

Masteroppgave

Atferdsvitenskap

Juni 2023

Effekten av tilbakemeldinger for å redusere
bekreftelsestendensen

Kandidatnavn: Svein Kevin Loe
Emnekode: MALK5000

30 studiepoeng

Fakultet for helsevitenskap
OSLO METROPOLITAN UNIVERSITY
STORBYUNIVERSITETET

Effekten av tilbakemeldinger for å redusere bekreftelsestendensen

Svein Kevin Loe

Institutt for atferdsvitenskap, fakultet for helsevitenskap, OsloMet – storbyuniversitetet

MALK5000

15.06.23

Sammendrag

Formålet med denne studien var å undersøke effekten av tilbakemeldinger for å redusere bekreftelsestendensen hos deltakere. Studien ble gjennomført ved å eksponere to grupper for samme treningstiltak i forskjellig rekkefølge. Til sammen deltok 10 voksne personer i studien, med fem deltakere i hver gruppe. Gruppe 1 gjennomgikk treningen i rekkefølgen ABACA, mens gruppe 2 gjennomgikk den i rekkefølgen ACABA og responsmønstrene ble registrert på et nettbrett. A-fasen representerte baselinefase, B-fasen representerte enkle tilbakemeldinger og til slutt C-fasen som representerte konkrete tilbakemeldinger. Et "mixed design" med vekt på innen-deltaker-design ble benyttet i eksperimentet. Resultatene viste at deltakere fra begge grupper viste en reduksjon i bekreftelsestendensen. En videre analyse viste at konkrete tilbakemeldinger hadde en større effekt sammenlignet med enkle tilbakemeldinger. Disse funnene indikerer at tilbakemeldinger kan spille en viktig rolle i å redusere bekreftelsestendensen hos enkeltpersoner. Å gi deltakerne konkrete tilbakemeldinger viste seg å være spesielt effektivt. Denne studien bidrar dermed til vår forståelse av hvordan tilbakemeldinger kan brukes som en intervensjon for å motvirke bekreftelsestendensen, og har potensiale til å ha praktisk betydning for ulike områder der bekreftelsestendensen kan påvirke beslutningsprosesser.

Nøkkelord: Bekreftelsesfellen, bekreftelsestendensen, seleksjonsoppgaver, tilbakemeldinger, atferdsanalyse

Abstract

The purpose of this study was to investigate the effect of feedback in reducing confirmation bias among participants. The study was conducted by exposing two groups to the same training intervention in different sequences. A total of 10 participants took part in the study, with five participants in each group. Group 1 underwent the training in the sequence ABACA, while group 2 underwent it in the sequence ACABA, and response patterns were recorded on a tablet. The A-phase represented the baseline phase, the B-phase represented simple feedback, and finally, the C-phase represented concrete feedback. A mixed design with an emphasis on within-participant design was employed in the experiment.

The results showed that participants from both groups exhibited a reduction in confirmation bias. Further analysis revealed that concrete feedback had a greater effect compared to simple feedback. These findings suggest that feedback can play a crucial role in reducing confirmation bias in individuals. Providing participants with concrete feedback proved to be particularly effective. Therefore, this study contributes to our understanding of how feedback can be used as an intervention to counteract confirmation bias and has the potential to have practical implications in various areas where confirmation bias can impact decision-making processes.

Keywords: Confirmation bias, confirmation tendency, selection tasks, feedback, behavior analysis

Effekten av tilbakemeldinger for å redusere bekreftelsestendensen

Ifølge Karl Popper (1959/2002) er det som skiller vitenskap fra pseudovitenskap at hypoteser er testbare og falsifiserbare. Han argumenterte for at vitenskapelige teorier må kunne testes gjennom eksperimentering eller observasjon. Dersom resultatene av testene ikke matcher hypotesen, må teorien revideres eller avvises (Popper, 1959/2002). Falsifisering er prosessen av å finne ut om en påstand eller hypotese er falskt ved å finne bevis som motsier dette. Motsetningen av falsifisering er å bekrefte der en søker etter bevis som støtter påstanden eller hypotesen. Når en hypotese er formulert, er det sannsynlig at en avleder ulike prediksjoner som gjennom eksperimenter blir testet. En kan si at hypotesen er falsifisert dersom resultater går mot prediksjonen. Et eksempel som illustrerer dette, er at ordtaket «alle svaner er hvite» blir en generalisering eller en hypotese som kan testes gjennom observasjon. Dersom noen observerer en svart svane, vil hypotesen bli falsifisert. På den annen side, helt til en har gjort seg opp den observasjonen, er ikke hypotesen testet og derfor ikke falsifisert. Alt som skal til, er én observasjon til å falsifisere påstanden. Popper (1959/2002) skriver at det er viktig at en søker etter informasjon som også kan falsifisere en hypotese, slik at en slutning er så gunstig som mulig. Til tross for det skriver Nickerson (1998) at det foreligger en stor tendens blant mennesker til å lete etter informasjon som skal bekrefte istedenfor informasjon som avkrefter. Den britiske filosofen Francis Bacon prøvde å beskrive allerede på 1600-tallet hvordan en selekterer informasjon vi blir presentert:

