende opstilling præsenteres børnene for begivenheder, hvor to objekter på skift bevæger sig ind og ud bag en skærm. Der opereres med to testbetingelser: En egenskabsbetingelse og en spatio-temporal betingelse (se figur 2a og 2b).

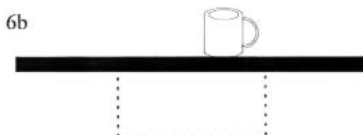
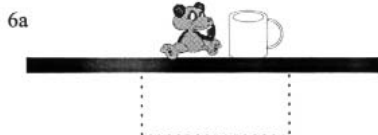
I *egenskabsbetingelsen* ser barnet ikke på noget tidspunkt begge objekter samtidig, men *kun* hver for sig. I den *spatio-temporale betingelse*, derimod, ser børnene indledningsvist begge objekter *samtidig*. Selve testen består i, at skærmen på et tidspunkt, hvor ingen objekter er synlige for barnet, bliver sænket ned i dukketeaterets scenegulv, hvorved antallet af objekter bag skærmen bliver afsløret for barnet. I hver testbetingelse opereres der med to mulige testbegivenheder, en forventet og en ikke-forventet testbegivenhed: I den forventede testbegivenhed svarer antallet af objekter bag skærmen til det antal, som man som voksen ville forvente, at der indgår i den pågældende begivenhed, nemlig *to* objekter. I den ikke-forventede testbegivenhed, derimod, kommer der kun *ét* objekt til syne bag

Figur 2a: Skematisk fremstilling af egenskabsbetingelsen



Forventet resultat

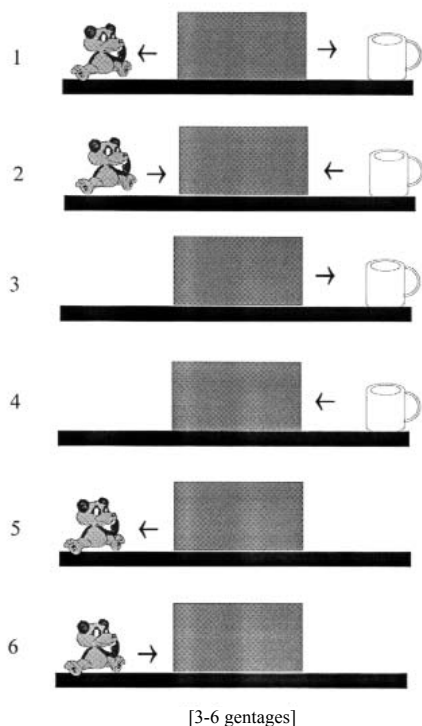
Ikke-forventet resultat



skærmen.³ Hvert barn præsenteres nu skiftevis for både forventede og ikke-forventede testbegivenheder i balanceret rækkefølge. Ved hjælp af et trykknapsystem, der er forbundet med en computer, tages der tid på, hvor længe børnene kigger på de pågældende begivenheder. Scoringen foregår via et kamera, hvorigennem man kan se barnet, men ikke objekterne. Den, der

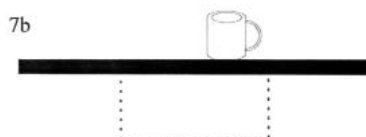
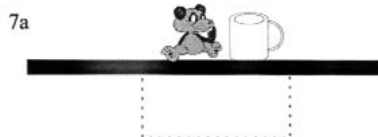
3 I den ikke-forventede testbegivenhed blev »forsvindingsnummeret« i praksis udført ved, at forsøgsassistenten, så snart bamsen var skjult bag skærmen, hurtigt fjernede bamsen gennem en skjult åbning i scenegulvet inden skærmen sænkedes. Under hele forsøget blev der tilført støj, så børnene ikke kunne høre, om objektet blev fjernet eller ej.

Figur 2b: Skematisk fremstilling af den spatio-temporale betingelse.



Forventet resultat

Ikke-forventet resultat

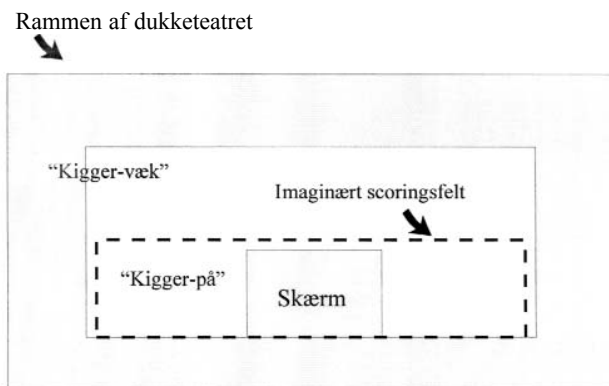


scorer begivenhederne, er således blind for, hvorvidt der i det enkelte tilfælde scores en forventet eller en ikke-forventet testbegivenhed. Blikfokusværdien er et resultat af summen af de »kigger på«-sekvenser, som barnet genererer, indtil barnet kigger væk i en sammenhængende periode på 2 se-

kunder, hvorved den enkelte *trial* afsluttes. Et computerprogram sørger dels for, at der bliver givet signal, når en *trial* er færdig og dels for, at de enkelte »kigger-på«-sekvenser bliver akkumuleret til en blikfokusværdi. Enkelte aspekter kan varieres, men essensen af denne procedure har været standard inden for denne type undersøgelser siden midten af 1980'erne (se f.eks. Baillargeon, 1986; 1993; 1995; Baillargeon, Spelke, & Wasserman, 1985; Simon, Hespos & Rochat, 1995; Wilcox & Baillargeon, 1998a, 1998b; Xu & Carey, 1996).

Medens ovennævnte procedure-karakteristikum altid fremgår af artiklernes forsøggennemgang, så er det stort set aldrig beskrevet i artiklerne, hvorledes »kigger-på«-feltet afgrænses. Der foreligger alene en implicit konvention herom: Barnet siges at kigge på begivenheden, når det kigger inden for den del af dukketeateret, hvor enten objekterne, skærmen eller »andet«, der måtte indgå i begivenhederne, har bevæget sig (personlig kommunikation, Renée Baillargeon; Amy Needham; Fei Xu). Nedenfor er afbildet det »imaginære« scoringsfelt, som blev anvendt i nærværende undersøgelse, og som er i overensstemmelse med ovennævnte implicitte konvention (se figur 3).

