

# Ekstinksjon etter partiell forsterkning: Hvilken rolle har verbalisering?

Frode Svartdal og Ivan Nygreen

Universitetet i Tromsø

Det er veldokumentert at partiell eller intermitterende forsterkning (forsterkning en gang i blant) gir en mer motstandsdyktig respons sammenlignet med kontinuerlig forsterkning. Dette påvises klart ved at uavhengige grupper av forsøksdeltakere først eksponeres for partiell vs. kontinuerlig forsterkning. I neste fase opphører så forsterkning (ekstinksjon). I slike situasjoner vil prestasjon etter partiell forsterkning gi en sterk respons under ekstinksjonsbetingelser, mens prestasjon etter kontinuerlig forsterkning gir en svakere ekstinksjonsrespons. Dette funnet omtales ofte som "partial reinforcement extinction effect" (PREE), og har vært godt etablert siden 1940-årene (Humphreys, 1939; Skinner, 1938).

En tidligere artikkel i *Diskriminanten* (Svartdal, 2001) har påvist at PREE likevel ikke er en så klar effekt som mange har trodd. Kort oppsummert er det i alle fall tre forhold som gjør at vi i dag er mer i tvil enn vi var for bare noen få år siden:

- PREE er ikke alltid resultatet når kontinuerlig vs. partiell forsterkning sammenlignes. For eksempel viste Pavlik og Carlton (1965) at når en og samme organisme eksponeres for begge typer forsterkningsbetingelser, kan resultatet være en omvendt effekt, en *reversert* PREE. Andre (Amsel, 1994) har vist at slike betingelser også kan resultere i en generelt forhøyet responsstyrke; dette kalles *generalisert* PREE.
- Det er uklarerhet rundt hvordan prestasjon under ekstinksjon skal måles. De to vanlige målene er *hvor raskt responsen avtar*, eller *hvor mange responser* som avgis under ekstinksjon. I noen situasjoner kan vi observere PREE, ingen effekt, eller en reversert PREE, alt etter hvilket av disse målene vi velger.
- Det er også uklarerhet om hvordan responsenheten skal avgrenses. Ved kontinuerlig forsterkning vil responsenheten per definisjon være den enkelte respons. Men ved partiell forsterkning får vi et problem. Hva skal vi velge som enheten, den enkelte respons, eller den funksjonelle respons (det antall responser som ble krevd for forsterkning i læringsfasen)?

Det er særlig de to første momentene her som gjør det vanskelig å trekke entydige konklusjoner fra empiriske data. For eksempel legger Nevin (1988) stor

vekt på målemetoden, og hevder at mange tilfeller av PREE egentlig er et resultat av feilanalyserte data.

Uklarheter som dette har selvsagt skapt mye usikkerhet rundt PREE. Denne usikkerheten skal imidlertid ikke overdrives. Det er ikke tvil om at PREE er en meget robust effekt under bestemte betingelser. Hvis vi manipulerer forsterkerrate for forskjellige grupper av personer (eller dyr), der en gruppe får forsterkere etter hver eneste ønskede respons (100 % forsterkning), mens en annen gruppe får forsterkere bare av og til når ønsket respons inntreffer (partiell forsterkning), ses PREE når ekstinksjon iverksettes.

Det faktum at PREE er en udiskutabel effekt under disse betingelsene, gjør at effekten kan fungere som et grunnlag for videre undersøkelser og teoridannelse. For eksempel er de viktigste teoriene om hvorfor PREE oppstår (se Amsel, 1994; Capaldi, 1994) forankret i denne godt etablerte empiriske effekten. Et annet eksempel, som er temaet for denne artikkelen, er våre undersøkelser av forholdet mellom ekstinksjonsatferd og personens verbalisering. Vi tar utgangspunkt i at PREE er en udiskutabel effekt, og spør: Gitt at personen viser PREE i atferd, vil denne effekten også vises i personens verbalisering?