The human understanding when it has once adopted an opinion (either as being the received opinion or as being agreeable to itself) draws all things else to support and agree with it. And though there be a greater number and weight of instances to be found on the other side, yet these it either neglects and despises, or else by some distinction sets aside and rejects; in order that by this great and pernicious

predetermination the authority of its former conclusions may remain inviolate (Bacon, 1620, s 9-10).

Denne påstanden til Bacon kan tolkes som et forsøk på å definere *confirmation bias*. Nickerson (1998) definerer bekreftelsestendensen som også er kjent som «confirmation bias» som tendensen til å søke, tolke og huske informasjon på en måte som bekrefter ens tro eller hypotese mens en ignorerer motstridende bevis. Med andre ord kan en si at mennesker har en tendens til å legge merke til og huske informasjon som støtter deres tro og ignorere eller glemme informasjon som strider imot troen (Klein, 2019a). En kan også kalle dette for en kognitiv skjevhet, og dette definerer Hogg og Vaughan (2017) som en systematisk feil som ligger til stede når en interpreterer og prosesserer informasjonen i omgivelsene våre. Mennesker, ifølge Festinger (1957), har en tendens til å være motivert til å holde atferden og holdninger konsekvent grunnet at handlinger som kan motstride egen holdning kan resultere i et psykisk ubehag. Olsen (2008) skriver at dersom en har oppfatning om noe, vil en som oftest ikke endre oppfatningen om en møter på bevis som motstrider. Tvert imot vil en sette søkelys på hva som kan støtte den allerede eksisterende oppfatningen. Noen av kjennetegnene til *confirmation bias* kan minne om Dunning-Kruger-Effekten. Dunning-Kruger-Effekten er en psykologisk feilslutning som forekommer når personer med lite kunnskap overvurderer sin evne til å prosessere og tolke informasjon (Kruger & Dunning, 1999).

Dersom en forsøker å bruke *confirmation bias* som en forklaring på atferden, vil det fra et atferdsanalytisk perspektiv være et problem, grunnet at de kan plasseres i forskjellige kategorier. En kan ikke si at en person leter etter bevis som støtter egen tro på grunn av *confirmation bias*, eller på den annen side at *confirmation bias* er årsaken til at en ser etter bevis som støtter ens egen tro. Denne forklaringen kommer en ingen vei med, og kan virke sirkulær (Ryle, 1949). En kan også se på dette som utfordrende da *confirmation bias* blir snakket om som om det er en ting som en finner inne i kroppen til et menneske. Dersom en

snakker om det som om det er en ting, kan en risikere at *confirmation bias* blir brukt som årsaksforklaring til atferden, og hindre andre variabler til å bidra til å forklare atferden. Skinner (1974) skriver at i leting på årsaksforklaring til en atferd, vil forklaringer som er mentalistiske være en svakhet, da resulterer i en større utfordring å finne en korrelasjon mellom atferd og miljøomgivelser.