Figur 3: Skematisk fremstilling af det imaginære scoringsfelt.



Der er altså konsensus omkring afgrænsningen af scoringsfeltet blandt de folk, der anvender den pågældende metode; men som vi skal se, så er det spørgsmålet om ikke denne snævre afgrænsning giver anledning til visse problemer, når man netop skal undersøge børns reaktioner over for objekter, der *forsvinder* fra synsfeltet.

Rationalet bag blikfokuspræference-metoden består i, at hvis barnet undrer sig over den ikke-forventede testbegivenhed, så vil det kigge i relativt længere tid på denne sammenlignet med den forventede testbegivenhed. Denne teknik har vist sig at være særdeles effektiv og har været medvirkende til at ændre vores opfattelse af, hvor meget spædbørn forstår af

deres fysiske omverden. Viden tilvejebragt på baggrund af denne metodik vil således være at finde i enhver internationalt anerkendt udviklingspsykologisk grundbog. Metoden har den umiddelbare fordel, at barnets motoriske formåen *ikke* er et potentielt begrænsende led, når barnet skal demonstrere, hvad det ved (Baillargeon, 1993; 1995 Spelke et al., 1992), hvilket har vist sig at være tilfældet med de klassiske studier af barnets motorisk baserede søgen efter skjulte objekter (Diamond, 1988; 1991), som oprindeligt blev designet af Piaget (1954).

I nærværende undersøgelse synes der imidlertid at være følgende problem: *Hvis* barnet vitterligt undrer sig, når skærmen i den ikke-forventede testbegivenhed sænkes og afslører, at der bag skærmen kun forefindes ét objekt i stedet for to, hvad gør barnet så? Det leder selvfølgelig (kigger efter) efter det forsvundne objekt. Og da objektet jo ikke er til at få øje på inden for scoringsfeltet, så kigger barnet undertiden efter det forsvundne objekt *uden for* scoringsfeltet. Men hvis barnet leder efter objektet uden for scoringsfeltet i en sekvens af kortere eller længere varighed, så bliver disse sekvenser *ikke* registreret som »kigger-på«-enheder som en følge af afgrænsningen af scoringsfeltet. Og hvis en sådan sekvens varer mere end to sekunder, så vil sekvensen ligefrem terminere den pågældende trial. Operationaliseringen af scoringsfeltet indebærer altså i dette tilfælde, at barnets naturlige »lede-respons« klassificeres som *manglende* interesse for begivenheden, hvilket det altså netop ikke behøver at være udtryk for.

Ville det så ikke blot være rimeligt at udvide scoringsfeltet? Nej, det ville det desværre ikke, for så ville man blot ende med det modsatte problem, så at sige: Man ville øge risikoen for at score børn, der reelt er uinteresserede, som om de faktisk ledte efter det forsvundne objekt – altså såkaldte »falsk positive«. En del af de sekvenser, hvor børnene kigger uden for scoringsfeltet *er* jo reelt udtryk for børnenes manglende interesse for den viste begivenhed og bør derfor også scores som sådanne. Man vælger derfor en konservativ strategi, hvor man foretrækker at minimere risikoen for »falsk positive« men med den bagside, at man risikerer at miste nogle reelt positive responser.

Som et eksempel på, hvordan forsøgspersonernes naturlige responser faktisk *blev* tilgodeset, selvom der var tale om et laboratorieforsøg, vil vi igen henlede opmærksomheden på Gibson & Walks (1960) *visual cliff*-forsøg. Trods den noget specielle forsøgssituation synes det ikke rimeligt at hævde, at de 6-14 måneder gamle kravlende spædbørn var berøvet deres naturlige responsmuligheder. Ved at have gode muligheder for at kravle til og fra kanten af kløften, bibeholdt børnene netop mulighederne for den eksplorativadfærd, som er helt central i denne sammenhæng.

Vi kan nu forsøge at sammenholde ovennævnte eksempler med den tredje af de af Schmuckler anførte anvisninger, der vedrører problemer med forsøgspersonernes indskrænkede responser. I det første af de ovenfor anførte eksempler (hvor man registrerede børnenes varighed af blikfokus

over for forventede og ikke-forventede testbegivenheder) er der måske ikke ligefrem tale om, at forsøgspersonerne direkte er berøvet deres naturlige handlemuligheder eller responser (bortset fra muligheden for at kravle), som Gibson så indtrængende advarede imod; børnene kan for så vidt kigge, hvorhen de vil. Problemet består derimod i, at aspekter af den naturlige »lede-adfærd« som en følge af scoringsfeltets afgrænsning bliver rubriceret som manglende interesse. Børnene blev så at sige ikke honoreret for deres naturlige adfærd, tværtimod. Eksemplet illustrerer hermed nogle af de problemer, der kan opstå som en følge af, at man med anvendelsen af en særlig metode samtidig indskrænker forsøgspersonernes handlemuligheder på en ikke altid hensigtsmæssig facon. I modsætning hertil bibeholdt børnene i Gibson & Walks' (1960) *visual cliff*-forsøg mulighederne for den naturlige eksplorative adfærd, selvom undersøgelsen foregik i laboratorieregi.

Ovenfor har vi set nogle eksempler, der falder inden for den systematisering af forskellige konkrete manifestationer af begrebet økologisk validitet, som Schmuckler havde opregnet. Det er imidlertid min opfattelse, at netop Bronfenbrenners definition af begrebet økologisk validitet, hvor der lægges særlig vægt på *forsøgspersonens oplevelse* af forsøget, er meget vigtig i nærværende sammenhæng. Som det så ofte er tilfældet med metodologiske problemstillinger, er det nemmest at illustrere relevansen af Bronfenbrenners udlægning af begrebet økologisk validitet ud fra konkrete eksempler, hvilket er emnet for det følgende afsnit.