## PREE og verbalisering

Det foreligger en stor forskningslitteratur på hvordan verbalisering påvirker tilpasning til *læringsbetingelser* (se for eksempel Svartdal, 1991), men forskning på hvordan verbalisering påvirker tilpasning til *ekstinksjonsbetingelser* er nærmest totalt fraværende. Dette er ganske merkelig. Vi vet at personens verbalisering er av stor betydning for tilegnelse av atferd under forsterkningsbetingelser; i noen tilfeller vil verbalisering dominere over ikke-verbale kontrollbetingelser. Noen har til og med hevdet at verbalisering er en forutsetning for tilpasning til læringskontingenser (se for eksempel Shanks & St John, 1994). I lys av dette skulle man tro at verbalisering kunne være av tilsvarende stor betydning for reduksjon av atferd under ekstinksjonsbetingelser. I dag er dette imidlertid bare en løs antakelse. Foruten vår forskning i Tromsø (som er temaet for denne artikkelen) er det kun gjort én undersøkelse som kan belyse sistnevnte spørsmål, og den ble utført i 1939 (Humphreys, 1939b).

Generelt sett kan to spørsmål formuleres for studier som ser på forholdet mellom atferd og verbalisering under ekstinksjonsbetingelser. (a) For det første kan man spørre om en atferdsmessig målt PREE ledsages av verbalisering. Sett at vi observerer PREE i atferd hos voksne forsøkspersoner. Vil det da være slik at denne atferdsmessige PREE også kan måles i personens verbalisering? (b) For det andre kan man undersøke spørsmålet motsatt vei: Vil beskrivelse av forsterkningsbetingelser gi PREE-lignende atferd under ekstinksjonsbetingelser? I prinsippet er dette samme spørsmål forskere innenfor operant psykologi har undersøkt i årevis når de har spurt om hvordan regler styrer atferd (se for eksempel Svartdal, 1992). Forskjellen er at vi i denne sammenheng ikke spør om hvordan regler som beskriver forsterkningsbetingelser styrer atferd når betingelsene er operative, men hvordan reglene styrer atferd etter at disse forsterkningsbetingelsene har virket *men nå er opphørt* (ekstinksjonsbetingelser). Vi har undersøkt begge disse spørsmålene.

## Ledsages PREE av verbalisering?

Hvis man undersøker litteraturen på PREE, er det åpenbart at et stort flertall av studiene av ekstinksjon er gjennomført på dyr. De studiene som er gjennomført på mennesker, er utført på samme måte, dvs. med atferdsmål. Konkret betyr dette at når PREE observeres, så er det en *atferdsmessig* PREE vi snakker om. Rotters *løpehastighet*, antall *rykkeresponser* per tidsenhet, og *reduksjon i responsrate* er typiske mål. Men hvis PREE er en så generell og robust effekt som forskningen tilsier, så skulle vi forvente at slike responsmål bare er en av flere mulige måter å uttrykke effekten på. En lang rekke studier av forholdet mellom verbalisering og atferd har påvist at atferdsmål korrelerer godt med alternative mål i innlæringsfasen. For eksempel har flere vist at hvis man ber forsøkspersoner gi *bedømmelser* av forsterkningsbetingelsene under læring, så samsvarer slike bedømmelser meget godt med atferdsmål (se for eksempel Chatlosh, Neunaber, & Wasserman, 1985; Shanks & Dickinson, 1991). Men hva med tilsvarende bedømmelser under ekstinksjon?

Humphreys (1939b) var, så vidt vi vet, den første og eneste som har undersøkt forsøkspersonenes verbalisering under ekstinksjonsbetingelser. Han brukte en klassisk betingingsprosedyre, og oppgaven for forsøkspersonen var ganske enkelt å predikere om US kom, gitt at BS hadde blitt presentert. Hver gang BS ble presentert, skulle forsøkspersonen skrive ned om US ville inntreffe eller ikke. I innlæringsfasen (24 trials) ble to forskjellige grupper gitt forskjellige læringsbetingelser. I en gruppe kom US etter BS hver eneste gang (100 % forsterkning), i en annen gruppe kom US etter BS i halvparten av alle trials (50 % forsterkning). Det interessante nå hva som skjedde under ekstinksjon. Humphreys fant at forsøkspersonenes forventning, målt på basis av gjetninger om US ville komme, var ”sterkere” (mer persistent) i 50 %-gruppen.

Humphreys' funn var altså helt i tråd med en konvensjonell PREE. Merk at effekten ble målt ved forsøkspersonenes gjetninger (prediksjoner), noe som er helt uvanlig i forskningslitteraturen innenfor ekstinksjon. Hvorvidt et mål som dette er like godt som (eller kanskje bedre enn) atferdsmål vet vi imidlertid lite om, ganske enkelt fordi man ikke har viet dette spørsmålet noen som helst interesse.