Svart-Svane teorien, som ble introdusert av Nassim Nicholas Taleb i hans bok «The Black Swan» fra 2010, refererer til et fenomen der mennesker har en tendens til å tro på at noe er sant eller usannsynlig basert på observasjoner og erfaringer fra tidligere, men blir overrasket når de opplever noe som utfordrer de tidlige oppfatningene. Det mest kjente eksempelet Taleb (2010) bruker er svart-svane teorien. I europeiske land var antagelsen om at det bare eksisterte hvite svaner reell, og dette var basert på flere århundrer med observasjoner med hvite svaner. Antagelsen ble betraktet som en sannhet eller et allment fakta og det var ingen antydning til at det skulle finnes svarte svaner. På den annen side endret dette seg da oppdagelsesreisende kom til Australia på 1600-tallet og observerte svarte svaner, og dette sto i strid mot den opprinnelige antagelsen om at det bare fantes hvite svaner. Dette eksempelet viser hvordan individer kan ha en tendens til å stole på det de har sett tidligere og at dette kan bli generalisert til en generell sannhet, til tross for at det foreligger begrenset bevis. Taleb (2010) skriver også at i en kompleks og uforutsigbar verden så er det utrolig viktig å være åpne for muligheten for at uventede hendelser kan oppstå, og at egen antagelse og oppfatning bør revurderes. Svart-svane teorien kan trekkes inn for å forstå hvordan individer har en tendens til å generalisere basert på det de har sett frem til klare unntak er observert.

Atferd som produserer en effekt på miljøet og som bli selektert gjennom læringshistorien med konsekvenser blir omtalt som operant atferd. Operant betinging er prosessen hvor atferdsendringen (læringen) forekommer som en funksjon av konsekvensen av atferden (Cooper et al., 2020). Bekreftelsestendensen er tendensen til å søke, tolke og huske

informasjon på en måte som bekrefter ens tro eller hypoteser. Å lete etter informasjon som bekrefter egen tro vil kunne øke sjansen for at atferden er lignende i framtiden. Dersom en ser på dette fra perspektiv som er atferdsanalytisk, kan en se på dette som forsterkning, dersom informasjonen stemmer overens med egen tro. Positiv forsterkning er referert til prosessen der en atferd blir positivt forsterket, i mål om at denne type atferd kan øke i frekvens. Teorien om positiv forsterkning kan bidra til å forsterke bekreftelsestendensen. Et eksempel på det er om en person mottar positiv forsterkning for deres tro eller hypoteser. Den informasjonen som støtter troen eller hypotesen kan fungere forsterkende. En kan forestille seg at en tror på at koronavaksinen inneholder giftstoffer og er farlig. Dersom denne personen mottar positiv forsterkning i form av artikler, enighet blant venner og kommentarer i forum på internett, vil dette styrke troen på at koronavaksinen er farlig. Dette kan også føre til en større sannsynlighet for at en fortsetter å søke etter bekreftende informasjon og tolker på en måte som støtter egen tro. Istedenfor å ta alle positive elementer ved koronavaksinen i betraktning, så kan bivirkningene være i størst fokus. Ved å eliminere eller ignorere alt som strider imot egen tro, vil kunne være negativt forsterkende. På den annen side vil det å komme i kontakt med noe som ikke er i tråd med egne meninger, i tillegg til å gå glipp av meninger som støtter egen tro, minske en sannsynlighet for en lik atferd i fremtiden. Catania (1998) skriver at dersom en blir utsatt for bekreftelse og avkreftelse, vil bekreftelse være mer forsterkende enn avkreftelse. En grunn til dette kan være at vi letter etter informasjon som henger sammen med forsterkning. Sammenhengen mellom atferdsanalyse og bekreftelsestendensen kan ses på gjennom et eksperiment som ble gjennomført av Dinsmoor, Browne og Lawrence i 1972.

Dinsmoor et al. (1972) gjennomførte et eksperiment der målet var å teste responsmønsteret til duer. I begynnelsen ble de lært opp til å hakke på en plate, etterfulgt av forsterkning i form av mat. Neste steg var å arrangere betingelser slik at hakking fra duene på en annen plate med en observasjonsknapp som resulterte i grønt eller rødt lys. Hakking i