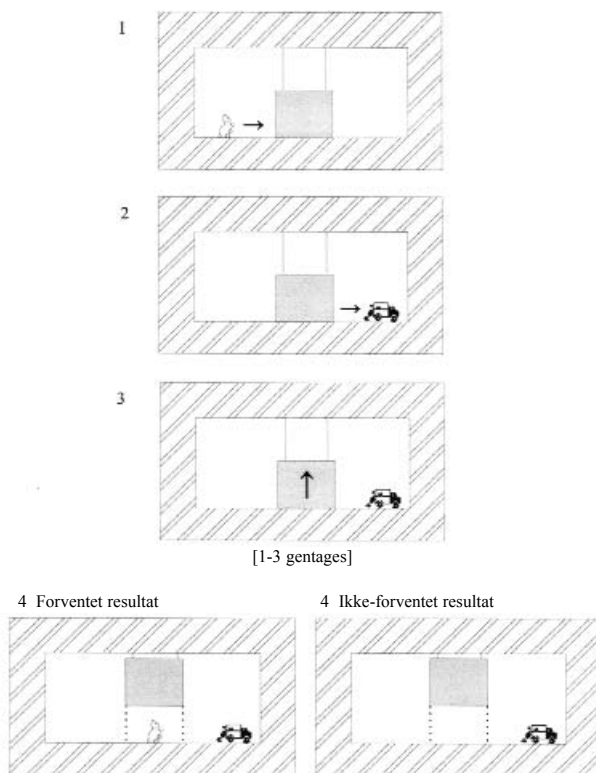
Bronfenbrenners pointe: Vigtigheden af korrespondens mellem forsøgspersonens oplevelse og forsøgslederens intention

Bronfenbrenners kriterie for økologisk validitet indebærer, at der skal være korrespondens mellem forsøgspersonens oplevelse af forsøget og dét, som er intenderet af forsøgslederen. Styrken ved denne definition er, at den ud over at bevare generaliseringsaspektet, også giver mulighed for at legitimere design, der måske umiddelbart ligger langt fra forsøgspersonens naturlige livsverden, men som af forsøgslederen antages at indebære vigtige aspekter. Et klassisk og illustrativt eksempel på, at dette kan være en effektiv strategi til at frembringe ny viden, finder vi i Milgrams klassiske socialpsykologiske eksperiment over lydighed over for autoriteter (for et review se Milgram, 1975). Hverken forsøgssituationen eller den opgave, som forsøgspersonerne blev stillet i Milgrams forsøg, minder om forsøgspersonernes dagligdag. Men det ændrer ikke på, at resultaterne fra undersøgelsen er tankevækkende og interessante.⁴ Vi vender os nu atter mod eksperimentelle undersøgelser af spædbørn.

4 At Milgrams forsøg var problematiske ud fra en etisk betragtning er en anden sag. Over for forsøgspersonerne var undersøgelsen åbenlyst uetisk. Dilemmaet består selvfølgelig i, at den viden, som forsøget på kløgtig vis tilvejebragte, samtidig er i menneskehedens interesse.

Med voksne forsøgspersoner har man ofte gode muligheder for at undersøge, om de nu også oplevede det ene eller andet aspekt ved en sådan undersøgelse, sådan som det var tilsigtet af forsøgslederen. Man kan jo blot spørge dem. Med helt små børn (og dyr) som forsøgspersoner forholder det sig imidlertid anderledes. Man kunne naturligvis godt stille verbale spørgsmål til 10 måneder gamle spædbørn om, hvordan forsøget blev oplevet, men »svarene« ville næppe være brugbare. Dertil kommer, at voksne forsøgspersoner undertiden selv ville gøre opmærksom på, hvis de mente, at der var noget fundamentalt galt med en opstilling – en mulighed, som heller ikke er til stede med spædbørn som forsøgspersoner. Det betyder, at når man foretager undersøgelser med helt små børn, så skal man være særligt varsom med, om de nu også formodes at opleve forsøget sådan, som det var tilsigtet. I det følgende skal vi se et eksempel, hvor det ikke var tilfældet. Igen vil jeg trække på min egen forskning.

I den første af de omtalte undersøgelser (Krøjgaard, 2000) anvendte jeg indledningsvist et design, hvor skærmen, som kunne skjule objekter, blev hævet for at afsløre, hvor mange objekter, der var gemt bag skærmen under testbegivenhederne (se figur 4).



Figur 4: Skematisk fremstilling af egenskabsbetingelsen (pilot design).

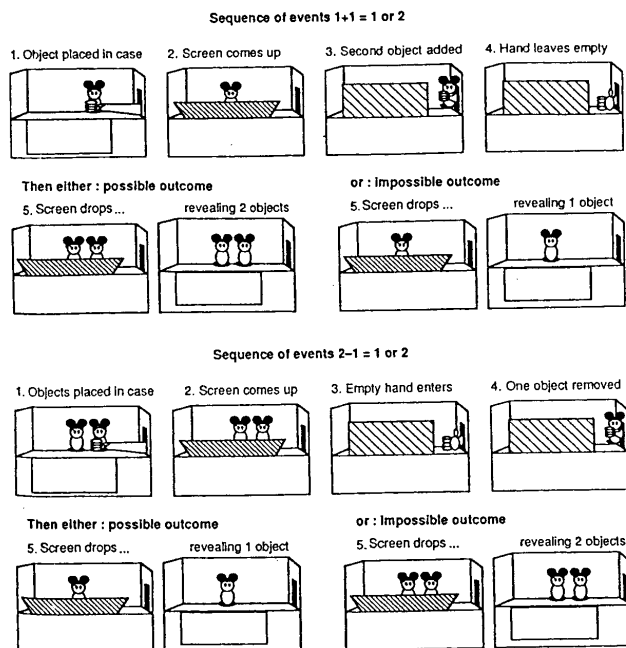
Da de første børn blev præsenteret for undersøgelsen, viste det sig imidlertid hurtigt, at opstillingen havde mindst én afgørende fejl. Så snart skærmen blev hævet, og vi skulle registrere, hvor længe børnene kiggede på det eller de objekter, der var gemt bag skærmen, så kiggede stort set ingen af børnene på objekterne, som jeg jo havde forventet, men derimod på *skærmen*, som blev hævet! Som forsøgsdesigner var jeg jo i min naivitet overbevist om, at børnene ville finde det rasende interessant, hvor mange objekter, der nu var skjult bag skærmen. Det var jo netop dét spørgsmål, mine hypoteser vedrørte. Men børnene delte ikke min interesse; de kiggede, som 10 måneder gamle børn jo gør, derimod derhen, hvor der skete noget i begivenheden – nemlig på skærmen, som var det eneste objekt i bevægelse på dette tidspunkt af testbegivenheden.