I en serie av eksperimenter (Svartdal, 2000b) angrep vi spørsmålet indirekte: La oss anta at folk flest har rikelig erfaring med ulike typer forsterkningsbetingelser, og at vi på basis av denne erfaringen etablerer en slags ”generalisert erfaring” som tilsier følgende: ”Hvis ønskede utfall inntreffer bare av og til når jeg gjør X, er det smart å stå på litt ekstra hvis utfallene uteblir.” En annen generalisert erfaring kunne være følgende: ”Hvis ønskede utfall inntreffer hver gang jeg gjør X, og X plutselig skulle utebli selv om jeg gjør X, da er det liten vits i å fortsette.” Merk at vi her ikke antar at folk flest verbaliserer disse reglene; vi bare antar at disse to typene erfaring sannsynligvis er forbundet med to forskjellige strategier. (Bruk dette på et tenkt eksempel. Hvis brusautomaten alltid er ”kranglete”, blir du vant til å jobbe litt ekstra for å få ut flasken du er ute etter. Hvis nå maskinen så slutter å fungere, vil du sannsynligvis ikke gi deg så lett. Omvendt: Hvis brusautomaten alltid fungerer på første forsøk, og den plutselig en dag skulle slutte å fungere, vil du sannsynligvis gi opp relativt raskt.)

Vår antakelse var at forsøkspersoner gjennom et langt liv ville ha etablert en slik generalisert kunnskap om slike situasjoner, og at dette ville komme til uttrykk i våre studier. Vi presenterte derfor små scenarier for forsøkspersoner, der informasjon om belønning var manipulert på en slik måte som typisk gir PREE. En gruppe kunne eksempelvis få presentert følgende historie:

”En familie har i lengre tid pleid å gi hunden sin mat ved middagsbordet. Ved hvert middagsmåltid har hunden kommet til bordet og tippet om mat, og den har fått en matbit hver eneste gang den har tippet. Vi kan altså anta at hunden har lært at *hvert forsøk* gir ønsket utfall (mat).

Nå har imidlertid familien satt seg fore å få hunden til å slutte å tigge ved bordet. De har derfor bestemt seg for ikke å gi hunden mat når den tigger ved middagsbordet. Bestemmelsen gjennomføres konsekvent: Hunden får nå *aldri* mat når den tigger ved middagsbordet (den får selvsagt mat i egen matskål på normal måte).”

Merk at hunden i dette scenarioet angivelig har fått mat hver gang den har tippet. En annen gruppe fikk annen informasjon om hundens matvaner: ”... den har fått en matbit av og til når den har tippet. Vi kan altså anta at hunden har lært at *noen men ikke alle forsøk* gir ønsket utfall (mat).” Her har vi altså informert en gruppe om kontinuerlig forsterkning for mas, en annen om partiell forsterkning.

Ifølge PREE skal nå folk, hvis de har en generalisert kunnskap om effekten av forsterkningsbetingelser, svare forskjellig når de så presenteres for følgende spørsmål: ”Hvor lenge tror du hunden vil fortsette å tigge ved middagsbordet før den slutter?” Vi lot forsøkspersonene svare på følgende måte:

Angi svarene ved antall ganger du tror hunden vil tigge i hver av de ti ukene etter at familien har bestemt seg for ikke å gi mat når hunden tigger. Sett ring rundt svarene (minimum 0, maksimum 7 ganger):

1. uke: 0 1 2 3 4 5 6 7 ganger
2. uke: 0 1 2 3 4 5 6 7 ganger
3. uke: 0 1 2 3 4 5 6 7 ganger
4. uke: 0 1 2 3 4 5 6 7 ganger
5. uke: 0 1 2 3 4 5 6 7 ganger
6. uke: 0 1 2 3 4 5 6 7 ganger
7. uke: 0 1 2 3 4 5 6 7 ganger
8. uke: 0 1 2 3 4 5 6 7 ganger
9. uke: 0 1 2 3 4 5 6 7 ganger
10. uke: 0 1 2 3 4 5 6 7 ganger