nærvær av grønt lys på den første platen førte til forsterkning, i motsetning til hakking i nærvær av det røde lyset som ikke resulterte i forsterkning. En kan si at hakkingen fra duene på observasjonsknappen var for å bekrefte/avkrefte om det resulterte i forsterkning på den første knappen. Dinsmoor et al. (1972) modifiserte betingelser til at hakking på knappen til venstre resulterte i et rødt eller grønt lys. I denne betingelsen ville fargen grønn fungere som en positiv diskriminativ stimulus (S+), og dette grunnet at i nærvær av det grønne lyset ville hakking føre til at mat ble levert. På den annen side vil ikke hakking på knappen til høyre resultere i forsterkning, og dermed får det funksjonen negativ diskriminativ stimulus (S-). Dinsmoor et al. (1972) ville se om denne informasjonen var legitim, og utelukke andre faktorer som kunne påvirke hakkingen, så endringer i betingelsene ble gjort. Det ble gjort slik at hakking på knapp til venstre resulterte i et lys som var grønt (S+) eller ingen lys i det hele tatt. Det var fortsatt en mulighet at mat kunne blitt gitt ved å hakke på knappen til høyre, men bare om det hendte i nærvær av det grønne lyset. Da duene var under de betingelsene ble det ikke registrert noen endring i responsfrekvensen på observasjonsknappen, og dette i sammenligning med den forrige fasen som både inkluderte S+ og S-. Til slutt ble det gjort en endring i betingelsene, og dette innebar at det ikke var forekomst av det grønne lyset, og at hakk på knappen til venstre kun ville utløse rødt lys eller ikke lys. Responsfrekvensen på observasjonsknappen hos duene falt kraftig ved denne intervensjonen. Det er fortsatt viktig å påpeke at det fortsatt er forsterkning når hakkingen på observasjonsknappen ikke leder til rødt lys. Konklusjonen var at dette resulterte med at når duene hakket på observasjonsknappen, så ble det gitt straff ved å tilføre S- eller ekstingvert (med å ikke resultere en forandring med lys). Catania (2013) definerer ekstinksjon som en atferd som har forekommet, ikke blir forsterket lengre og slutter å forekomme. Det var i den grunn samme mengde med informasjon om tilgjengelighet av forsterker i de tre informasjonsbetingelsene. En kan oppsummere at i denne studien så fant Dinsmoor et al. (1972) ut at det ble registrert mest

hakking på observasjonsknappen til venstre når det ble produsert S+. De ble konkludert med at når S+ var til stede, så hadde duene en responsfrekvens som var atskilligere høyere enn når S- var til stede. En kan si at det foreligger en indikasjon på at duene var motiverte i større grad til å få bekreftelse på om maten var tilgjengelig i motsetning til å få bekrefte at maten ikke var tilgjengelig. Differansen av responsfrekvensen kan ses i sammenheng med bekreftelsesskjevheter, og dette grunnet at vi forsøker å unngå å se på informasjon eller på omstendigheter som resulterer i at vi mislykkes. Catania (2013) skriver at å se «korrekt» informasjon er forsterking i seg selv, og dette til tross for at det vi finner ikke er rasjonelt. Catania (2013) skriver også at duenes læringshistorie kan påvirke atferden og effektiviteten i betingelsesprosedyrene. Duenes tidligere læringshistorie med forsterkning kunne påvirke hvor raskt de lærer seg ny atferd og hvor sannsynlig det er for at de engasjerer seg i visse atferder i framtiden. I likhet med denne studien, så kan en trekke likheter fra studien til Dinsmoor et al. (1972) til Wasons *four card problem* som ble forsket på i 1968.

Den britiske psykologen Peter Wason var blant de første som benyttet seg av begrepet *confirmation bias*. Wason utførte innflytelsesrike studier på 1960- og 1970-tallet der han undersøkte hvordan individer har en tendens til å søke bevis som bekrefter troen en han og hypoteser, istedenfor å aktivt oppsøke avkreftende bevis. Hans arbeid med seleksjonsoppgaven, som regnes å være et klassisk problem innen psykologien, førte til identifisering og beskrivelse av *confirmation bias* som en gjennomgripende kognitiv skjevhet i menneskelig beslutningstaking (Wason, 1960, 1968).

I 1966 presenterte Wason (gjengitt i Wason, 1968) en seleksjonsoppgave som blir omtalt som «fire kort-problemet». Fire kort-problemet var en berømt oppgave innen kognitiv psykologi, og denne seleksjonsoppgaven gikk ut på å presentere fire kort til deltakerne. Wason (1968) var interessert i å se hvordan deltakerne resonerte rundt regler som «dersom P, så Q». P, \bar{P} (ikke-p), Q og \bar{Q} (ikke-Q). Påstanden kunne være «dersom det er en vokal på den