Læg mærke til, at dette forsøg ikke nødvendigvis forbyder sig mod de tidligere anførte forskrifter, der er at finde i Schmucklers systematik. Stimuli (bamserne) var ikke specielt kunstige (jf. punkt 1); forsøgssituationen adskilte sig ikke fra en lang række andre forsøg af denne type (jf. punkt 2), og barnets responser var ikke yderligere begrænset set i forhold til lignende undersøgelser (jf. punkt 3). Alligevel synes det evident, at designet er problematisk. Medens det ikke umiddelbart er indlysende, at designet mangler økologisk validitet ud fra de ovenfor anførte anvisninger, systematiseret af Schmuckler, så er problemet imidlertid ikke til at overse, hvis vi inddrager Bronfenbrenners definition af begrebet: Designet var ikke økologisk validt i Bronfenbrenners forstand, da børnene ikke oplevede begivenheden, som den var tilsigtet fra min side. Problemet blev løst ved at tilføje en plade i dukketeaterets ramme således, at skærmen forsvandt ind bag pladen og dermed ud af børnenes synsfelt, når den blev hævet.

Det var mit held, at børnene med deres tydelige respons gjorde det klart, at de fokuserede på noget andet i opstillingen end dét, som var tilsigtet. Jeg blev på en meget direkte facon gjort opmærksom på designets manglende økologiske validitet, hvilket gav mig mulighed for at foretage de fornødne ændringer. Det er imidlertid langt fra sikkert, at børn som forsøgspersoner altid er helt så meddelsomme, som det var tilfældet i dette forsøg. Nogle gange er det formentlig langt vanskeligere at afgøre, præcis hvordan børnene forstår forsøget. Det synes netop at være tilfældet i det næste eksempel, vi skal se på.

Karen Wynn har med afsæt i en række undersøgelser, der også gør brug af blikfokuspræference-metoden, argumenteret for, at spædbørn er i stand til at udføre simpel aritmetik (Wynn, 1992; 1995; Wynn & Chiang, 1998). Vi skal se nærmere på det første af hendes studier. I en dukketeater-lignende opstilling blev fem måneder gamle børn præsenteret for en række simple »regnestykker«. Det foregik ved, at børnene så et objekt blive placeret bag en skærm, hvorefter skærmen blev vippet op, så objektet blev skjult. Dernæst så barnet en hånd, der placerede endnu et objekt inde bag skærmen (se figur 5).

Figur 5: Skematisk fremstilling af de mulige/forventede og de umulige/ikke-forventede testbegivenheder af såvel » $1+1=1$ eller 2-sekvensen« som af » $2-1=1$ eller 2-sekvensen« (fra: Wynn, 1992).



Umiddelbart herefter blev skærmen vippet ned, og barnet var vidne til en forventet og en ikke-forventet testbegivenhed: I den forventede testbegivenhed var antallet af objekter, der kom til syne efter skærmfald i overensstemmelse med det regnestykke, som barnet umiddelbart forinden havde været vidne til. I den ikke-forventede testbegivenhed, derimod, passede regnestykket ikke. I stedet for de to objekter, der skulle have været bag skærmen, hvis regnestykket skulle passe, var der kun ét objekt. Ud over dette »additionsstykke« ($1+1=2$ eller 1) blev børnene også præsenteret for et analogt »subtraktionsstykke« ($2-1=1$ eller 2). Varigheden af barnets blikfokus blev registreret og sammenlignet for de respektive testbegivenheder. Resultaterne fra undersøgelsen viste, at børnene reagerede med overraskelse (kiggede længere), når de blev præsenteret for de ikke-forventede testbegivenheder (regnestykkerne med et fejlagtigt resultat), sammenlignet med de forventede – uanset om der var tale om et additionsstykke eller et subtraktionsstykke. Wynn konkluderede, at resultaterne viste, at 5 måneder gamle børn var i stand til at udføre simple aritmetriske opgaver, og undersøgelsen blev publiceret i *Nature* under i overskriften *Addition and subtraction by human infants*.

Selve Wynns undersøgelse er uden tvivl velgennemført, og den interne logik i designet er konsistent. Det er imidlertid spørgsmålet, om børnenes systematiske responser beror på ægte aritmetiske færdigheder, eller om de blot skal tilskrives børnenes opfattelse af, hvorledes objekter kan, og ikke kan, bevæge sig i tid og rum (eller en helt tredje fortolkning). Som læseren sikkert har bemærket, så minder Wynns undersøgelse lidt om mine egne (eller omvendt!). Og nogle af forfatterne bag de undersøgelser, som mine egne undersøgelser eksplicit tager afsæt i (Simon, Hespos & Rochat, 1995; Xu & Carey, 1996), forsøger mere eller mindre eksplicit at refortolke Wynns resultater. En alternativ fortolkning af Wynns undersøgelser består i, at børnene undrer sig, hvis objekter enten *forsvinder* i tid og rum (jf. den ikke-forventede testbegivenhed, hvor $1+1=1$), eller hvor objekter *dukker op* ud af den blå luft (jf. den ikke-forventede testbegivenhed, hvor $2-1=2$). Kort sagt: Børnene undrer sig, hvis objekter bevæger sig *diskontinuert* (f.eks. Spelke et al., 1992; Spelke et al., 1995). Såvel den aritmetiske som den fysiske forklaring er logisk konsistent og i overensstemmelse med resultaterne fra Wynns undersøgelse.