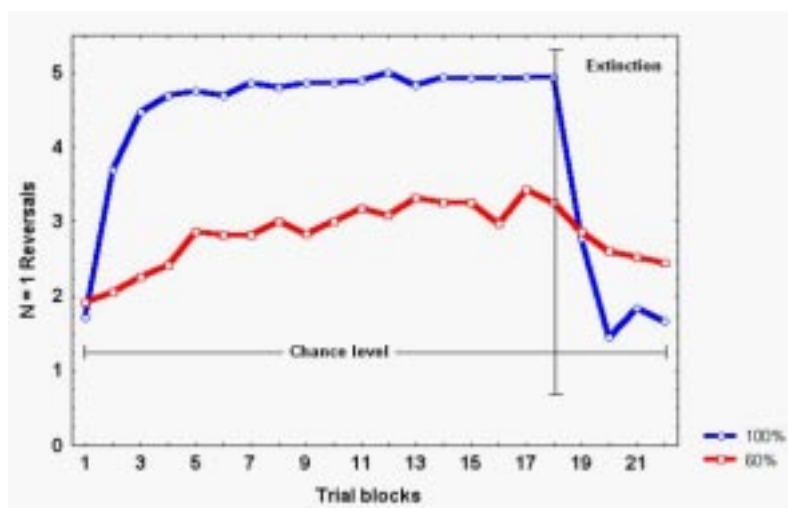
For å gjøre en lang historie (Svartdal, 2000b) kort, så viste det seg at folk her svarte stort sett det samme, uavhengig av hvilken informasjon vi ga i scenarioet. Sagt på en annen måte: Folk tar ikke hensyn til informasjon om hvordan en (u)vane er etablert når de skal gi en bedømmelse av hvor motstandsdyktig den er overfor ekstinksjonsbetingelser. Vi stilte flere hundre forsøkspersoner overfor denne

og tilsvarende bedømmelsesoppgaver, og vi spurte på alternative måter. Uansett klarte vi ikke å påvise noen form for sensitivitet i bedømmelser til relevant informasjon om forsterkningsbetingelser. Kun i to tilfeller påviste vi slik sensitivitet, nemlig hos folk med ekspertkunnskap (grunnfagsstudenter som nettopp hadde lest om PREE, og vernepleiere ved Akershus vernepleierhøgskole). Så selv om disse ekspertene gjennom våre oppgaver fikk vist at de kunne sine ting, var konklusjonen for folk flest negativ: Vi har generelt sett ingen generalisert kunnskap om effektene av forsterkningsbetingelser på ekstinksjon.

”Ja, og så?”, vil kanskje mange nå si. Vi andre vil si ”så interessant!”. Vi vet fra annen forskning at generalisert kunnskap om læringsbetingelser er viktig for tilpasning til nye situasjoner. Harlows (1949) studier av *learning set* er et velkjent eksempel. Harlow viste at økende erfaring med en type lærings situasjon (diskriminasjonslæring) førte til at aper lærte på mer effektive måter; de ”lærte seg å lære”. I nyere tid har forskning på ”metakognisjon”, ”metahukommelse” og ”metalæring” kastet lys over hvordan vi etablerer ulike former for *overordnet* kunnskap, metakunnskap om den erfaringen vi gjør. Bedømmelsesstudiene vi har diskutert viser så langt at folk ikke har slik overordnet kunnskap når det gjelder effektene av ulike forsterkningsbetingelser på ekstinksjon.

Dette funnet fikk oss til å stille et annet og mer dristig spørsmål: Kunne det tenkes at fraværet av metakunnskap i disse bedømmelsesstudiene innebærer at det foreligger en dissosiasjon (spaltning) mellom atferd og verbalisering når det gjelder ekstinksjon? Med andre ord: Kan det tenkes at ekstinksjonseffekter først og fremst er atferdsmessige effekter, og at verbalisering ikke ”henger med”?