ene siden så er det et partall på den andre siden». Deltakerne fikk beskjed om at hvert av kortene hadde et tall på den ene siden, og på den andre siden en bokstav. Det neste steget var å informere deltakerne om at de skulle velge de kortene som måtte snus for å godkjenne at regelen stemte. I studiet til Wason (1966, gjengitt i Wason, 1968) viste resultatene at majoriteten av de selekterte kortene var kombinasjonen bare P, eller bare Q eller kombinasjonen av PQ. Svært få av deltakerne valgte å snu \bar{Q} og det var heller ikke populært å velge \bar{P} . Wason (1968) skriver at dersom en kombinerer valget av P og \bar{Q} så vil man ikke bare bekrefte, men en vil også avkrefte regelen. Wason (1968) skriver at de fleste valgte P og Q, og trekker at individer er preget av bekreftelsesskjevheter gjennom ulike læringshistorier, som kan henge sammen med at vi har forventninger til at våre antakelser og settinger har en sammenheng.

Selv utenfor laboratoriet er bekreftelsestendensen et relevant og betydningsfullt tema som påvirker menneskers beslutninger og oppfatninger. Et eksempel på dette er «algoritmer» innen sosiale medier. En algoritme er en oppskrift, som i denne sammenheng sørger for at mennesker får tilgang til enda mer informasjon som sannsynligvis er i samsvar med de tidligere søkene deres på internett. En kan forstå at innen netthandel for eksempel, så kan dette være gunstig. På en annen side blir det også argumentert for at dette utgjør en fare for demokratiet, grunnet at mange samfunnsaspekter blir ignorert (Baer, 2016). Algoritmer kan henge sammen med bekreftelsestendensen fordi en blir presentert mer informasjon som støtter troen en allerede har. Gitt at bekreftelsesfellen er et fenomen som kan oppstå i ulike former, kan det være betydningsfullt å bli informert om hvordan en kan unngå å havne i fellen selv.

Dette eksperimentet er en systematisk replikasjon av Wasons (1968) seleksjonsoppgaver, med en digital utformelse. Gruppe 1 og gruppe 2 ble opprettet, der gruppe 1 ble utsatt for ABACA, og gruppe 2 ble utsatt for ACABA. I den første treningsfasen fikk gruppe 1 enkle tilbakemeldinger som riktig/galt, og i den siste treningsfasen fikk de

konkrete tilbakemeldinger som forklarte hvorfor det var riktig/galt. Gruppe 2 hadde motsatt rekkefølge på treningsfasene og ble først presentert konkrete tilbakemeldinger og til slutt enkle tilbakemeldinger. Hensikten med denne studien er å se på hvordan to ulike treningstiltak med enkle og konkrete tilbakemeldinger med ulik rekkefølge har en effekt på bekreftelsestendensen. Det var også interessant å se om effekter var generaliserbare til abstrakte oppgaver med større omfang.

Metode

Deltakere

I denne studien ble 10 deltakere rekruttert. Disse var av begge kjønn og var i aldersgruppen 21-30 år. Deltakerne ble rekruttert etter et ikke sannsynlighetsutvalg og alle var tilfeldige familiemedlemmer og bekjente. Eksperimentator tok kontakt og ga informasjon om at det skulle bli gjort et eksperiment som var relevant for masteroppgave. Det ble ikke gitt noe mer informasjon til deltakerne, for at resultatene i studien ikke skulle påvirkes av denne. Det var heller ikke noen spesielle kriterier bortsett fra å være minst 18 år gammel. Det ble hentet inn en signatur gjennom et samtykkeskjema. Gjennom samtykkeskjemaet fikk deltakerne opplysning om at de ville være anonyme i oppgaven, og dersom de ville, kunne angre seg og trekke sitt samtykke når de ønsket.

Apparatur og setting

Datainnsamling er gjennomført på et nettbrett. Nettbrettet kalles for Microsoft Surface TM Go, og det hadde en berøringsskjerm. Mer spesifikt er spesifikasjonene Microsoft Desktop-1D5oROC Intel(R) Pentium (R) CPU 4415Y @ 1.60GHz. Microsoft Surface TM Go har et 64-bits operativsystem og Windows 10. Datakilden er programmet med tittelen BB1, utviklet i vår Bias-Busters labgruppe ved OsloMet og skrevet i Microsoft Visual Basic 1.0, 2010 Express. Programvaren arrangerte eksperimentelle forhold og registrerte kjønn, alderskategori (i intervaller på 5 år), gruppe (1 eller 2), samt avgitte svar (hvilke Wason-kort

som ble selektert) på hver oppgave i forsøket. Teksten i programmet er på norsk bokmål, i likhet med alle deltakeres morsmål i dette eksperimentet. Datainnsamlingen ble gjennomført på ulike arenaer som skole, bibliotek, hjemme hos eksperimentator og jobb. Rollen til eksperimentator var å bistå deltakeren med praktisk hjelp, men ikke noe som kunne påvirke resultatene i studien. Dersom deltakerne hadde noen kommentarer, ble disse notert ned av eksperimentator.