I hvert enkelt eksperiment er det snarere reglen end undtagelsen, at der er flere logisk konsistente fortolkninger. Af samme grund er det sjældent, at en undersøgelse blot indeholder ét eksperiment. Ofte er det nødvendigt med en mindre serie af flere på hinanden følgende forsøg, hvor konkurrerende fortolkninger ved hjælp af kontrolforsøg skridt for skridt kan udspilles mod hinanden. Med et eller flere kontrolforsøg er det således ofte muligt at sandsynliggøre, hvilken af to eller flere konkurrerende fortolkninger af et forsøgsresultat, der er den mest sandsynlige. Men netop i Wynns undersøgelse er det formentlig slet ikke muligt inden for forsøgsrammen at afgøre, om børnenes responser skal forklares *enten* med henvisning til en aritmetisk forståelse *eller* med henvisning til en forståelse af objekters bevægelser i tid og rum. Dette beror på, at man ikke i analog form kan præsentere børnene for et simpelt additions- eller subtraktionsstykke, hvor resultatet er forkert uden, at fremstillingen *samtidig* indebærer diskontinuerede objektbevægelser og vice-versa. I bestræbelserne på at afgøre, hvilken fortolkning, der er den mest korrekte, synes man derfor tvunget til at ty til evidens *uden for* forsøgsrammen. Haith (1998; Haith & Benson, 1998) forholder sig således kritisk til Wynns fortolkning af hendes resultater. Ifølge Haith er der grund til at nære skepsis over for Wynns fortolkning – at 5 måneder gamle børn forstår simpel aritmetik – ud fra den betragtning, at det samtidig er velkendt at 3-årige har vanskeligheder ved sådanne opgaver!

Afslutningsvis bør det nævnes, at selv om en undersøgelse skulle vise sig at honorere alle tænkelige krav til økologisk validitet, så ville dette ikke i sig selv være nogen garanti for, at der var tale om en interessant og videns-skabende undersøgelse. Der er nemlig intet principielt til hinder for, at en undersøgelse, der er økologisk valid, kan være behæftet med *andre*

graverende typer af fejl, såsom f.eks. logiske fejlslutninger i hypotesedannelsen, eller uhensigtsmæssig udvælgelse af forsøgspersoner.

Sammenfatning

Det overordnede budskab i begrebet økologisk validitet, nemlig at resultater opnået i laboratoriet bør kunne generaliseres til forhold uden for laboratoriet, er utvivlsomt både relevant og vigtigt – også for spædbarnsforskningen. Fra flere sider er det imidlertid blevet påpeget, at hvis begrebet økologisk validitet skal være anvendeligt som en retningslinje for udførelse af forsøg alment og laboratorieforsøg i særdeleshed, så må begrebet antage en mere *konkret* karakter (Neisser, 1976; Lewkowicz, *in press*).

Af Schmucklers systematisering fremgik, at man netop har forsøgt en sådan konkretisering ved at pege på *stimuli*, *forsøgssituationer* og *forsøgspersonernes responser* som tre specifikke områder, der skal opfylde særlige krav, såfremt laboratorieforsøg skal kunne kaldes økologisk valide. Det var imidlertid spørgsmålet, hvorvidt netop de specifikke anvisninger, som blev angivet af Schmuckler, reelt rummede et sæt kriterier, der på den ene side sikrede, at undersøgelser vitterligt informerede os om egentlige psykologiske forhold, og på den anden side ikke på forhånd diskvalificerede undersøgelser, der måske på overfladen forekom arbitrære, men som faktisk gjorde os klogere. Lad os se på de enkelte anvisninger i forhold til de fremførte eksempler fra den eksperimentelle spædbarnsforskning:

For det første skulle de anvendte stimuli afspejle subjektets naturlige omgivelser (jf. punkt 1 i Schmucklers systematik). Jeg fremførte i den forbindelse et eksempel fra min egen forskning (Krøjgaard, 2000), hvor anvendelsen af (for) kunstige stimuli reelt *var* et problem. Men der blev også præsenteret et eksempel på en undersøgelse fra Baillargeons laboratorium (Baillargeon & Graber, 1988), hvor der ganske vist blev anvendt kunstige stimuli, men hvor forsøgets resultater ikke desto mindre var relevante, velkontrollerede og begribelige. Det forekommer derfor tvivlsomt om anvendelsen af kunstige stimuli *per se* skulle være en indikator for manglende økologisk validitet. Lewkowicz (*in press*) forholder sig også kritisk til denne radikale udlægning af kriterierne og foreslår i stedet, at de stimuli, der anvendes i laboratoriet ikke nødvendigvis behøver at være præcis, som de er uden for laboratoriet, men at stimuli, der anvendes i laboratorieforsøg, skal bære den information, som er relevant i de naturlige omgivelser, eller at informationen skal være af en type, som subjektet er udviklet til at behandle uden for laboratoriet (Lewkowicz, *in press*, p. 8).

For det andet skulle forsøgssituationen så vidt muligt minde om den kontekst, som forsøgspersonen lever i uden for laboratoriet (jf. punkt 2 i Schmucklers systematik). Da det forekommer indlysende, at meget kunstige forsøgssituationer kan give anledning til problemer, tillod jeg mig helt

at udelade eksempler herpå. Derimod blev der præsenteret eksempler på elegante design (Murray & Trevarthen, 1985; Gibson & Walk, 1960), hvor de anvendte forsøgssituationer befandt sig langt fra forsøgspersonernes dagligdag, men hvor resultaterne trods dette alligevel ikke umiddelbart kunne afvises. Hvad var det, der gjorde, at det i disse design ikke syntes problematisk at anvende en forsøgssituation, der adskilte sig fra forsøgspersonernes dagligdag uden for laboratoriet? Det skyldtes formentlig, at de særlige forsøgsomstændigheder ikke kompromitterede netop *de* specifikke fænomener, der var genstand for de respektive undersøgelser (hhv. 1. kommunikationen mellem mor og barn, og 2. barnets dybdeperception), men alene andre aspekter. I Murray & Trevarthens (1985) undersøgelse forløb det præverbale kommunikative samspil således problemfrit indtil forstyrrelserne satte ind; og i Gibson & Walks (1960) forsøg var de kinæstetiske dybde-*cues* fuldt ud lige så tilgængelige, som de ville have været uden for laboratoriet. Disse undersøgelser synes altså at vise, at kravet om, at hele forsøgssituationen i videst muligt omfang bør minde om en analog dagligdagssituation uden for laboratoriet, ikke blot forekommer en kende for restriktivt, men samtidig indebærer en risiko for, at visse potentielt indsigtgivende design på forhånd diskvalificeres. En mere afbalanceret udlægning af kravene til forsøgssituationen synes derfor at være, at det forekommer tilstrækkeligt, hvis blot de *aspekter* af forsøgssituationen, der direkte er relevante for undersøgelsesgenstanden, præsenteres på en måde, så de i videst muligt omfang minder om den analoge dagligdagssituation uden for laboratoriet. Men selv med denne præcisering mangler vi stadig kriterier for, hvorledes vi afgør, hvad der er de »direkte relevante aspekter af en forsøgssituation«.