For å klarlegge dette spørsmålet, var det nødvendig angripe saken mer direkte. Den enkleste og mest direkte måten spørsmålet kan belyses på, er simpelt hen å sette forsøkspersonen i en lærings situasjon, la personen lære en ferdighet, og så – umiddelbart etter læringsfasen – undersøke hvordan personen vil vurdere sin egen tilbøyelighet til å utføre den lærte responsen under ekstinksjonsbetingelser. Hvis to



Figur 1. Konvensjonell atferdsmessig PREE

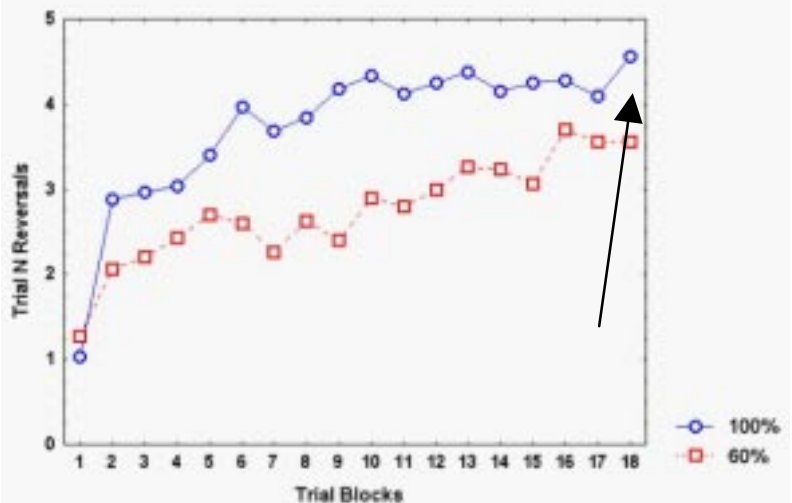
forskjellige grupper av forsøkspersoner lærer ferdigheten under 100 % forsterkning vs. for eksempel 60 % forsterkning, skulle vi – i tråd med PREE – forvente at den første gruppen ville vurdere sin egen tilbøyelighet til å fortsette som ganske svak, mens den siste gruppen skulle vurdere tilbøyeligheten som sterkere. Vi har mengder av data som viser av disse betingelsene fører til en *atferdsmessig* PREE (Svartdal, 2000a; se Figur 1); spørsmålet nå er om personenes verbalisering henger med.

Tabell 2 og Figur 2 viser oppsettet for et typisk eksperiment (se Svartdal, 2001). To uavhengige grupper av forsøkspersoner lærer en relativ kompleks ferdighet med henholdsvis 100 % og 60 % forsterkning for korrekt respons. Umiddelbart etter tilegnelsesfasen avbrytes eksperimentet, og forsøkspersonen blir stilt spørsmål for å måle persistens. Konkret ber vi forsøkspersonen her forutsi (predikere) hvor mange ganger hun/han vil gjenta den lærte responsen i fravær av forsterker før hun/han gir opp og prøver noe annet.

Vi har gjennomført flere runder av slike eksperimenter, og målt verbalisert persistens på ulike måter. Alle studiene viser samme utfall: Når forsøkspersoner skal forutsi hvor mange ganger de vil gjenta den lærte responsen, har læringshistorien de nettopp har lagt bak seg ingen effekt på prediksjonene (se Svartdal, 2002). Svartdal og Silvera (2002) malte prediksjon på en alternativ måte, men heller ikke

Tabell 2. Oppsett for eksperiment for å undersøke verbalisering av PREE

	Læringsfase	Prediksjon
Gruppe CRF	100 % forsterkning	"Hvor mange ganger vil du gjenta responsen du har lært i fravær av forsterker?"
Gruppe PRF	60 % forsterkning	"Hvor mange ganger vil du gjenta responsen du har lært i fravær av forsterker?"

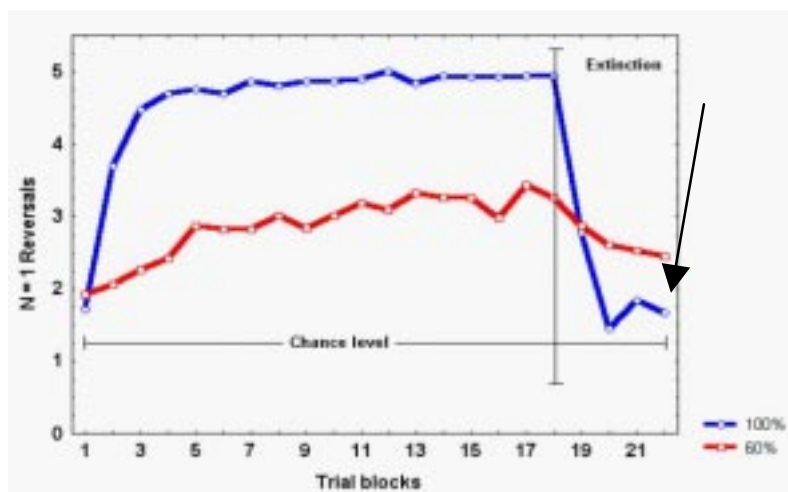


Figur 2. Læringsfase under to betingelser, 60% og 100% forsterkerrate. Pilen indikerer tidspunkt for når forsøkspersonene ble bedt om å predikere egen persistens