Design

I denne studien ble det anvendt et «mixed design» med hovedfokus på innen-deltaker-design. Dette kan forklares med at målingene av deltakerens atferd under baselinefase ble anvendt som kontroll og grunnlag for å sammenligne effektene av intervensjoner i forskjellige faser i denne datainnsamlingen. Datainnsamlingen bestod av tre baselinefaser der informasjonen om en avhengig variabel ble samlet inn, og to intervensjonsfaser der ulike uavhengige stimuli ble presentert i hver fase. De ulike bokstavene representerer ulike faser der A er baseline, B er tiltak med enkle tilbakemeldinger og C et annet tiltak med konkrete tilbakemeldinger. Deltakerne ble i tillegg tilfeldig plassert i to grupper. Deltaker 1, 2, 3, 4, 5 i gruppe 1 og deltaker 7, 8, 9, 10 og 11 i gruppe 2. Deltaker 6 ble kastet ut av programmet, og er derfor ekskludert fra denne studien. Gruppene ble presentert for de samme fasene, men med forskjellig rekkefølge på de eksperimentelle manipulasjonene. Gruppe 1 ble presentert for ABACA og gruppe 2 ble presentert for ACABA. En fordel med dette designet er at det gir mulighet for eksperimentell kontroll, noe som betyr at forskerne kan styre og manipulere uavhengige variabler for å se hvordan det påvirker den avhengige variabelen. I denne studien er de uavhengige variablene to forskjellige typer feedback som blir presentert for deltakerne for å se hvordan det påvirker responsmønstrene deres. Responsmønstrene til deltakerne utgjør den avhengige variabelen. Etter at dataene ble samlet inn, ble det gjennomført en visuell analyse slik at en kunne få en god oversikt over atferdsendringer som kan være signifikante.

For å gjøre dette, ble dataene organisert i tabeller i Excel og visualisert i figurer. Dette gjorde det lettere å se endringer over tid.

Prosedyre

I begynnelsen av eksperimentet mottok deltakerne skriftlige instruksjoner om de kommende fasene og oppgavene de ville bli bedt om å utføre i hver påfølgende fase. De ulike faser inneholdte seks oppgaver med fire kort hver, unntatt generaliseringsfasene som kun inneholdte tre oppgaver, med åtte kort per oppgave. Hver oppgave inneholdte opplysninger og en regel, mens kortene inneholdt enten et ord, et symbol eller et bilde fra en spesifikk kategori på den ene siden, og et ord, et symbol eller et bilde fra en annen kategori på den andre siden. Bare én side av kortene var synlige for deltakerne, og kortene ble presentert i tilfeldig rekkefølge. Deltakerne skulle velge hvilke(t) kort som måtte snus for å teste om en gitt regel var riktig eller ikke. Etter å ha gjort sitt valg, ble de bedt om å klikke på avgitt svar for å sende inn svaret, og for å gå videre til neste oppgave. Alle oppgaver hadde en tidsbegrensning på enten 180 eller 300 sekunder, der den første oppgaven i fasen med meningsfulle oppgaver og tilbakemeldinger hadde lengre tidsbegrensning på 300 sekunder. Om tiden gikk ut før deltakerne hadde svart, ble de automatisk sendt videre til neste oppgave, og svaret ble registrert som feil. For å få riktig svar i alle oppgavene, måtte deltakerne velge kombinasjonen av P-kort og \bar{Q} -kort, da dette ville vise at de både bekreftet og falsifiserte påstandene. Se tabell 1 for hvilke responsalternativer som ble presentert

Fase 1 (baseline)