For det tredje skulle man fokusere på, hvorvidt forsøgspersonens naturlige responser reelt var mulige i forsøgssituationen (jf. punkt 3 i Schmucklers systematik). I den forbindelse blev fremdraget to forskellige eksempler: I det første eksempel risikerede børnenes naturlige »lede-adfærd« som en følge af den restriktive afgrænsning af scoringsfeltet at blive scoret som manglende interesse fra børnenes side (Krøjgaard, *in prep.*). Dette potentielle problem kunne ansues som et udtryk for, at børnenes naturlige responser ikke var blevet tilgodeset i forsøgssituationen. Man bør dog lægge mærke til, at denne restriktive afgrænsning af scoringsfeltet gav anledning til en asymmetrisk og »konservativ« forvriddning af resultaterne, der alt andet lige gjorde det vanskeligere at opnå positive resultater. I forlængelse heraf er det værd at bemærke, at man selv med denne restriktive strategi faktisk har opnået konsistente positive resultater i forskellige laboratorier (f.eks. Xu & Carey, 1996; Wilcox & Baillargeon, 1998a; 1998b; Wynn, 1992). Problemet ville selvsagt have været af en mere alvorlig karakter, hvis begrænsningen af børnenes naturlige responser havde været *faciliterende* for opnåelsen af positive resultater. Alligevel bør man naturligvis generelt være varsom med at afskære forsøgspersonerne fra at udføre deres naturlige responser.

I det andet eksempel, der blev præsenteret i denne forbindelse (*visual cliff*-forsøget), havde man – i modsætning til det ovenfor anførte eksempel – formået at tilrettelægge designet således, at børnenes naturlige responser (først og fremmest at kunne kravle relativt frit rundt) syntes bibeholdt trods det, at der var tale om et laboratorieforsøg. Disse eksempler giver anledning til flere overvejelser: At forsøgspersonerne kan være forhindret i at udføre deres naturlige responser på en sådan måde, at det er potentielt problematisk; at et sådant design alligevel kan være anvendeligt, hvis børnenes responser er så stærke, at de trods restriktive betingelser manifesterer sig signifikant i resultaterne; at det undertiden er muligt at designe eksperimentelle undersøgelser med spædbørn *uden*, at deres naturlige eksplorative adfærd begrænses.

Jeg argumenterede endvidere for, at Bronfenbrenners insisteren på, at der skulle være korrespondens mellem dét forsøgspersonen oplever og dét, som forsøgslederen havde intenderet, at forsøgspersonen skulle opleve, var en vigtig pointe, som *ikke* nødvendigvis blev begrebet af de tre anvisninger, som Schmuckler havde opregnet. Til illustration heraf inddrog jeg to eksempler, som ikke umiddelbart forbrød sig mod de tre tidligere anførte forskrifter, men hvor det alligevel i begge tilfælde syntes tvivlsomt, om forsøgspersonerne reelt oplevede dét, som var intenderet af forsøgslederen. I det første eksempel fra min egen forskning fik jeg via børnenes reaktioner (hvor de kiggede efter skærmen og ikke efter objekterne) tydeligt besked om, at de oplevede begivenheden anderledes, end den var intenderet fra min side (Krøjgaard, 1999). Var dette problematiske forhold ikke blevet opdaget, ville det selvsagt have haft negative konsekvenser for generaliserbarheden af resultaterne og dermed for den økologiske validitet. I det andet eksempel fra Wynns studier (Wynn, 1992) var det efter min vurdering derimod reelt ikke muligt inden for det enkelte forsøgs logik at afgøre, hvorvidt børnene opfattede testbegivenhederne, som de var intenderet af Wynn. Hvorvidt Wynns forsøg mangler økologisk validitet ud fra Bronfenbrenners kriterier synes således ikke muligt at afklare endeligt.

Det er min opfattelse, at de potentielle problemer med den økologiske validitet i Bronfenbrenners forstand i Wynns forsøg er et godt eksempel på, hvorledes begrebet økologisk validitet undertiden manifesterer sig *implicit* i eksperimentel spædbarnsforskning: Som tidligere anført, kan det diskuteres, hvorvidt der reelt er et problem i Wynns fortolkning. Men hvis vi for en stund tillader os at antage, at der *er* et problem, så lad os se på, hvornår problemet opstår. Problemet kan formentlig spores tilbage til selve *operationaliseringen* af det stillede spørgsmål. Wynn har, som følge af den problemstilling hun ville undersøge, ønsket at præsentere spædbørnene for et simpelt regnestykke, og da det pga. børnenes alder ikke gav mening at præsentere regnestykkerne med symboler, var det nærliggende at gøre brug af en analog fremstilling. Men netop denne operationalisering indebar, at børnenes opfattelse af objekters bevægelser i tid og rum uundgåeligt blev

medtaget i designet som en konfunderende variabel.⁵ Og så fanger bordet så at sige, da fejl i operationaliseringsfasen også vil manifestere sig, når de tilvejebragte resultater skal fortolkes. Pointen er altså den banale, at laboratorieforsøg, der rummer andre plausible fortolkninger end de, der indgår i hypoteserne, dels har medtaget konfunderende arvegods fra operationaliseringsfasen og dels, at tilstedeværelsen af alternative fortolkninger kan anskues som en direkte manifestation af risikoen for manglende økologisk validitet i Bronfenbrenners forstand.