her hadde læringshistorien noen som helst effekt. I et av eksperimentene i sistnevnte artikkel lot vi forsøkspersonene få begrenset erfaring med ekstinksjonsbetingelsene ved at de etter læringsfasen fikk 5 ekstinksjonstrials. For gruppen med kontinuerlig belønning i læringsfasen skulle dette øke sannsynligheten for verbalisering ved at overgangen fra læring til ekstinksjon her er nokså åpenbar. Analyser av atferdsdata viste da også at forsøkspersonene i denne gruppen viste *atferdsmessigs*sensitivitet til ekstinksjonsbetingelsene – responsraten sank umiddelbart etter at ekstinksjon ble innført. Men prediksjonsdata ble ikke påvirket. Som i våre tidligere eksperimenter var forsøkspersonenes prediksjoner helt upåvirket av forsterkningsbetingelsene i læringsfasen.

Funnet er dermed nokså klart: Forsøkspersoner er under disse betingelsene ikke i stand til å si hva de kommer til å gjøre om noen få sekunder! Dette står i åpenbar kontrast til atferdsdata (se Figur 1), som viser en klar effekt av læringshistorien på ekstinksjonsprestasjon.

Nå kunne man innvende at forsøkspersoner i denne situasjonen ikke kan predikere sin egen persistens fordi de ikke kan beskrive hva de gjør i situasjonen. Vi har imidlertid data som viser at våre forsøkspersoner er meget dyktige til å beskrive ikke bare hva de gjør i situasjonen, men også hvordan læringsbetingelsene er. I et eksperiment (Svartdal, 2002, Experiment 4) lot vi to grupper av forsøkspersoner først gjennomgå læringsfasen, deretter en ekstinksjonsfase. Atferdsdata, som viser en konvensjonell PREE, ses i Figur 3. Etter at ekstinksjonsfasen var gjennomført, ba vi så forsøkspersonene bedømme egen atferd under så vel læring som ekstinksjon, og vi ba dem bedømme forsterkerrate under læringsfasen. Alle disse målene viste seg å samsvare meget godt med de faktiske antall responser og forsterkere. Med andre ord: Forsøkspersonene var meget dyktige til å beskrive alle relevante sider ved eksperimentsituasjonen. Av størst interesse her er at de beskrev egen persistensatferd helt korrekt, helt i tråd med atferdsdata vist i Figur 3.



Figur 3. Læringsfase og ekstinksjonsfase under to betingelser, 60% og 100% forsterkerrate. Pilen indikerer tidspunkt for når forsøkspersonene ble bedt om å bedømme egen persistens

Disse funnene kan oppsummeres slik: Når forsøkspersoner etter endt læringsfase blir bedt om å *predikere* egen persistens, skjer dette helt upåvirket av de læringsbetingelsene de nettopp er blitt eksponert for. Men når en annen gruppe forsøkspersoner, som gjennomgår så vel læringsfase som ekstinksjonsfase, blir bedt om å *bedømme* egen persistens under læring, er deres bedømmelser meget nøyaktige. Av dette kan man slutte at persistensatferd ikke verbaliseres før etter at den er vist. Sagt på en annen måte: Verbalisering av persistensatferd er ikke en funksjon av læringsbetingelsene, men av den atferd folk viser under ekstinksjonsbetingelser. Sagt på en tredje måte: Persistens er sannsynligvis primært et atferdsmessig fenomen.

## Påvirkes persistensatferd av informasjon om læringskontingensen?

Så langt har vi diskutert om ulike læringsbetingelser fører til ulik grad av vurdert persistens. Svaret er altså "nei!". Oppildnet av dette (for oss overraskende) funnet stilte vi også motsatt spørsmål: Vil persistensatferd påvirkes av kontingensinformasjon? Kunne det være slik at vi også her ville finne en "frikopling" mellom verbalisering og faktisk atferd?