I fase 1 av studien gjennomførte begge gruppene en baseline-testing. Hensikten med denne fasen var å undersøke om deltakerne hadde et forbedringspotensial i forhold til bekreftelsestendensen. Ved denne fasen fikk ikke deltakerne noe tilbakemelding på om svarene deres var riktige er ukorrekte, og dette gjelder også for de andre fasene som var kategorisert som baseline. Før deltakerne skulle begynne på oppgaver i baselune, fikk de en

beskrivelse av hva som ville skje i de kommende oppgavene. På skjermen ble det vist at deltakerne ville bli presentert for flere kort. Hver med enten et navn, figur eller symbol på den ene siden, og et navn eller symbol fra en annen kategori på den andre siden. I hver oppgave ville det også bli gitt en påstand om kortene, og deltakerens oppgave var å identifisere hvilke(t) kort som måtte snus for å sjekke om påstanden var sann. Denne informasjonen ble vist til deltakerne, slik at de skulle få en forståelse av hva som skulle skje i oppgavene, for å forberede de til oppgavene de skulle løse. Ifølge instruksjonene kunne deltakerne markere kortene de ønsket å snu ved å klikke i boksen som ble presentert på skjermen. De fikk også beskjed om at de kunne endre svaret sitt ved å klikke på kortet på nytt dersom de ombestemte seg. De kunne velge mellom «*touch*-skjerm» eller bruke musepeker når de skulle svare på alle oppgavene. Det ble gjort tester på programmet av eksperimentator, der det ble forsøkt å bruke både *touch* og musepeker, og dette kunne resultere i at programmet avsluttet. Så deltakerne måtte velge mellom en av de i begynnelsen og fortsette med samme måte og svare på oppgavene på. I løpet av baselinefasen ble deltakerne gitt seks oppgaver de skulle løse. I oppgave 1 under baselinefasen ble deltakerne presentert for en regel. «Hvis det er en sirkel på den ene siden, er det rødt på den andre siden. Det ble vist 4 ulike kort på skjermen og det første var en sirkel (P), en firkant med fargen grønn (\bar{Q}), fargen rød (Q) og til slutt en triangel (\bar{P}).

I oppgave 2 fikk deltakerne lese «Hvis det står Volvo på den ene siden, er det bilde av et glass på den andre siden». Det ble vist 4 ulike kort på. Det første kortet var ordet Volvo(P), etterfulgt av neste kort som var ordet Toyota (\bar{P}). Neste kort var et bilde av et glass (Q), og det siste kortet var et bilde av en blå kopp (\bar{Q}). Deltakerne ble presentert påstanden «Hvis det er en vokal på den ene siden, er det et partall på den andre siden». Kortene som ble vist på denne siden var bokstaven A (P), bokstaven D (\bar{P}), tallet 7 (\bar{Q}) og tallet 4 (Q). Oppgave 4 ble deltakerne presentert for informasjon om at alle korte kortene nedenfor har et fjellnavn på den

ene siden og et kjønnsymbol på den andre. Regelen i denne oppgaven var «Hvis det står Galdhøpiggen på den ene siden, er det et hannkjønn-symbol på den andre siden. De 4 kortene de kunne velge mellom var et bilde av hannkjønn-symbol (Q), hunnkjønn-symbol (\bar{Q}), ordet Glittertind (\bar{P}) og ordet Galdhøpiggen (P).

Neste oppgave var oppgave 5 og instruksjonen startet med teksten: Alle kortene nedenfor har bilde av et pålegg på den ene siden og et skriveredskap på den andre. Hvilke(t) av kortene må du snu for å sjekke om følgende påstand er riktig. Regelen er «Hvis det er 'ost' på den ene siden, står det «Penn» på den andre. De fire kortene som ble vist var ulike, og det første bilde var et bilde av syltetøy (\bar{P}), neste kort var ordet «penn» (\bar{Q}), neste kort var bilde av en ost (P) og til slutt et ordet «blyant» (Q). Oppgave 6 var den siste og deltakerne fikk se teksten «Alle kortene nedenfor har navn på en stor primat på den ene siden og navn på en fugleart på den andre siden». Regelen for denne oppgaven var «Hvis det står gorilla på den ene siden, står det undulat på den andre». Det var også fire kort i denne oppgaven som ble presentert. Det første kortet var ordet «sjimpanse» (\bar{P}), ordet «undulat» (Q), ordet «arapapapegøye» (\bar{Q}) og til slutt ordet «gorilla» (P). Selv om påstandene var like i alle oppgavene, så ble navnene/utseende på kortene endret for hver oppgave. Tabell 1 viser stimuliene som ble presentert for begge gruppene. Se figurer for responser som ble registrert hos deltakerne. De seks oppgavene overfor målte baseline, men hadde også for å velge ut hvilke deltakere som fikk være med i studien. Grunnen til dette var fordi det var viktig å luke ut de som kanskje hadde en læringshistorie med oppgaver som var lik. De som svarte riktig på alle oppgavene i baselinefasen ble ført til en avsluttende side med følgende teksten «Takk for deltakelsen». Kriteriet for å bli med videre var å svare feil på minst en av de seks oppgavene i baselinefasen.