Begrebet reduktionisme vækker ofte bekymrede miner inden for psykologien. Der findes imidlertid mindst én form for reduktionisme, som aldrig burde give anledning til bekymring, men som derimod er en nødvendighed, og det er den *metodiske reduktionisme*. Enhver tænkelig undersøgelse indebærer nødvendigvis, at noget gøres til figur, og at noget andet gøres til baggrund (Køppe, 1990). Dette karakteristikum kendetegner også vores dagligdags erkendelse og er således ikke blot noget, der alene er forbeholdt den videnskabelige praksis. I laboratorieforsøg strækkes dette karakteristikum til sin yderste grænse med de fordele og ulemper dette indebærer. Laboratorieforsøg er ikke attråværdige, *fordi* de er laboratorieforsøg. Laboratorieforsøg er derimod nødvendige og vigtige, da man kun under sådanne betingelser kan opnå den grad af kontrol, der skal til for at kunne sige noget om kausale årsagsrelationer mellem variable. Styrken ved laboratorieforsøg er således åbenbar. Det er svagheden imidlertid også. Den består bl.a. i risikoen for, at man i bestræbelserne på at isolere variable, skaber en forsøgssituation, hvor et eller flere relevante aspekter af den situation, man gerne ville undersøge, er blevet forandret eller måske ligefrem elimineret på en sådan måde, at forsøgspersonen ikke helt oplever forsøgssituationen sådan, som den af forsøgslederen var tiltænkt. I bestræbelserne på at minimere denne potentielle svaghed ved undersøgelser generelt og laboratorieforsøg i særdeleshed er begrebet økologisk validitet et både nødvendigt og vigtigt begreb.

Den økologiske tilgangs insisteren på, at subjektet er indlejret og aktivt handlende i en økologisk niche, og at denne forståelse bør tjene som udgangspunkt for videnskabelige undersøgelser, er i den forbindelse en vigtig påpegning. Imidlertid kan en sådan anvisning ikke stå alene (Lewkovicz, *in press*). Det er derfor nødvendigt at specificere de kriterier, der skal opfyldes, for at undersøgelser kan siges at være økologisk valide. Ellers risikerer begrebet økologisk validitet at forblive en relativt ufrugtbar og distanceret standardafvisning af laboratorieforsøg. De inddragede empiriske eksempler skulle gerne have illustreret, at det imidlertid ikke er nogen

5 Det bør retfærdigvis nævnes, at i og med, at de simple additions- og subtraktionsstykker, når de fremstilles i analog form, ikke kan adskilles fra at korrespondere med objekters bevægelser i tid og rum, kunne man naturligvis lige såvel hævde, at det var Xu & Carey (1996) eller undertegnede, der havde lavet konfunderede design.

simpel opgave at specificere sådanne kriterier. For samtlige de tre konkretiseringer, som blev opregnet af Schmuckler, var det muligt at anføre eksempler på undersøgelser, som var *vellykkede* trods det, at de faktisk forbrød sig mod én (eller flere) af de tre anvisninger. Vi synes derfor at stå tilbage med det aktuelle problem, at det på den ene side er nødvendigt at specificere begrebet økologisk validitet, hvis begrebet skal være anvendeligt, men at vi samtidig med de hidtidige forsøg på konkretiseringer på forhånd risikerer at diskvalificere potentielt frugtbare undersøgelser. Der synes således at foreligge et betragteligt stykke arbejde med at finpudse konkretiseringer af begrebet økologisk validitet. En løsning på disse begrebsmæssige problemer vil formentlig ikke blot være af betydning for den økologiske psykologi og den eksperimentelle spædbarnsforskning, men for alle grene af psykologien, der gør brug af empiriske undersøgelser.

REFERENCER

- BAILLARGEON, R. (1986): Representing the existence and the location of hidden objects: Object permanence in 6- and 8-month-old infants. *Cognition*, 23, 21-41.
- BAILLARGEON, R. (1993): The object concept revisited: New directions in the investigation of infants' physical knowledge. In C.E. Granrud (Ed.), *Visual perception and cognition in infancy: Carnegie-Mellon symposia on cognition*, vol. 23 (pp. 265-315). Hillsdale, NJ: Erlbaum.
- BAILLARGEON, R. (1995): A model of physical reasoning in infancy. In C. Rovee-Collier and L. Lipsitt (Eds.), *Advances in Infancy Research*, vol. 9 (pp. 305-371). Norwood, N.J.: Ablex.
- BAILLARGEON, R. (1998): Infants' understanding of the physical world. In M. Sabourin, F. Craik, & M. Robert (Eds.), *Advances in Psychological Science*, Vol. 2 (pp. 503-529). London: Psychology Press.
- BAILLARGEON, R., & GRABER, M. (1988): Evidence of location memory in 8-month-old infants in a nonsearch AB task. *Developmental Psychology*, 24, 502-511.
- BAILLARGEON, R., SPELKE, E.S., & WASSERMAN, S. (1985): Object permanence in five month old infants. *Cognition*, 20, 191-208.
- BERNTSEN, D. (1996): Involuntary autobiographical memories. *Applied Cognitive Psychology*, 10, 435-454.
- BERNTSEN, D. (1998): Voluntary and involuntary access to autobiographical memory. *Memory*, 6, 113-141.
- BERNTSEN, D. (*in press*): Involuntary memories of emotional events: Do memories of traumas and extremely events differ? *Applied Cognitive Psychology*.
- BRONFENBRENNER, U. (1977): Toward an experimental ecology of human development. *American Psychologist*, 32, 513-531.
- BRONFENBRENNER, U. (1979): *The ecology of human development*. Cambridge, Massachusetts: Harvard University Press.
- BRONFENBRENNER, U., & MORRIS, P.A. (1998): The ecology of developmental processes. In W. Damon (Series Ed.) & R.M. Lerner (Vol. Ed.), *Handbook of Child Psychology*, Vol. 1. (pp. 993-1028). New York: Wiley.
- BRUNSWIK, E. (1956): *Perception and the representative design of psychological experiments*. Berkeley and Los Angeles: University of California Press.
- COOLICAN, H. (1999): *Research methods and statistics in psychology*. London: Hodder & Stoughton.