Dette spørsmålet er interessant av mange grunner. Generelt har vi her å gjøre med forholdet mellom informasjon og faktisk atferd. Vi antar at det vi gjør i stor grad er påvirket av relevant informasjon. Får vi vite at det er en spesielt stor lottogevinst en uke, øker dette sannsynligheten for at vi vil spille i lotto. Får vi vite at vi risikerer å bli tatt hvis vi kjører for fort, så senker vi farten.

Innenfor foreliggende prosjekt undersøker vi en variant av denne type situasjon, hvor ekstinksjonsbetingelser legges inn. Vi gir informasjon om sannsynlighet for et bestemt utfall, og så lar vi forsøkspersonen forsøke å oppnå dette utfallet. Utfallet vil imidlertid aldri inntreffe (ekstinksjon). Spørsmålet er nå om den informasjon om sannsynlighet vi har gitt, vil påvirke personens persistens under ekstinksjonsbetingelser. Tabell 3 gir en oversikt over hvordan eksperimentene er satt opp.

Sunn fornuft tilsier at persistens påvirkes av informasjon om sannsynlighet for ønskede utfall. Sett at du får vite at en bestemt brusmaskin er "kranglete". Du forsøker så å kjøpe en Cola, men maskinen gir deg ikke noen flaske. I en slik situasjon er det sannsynlig at du vil forsøke ganske lenge før du gir opp. Her er i så fall sunn fornuft på linje med prinsippet om PREE. Men er våre forsøkspersoner med på dette?

I et eksperiment (Svartdal, Breivik, Nygreen og Hansen, 2002) lot vi forsøkspersoner snurre et "lykkehjul" for å oppnå "treff". Tre uavhengige grupper av forsøkspersoner fikk ulik informasjon om sannsynlighet for å oppnå treff (lav, middels, høy). Snurring ga imidlertid aldri "treff" (ekstinksjon). Spørsmålet var nå

*Tabell 3. Oppsett av eksperiment for å undersøke effekt av informasjon om læringsbetingelser på ekstinksjonsatferd*

Informasjon	Ekstinksjon	Persistens
Høy sannsynlighet for ønsket utfall	Personen prøver, men oppnår ikke ønsket utfall	Hvor lenge holder personen på før han/hun gir opp?
Lav sannsynlighet for ønsket utfall	Personen prøver, men oppnår ikke ønsket utfall	Hvor lenge holder personen på før han/hun gir opp?



om informasjonen de hadde fått ville påvirke persistens under ekstinksjonsbetingelser. PREE tilsier at de som fikk informasjon om lav sannsynlighet, ville holde på lenger før de ga opp enn forsøkspersoner som hadde fått informasjon om høy sannsynlighet for utfall. Våre data viser at det generelt ikke er en slik effekt. Forsøkspersonene holdt det gående i ½ - 1 time før de ga opp, men det var ingen sammenheng mellom informasjon (høy, middels, lav sannsynlighet for ønsket utfall) og persistens under ekstinksjon. Riktignok viste menn en tendens til å følge PREE, mens kvinner ikke gjorde det.

I et annet eksperiment gjennomførte vi dette eksperimentet på en noe annen måte, med spesiell fokus på kjønnsforskjeller. Forsøkspersoner trykket på to knapper på dataskjermen på en PC for å produsere tall mellom 1 og 100. Hvert tall de fikk ble så sjekket mot et ark med trykte tall personen hadde foran seg på bordet. Poenget var å oppnå 4 på rad. Sannsynligheten for å oppnå de korrekte tallene ble manipulert mellom uavhengige grupper (den ble sagt å være svært lav, middels, eller høy). Vi registrerte hvor lenge forsøkspersonene holdt på før de ga seg. Igjen viste det seg at den informasjon forsøkspersonene hadde fått om sannsynlighet for utfall ikke viste noen sammenheng med faktisk persistensatferd. I dette eksperimentet var det heller ikke kjønnsforskjeller, noe som tilsier at de kjønnsforskjellene vi observerte i det første eksperimentet høyst sannsynlig var knyttet til den konkrete eksperimentsituasjonen.