- CORSINI, R.J. (Ed.) (1994): *Encyclopedia of psychology*. New York: Wiley & Sons.
- DIAMOND, A. (1988): Abilities and neural mechanisms underlying AB performance. *Child Development*, 59, 523-527.
- DIAMOND, A. (1991): Neuropsychological insights into the meaning of object concept development. In S. Carey, & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 67-110). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- FABES, R.A., MARTIN, C.L., HANISH, L.D., & UPDEGRAFF, K.A. (2000): Criteria for evaluating the significance of developmental research in the 21st Century: Force and counterforce. *Child Development*, 71, 212-221.
- GARDNER, H. (1987): *The mind's new science: A history of the cognitive revolution*. (2. ed.) New York: Basic Books.
- GIBSON, J.J. (1966): *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
- GIBSON, J.J. (1972): A theory of direct visual perception. In J.R. Royce, & Wm.W. Rozeboom (Eds.), *The psychology of knowing* (pp. 215-225). NY: Gordon and Breach.
- GIBSON, J.J. (1979): *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- GIBSON, E.J., & WALK, R.D. (1960): The »visual cliff«. *Scientific American*, 202, 64-71.
- GRATCH, G., APPEL, K.J., EVANS, W.F., LECOMPTE, G.K., & WRIGHT, N.H. (1974): Piaget's Stage IV object concept error: Evidence of forgetting or object conception? *Child Development*, 45, 71-77.
- HAITH, M.M. (1998): Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior and Development*, 21, 167-179.
- HAITH, M.M., & BENSON, J.B. (1998): Infant cognition. In W. Damon (Series Ed.) & D. Kuhn & R. Siegler (Eds.), *Handbook of Child Psychology*, Vol. 2. (pp. 199-254). New York: Wiley.
- HARRIS, P.L. (1973): Perseverative errors in search by young infants. *Child Development*, 44, 28-33.
- KRØJGAARD, P. (1998): Forskningen vedrørende objektpermanens: Et review af de undersøgelser, der anvender velkendte objekter. *Nordisk Psykologi*, 50, 258-279.
- KRØJGAARD, P. (1999): *Spædbørns erkendelse af objekters numeriske identitet*. Ph.D.-afhandling. Psykologisk Institut, Aarhus Universitet.
- KRØJGAARD, P. (2000): Object individuation in 10-month-old infants: Do significant objects make a difference? *Cognitive Development*, 15, 169-184.
- KRØJGAARD, P. (*in prep.*): Reasoning about object individuation in 8- and 10-month-old infants: The role of familiar objects and amount of introduction. Manuscript in preparation.
- KÖPPE, S. (1990): *Virkelighedens niveauer*. København: Nordisk Forlag.
- LARSEN, S.F., & CONWAY, M.A. (1997): Reconstructing dates of true and false autobiographical memories. *European Journal of Cognitive Psychology*, 9, 259-272.
- LARSEN, S.F., & THOMPSON, C.P. (1995): Reconstructive memory in the dating of personal and public news events. *Memory & Cognition*, 23, 780-790.
- LEWKOWICZ, D.J. (*in press*): The concept of ecological validity: What are its limitations and is it bad to be invalid? *Infancy*.
- MAMMEN, J. (1993): The Elements of Psychology. In N. Engelsted, M. Hedegaard, & A. Mortensen (Eds.), *The Societal Subject* (pp. 29-44). Aarhus: Aarhus University Press.
- McCALL, R.B. (1977): Challenges to a science of developmental psychology. *Child Development*, 48, 333-344.

- MILGRAM, S. (1975): *Obedience to authority*. New York: Harper Colophon.
- MURRAY, L., & TREVARTHEN, C. (1985): Emotional regulation of interactions between two-month-olds and their mothers. In T.M. Field, & N.A. Fox (Eds.), *Social perception in infants* (pp. 177-197). Norwood, NJ: Ablex.
- NEISSER, U. (1967): *Cognitive psychology*. New York: Appleton Century Crofts.
- NEISSER, U. (1976): *Cognition and reality*. San Francisco: Freeman.
- NEISSER, U. (1985): Toward an ecologically oriented cognitive science. In T.M. Shlechter, & P.M. Togliola (Eds.), *New directions in cognitive science* (pp. 17-32). Norwood, NJ: Ablex.
- PIAGET, J. (1954): *The construction of reality in the child*. New York: Basic Books. (Fransk udgave 1937).
- SIMON, T.J., HESPOS, S.J., & ROCHAT, P. (1995): Do infants understand simple arithmetic? A replication of Wynn. *Cognitive Development*, 10, 253-269.
- SPELKE, E.S. (1985): Preferential looking methods as tools for the study of cognition in infancy. In G. Gottlieb, & N. Krasnegor (Eds.), *Measurement of audition and vision in the first year of life* (pp. 323-363). Norwood, NJ: Ablex.
- SPELKE, E.S., BREINLINGER, K., MACOMBER, J., & JACOBSEN, K. (1992): Origins of knowledge. *Psychological Review*, 99, 605-32.
- SPELKE, E.S., KESTENBAUM, R., SIMONS, J.D., & WEIN, D. (1995): Spatio-temporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology*, 13, 113-142.
- STILLINGS, N.A., FEINSTEIN, M.H., GARFIELD, J.L., RISSLAND, E.L., ROSENBAUM, D.A., WEISLER, S.E., BAKER-WARD, L. (1987): *Cognitive science: An introduction*. Cambridge, MA: Bradford/MIT Press.
- WELLMAN, H.M., CROSS, D., & BARTSCH, K. (1987): Infant search and object permanence: A meta-analysis of the A not B error. *Monographs of the Society for Research in Child Development*, 51, 1-51.
- WILCOX, T., & BAILLARGEON, R. (1998a): Object individuation in infancy: The use of featural information in reasoning about occlusion events. *Cognitive Psychology*, 37, 97-155.
- WILCOX, T., & BAILLARGEON, R. (1998b): Object individuation in young infants: Further evidence with an event-monitoring task. *Developmental Science*, 1, 127-142.
- WYNN, K. (1992): Addition and subtraction by human infants. *Nature*, 358, 749-750.
- WYNN, K. (1995): Infants possess a system of numerical knowledge. *Current Directions in Psychological Science*, 4, 172-177.
- WYNN, K., & CHIANG, W. (1998): Limits to infants' knowledge of objects: The case of magical appearance. *Psychological Science*, 9, 448-455.
- XU, F., & CAREY, S. (1996): Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, 30, 111-153.