Data så langt viser altså – igjen noe overraskende – at persistensatferd ikke, eller i liten grad, er påvirket av relevant informasjon om sannsynlighet for utfall. Hvis dette skulle vise seg å være en riktig konklusjon (vi holder nå på med ytterligere studier i ulike situasjoner), er et slikt funn meget interessant. Generelt angår dette situasjoner hvor vi angir informasjon om mulighet for utfall, men hvor utfallene ikke inntreffer. Våre funn tilsier at folk er "blinde" for forholdet mellom informasjonen og egen atferd under slike betingelser. "Kjører du for fort, får du bot" er viktig informasjon, men hvis vi har erfaring for at håndhevelse av regelen forekommer sjelden eller aldri (ekstinksjonsbetingelser), mister informasjonen tydeligvis sin effekt. Våre data tilsier at informasjon om sannsynlighet for utfall i slike situasjoner i uhyre liten grad påvirker det folk gjør. Fra dagliglivet kjenner vi igjen situasjoner av samme type: "Det-skjer-ikke-meg"-effekten (til tross for massiv informasjon om farene ved røyking, fortsetter røykeren å røyke; "hvis man ikke deltar, har man i hvert fall ingen vinnerjans" (unnskyldning for å kjøpe lodd selv om vinnerjansene er mikroskopiske), er et par mulige eksempler.

## Avslutning

Det er lett å bli fascinert av dokumentasjon av misforholdet mellom det vi sier og det vi faktisk gjør (se Nisbett & Wilson, 1977, for eksempler). I denne artikkelen har vi beskrevet noen funn fra vår forskning på ekstinksjon, som påviser dette misforholdet på to måter: (a) Atferdsmessig viser folk de samme regelmessigheter i ekstinksjonsatferd som andre dyr, men verbaliseringen vi gir er i utakt. (b) Før vi handlingsrelevant informasjon, har denne informasjonen (i alle fall ifølge våre studier så langt) fint liten effekt på hva vi faktisk gjør.

## Referanser

- Amsel, A. (1994). Précis on *Frustration theory: An analysis of dispositional learning and memory*. *Psychonomic Bulletin & Review*, *1*, 280-296.
- Capaldi, E. J. (1994). The sequential view: From rapidly fading stimulus traces to the organization of memory and the abstract concept of number. *Psychonomic Bulletin & Review*, *1*, 156-181.
- Chatlosh, D. L., Neunaber, D. J., & Wasserman, E. A. (1985). Response-outcome contingency: Behavioral and judgmental effects of appetitive and aversive outcomes with college students. *Learning and Motivation*, *16*, 1-34.
- Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, *56*, 51-65.
- Humphreys, L. G. (1939a). The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *Journal of Experimental Psychology*, 141-158.
- Humphreys, L. G. (1939b). Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *Journal of Experimental Psychology*, *25*, 294-301.
- Nevin, J. A. (1988). Behavioral momentum and the partial reinforcement effect. *Psychological Bulletin*, *103*(1), 44-56.
- Nisbett, R., & Wilson, T. D. W. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.
- Pavlik, W. B., & Carlton, P. L. (1965). A reversed partial reinforcement effect. *Journal of Experimental Psychology*, *70*, 417-423.
- Shanks, D. R., & Dickinson, A. (1991). Instrumental judgment and performance under variations in action-outcome contingency and contiguity. *Memory and Cognition*, *19*, 353-360.
- Shanks, D. R., & St John, M. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, *17*, 367-447.
- Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.
- Svartdal, F. (1991). Operant modulation of rule-governed behavior by nonverbal contingencies. *Learning and Motivation*, *22*, 406-420.
- Svartdal, F. (2000a). Persistence during extinction: Conventional and reversed PREE under multiple schedules. *Learning and Motivation*, *31*, 21-40.
- Svartdal, F. (2000b). Persistence during extinction: Are judgements of persistence affected by contingency information? *Scandinavian Journal of Psychology*, *41*, 315-328.
- Svartdal, F. (2001). Ekstinksjon etter intermitterende forsterkning: Hva vet vi i dag? *Diskriminanten*, *28*.
- Svartdal, F. (2002). Extinction after partial reinforcement: Predicted vs. judged persistence. *Scandinavian Journal of Psychology*, in press.
- Svartdal, F., Breivik, C., Nygreen, I., & Hansen, T. (2002). Is persistence under no-reward affected by expectation of reward? Ms. in preparation.
- Svartdal, F., & Silvera, D. (2002). Cognitive measures of PREE: Predicted persistence after continuous vs. partial reinforcement in humans. Ms. submitted for publication.