Fase 2 (treningsfase)

Denne delen består av samme typer oppgaver som den første, men her kunne deltakerne snu hvert kort ved å klikke direkte på det. De får også tilbakemelding om svaret er korrekt eller ukorrekt. Før deltakeren kan klikke på knappen «Avgi svar» kan de ombestemme seg ved å klikke om igjen. Når deltakeren klikker «Avgi svar» så får de tilbakemelding på om svaret er korrekt eller ukorrekt. Dersom svaret er feil så kan de klikke «Prøv igjen!». Dersom svaret er riktig, kan de klikke på «Neste oppgave» for å gå videre. Makstiden på den første oppgaven i denne fasen er forlenget til fem minutter. De andre oppgavene er tre minutter lange. I fase 2 av studien gjennomgikk deltakerne i gruppe 1 en treningsfase. I denne fasen ble deltakerne presentert for seks oppgaver med fire forskjellige kort. Makstiden på den første oppgaven ble forlenget til fem minutter slik at deltakerne fikk ekstra god tid på å lese oppgaveteksten. «Hvis det er bilde av en gitar på den ene siden, står det 'Eple' på den andre siden», er et eksempel på en av oppgavene som ble gitt i denne fase. Det ble presentert samme påstander i de resterende oppgavene, bortsett fra at bilde/navn på kortene ble endret for hver oppgave. Deltakerne ble ikke lov å fortsette til neste oppgave før de klarte å svare riktig, eller før tiden gikk ut på hver enkelt oppgave.

I fase 2 gjennomgikk også deltakerne i gruppe 2 en treningsfase som inneholdte de samme oppgavene som deltakerne i gruppe 1 fullførte i treningsfasen. Ulikheten her var at deltakerne i gruppe 2 fikk utfyllende tilbakemeldinger på om svarene var riktige eller gale. Et eksempel på tilbakemelding kunne være at når en for eksempel trykte på nederste boksen på bildet av gitaren, så kunne det stå «Hvis det står noe annet enn 'Eple' på den andre siden av dette kortet, er påstanden feil, og derfor er det korrekt at dette kortet må snus». Alle kortene hadde forklaringer om det var korrekt eller ukorrekt å velge. Deltakerne måtte ikke velge et kort for å få forklaring, men de kunne trykke på en liten boks i hjørnet på kortet både før og etter de skulle velge kort. Dersom en trykte på forklaringen til kortet «Eple» som var feil å snu, ville forklaringen fortelle at «Påstanden sier ikke at det ikke kan stå 'Eple' på andre siden

av kort med bilde av et annet musikk-instrument enn gitar, og derfor er det unødvendig å snu dette kortet.

Fase 3 (baseline og generalisering)

I fase 3 av eksperimentet ble det utført en ny baseline-test for å undersøke om deltakerne hadde oppnådd kunnskap fra simple tilbakemeldinger som ble gitt i den tidligere fasen, etterfulgt av de tre generaliseringsoppgavene med åtte responsalternativer. I fase 3 for gruppe 2, gjennomgikk deltakerne en helt lik baseline-test som den første gruppen, etterfulgt av de tre generaliseringsoppgaver med åtte responsalternativer. Disse generaliseringsoppgavene var av samme type som gruppe 1 gikk gjennom.

Fase 4 (andre treningsfase)

I fase 4 for gruppe 1 gjennomgikk deltakerne samme typer oppgaver som i fase 2, men med den forskjell at de nå fikk mer konkrete tilbakemeldinger på hvorfor et responsalternativ var korrekt eller ukorrekt. Stimuliene var de samme som fase 1, og deltakerne fikk 5 minutter på den første oppgaven, mens resten av oppgavene hadde tidsbegrensning på 3 minutter. Fase 4 for gruppe 2 gjennomgikk deltakerne også like typer oppgaver som i fase 2, men de fikk kun enkle tilbakemeldinger. De enkle tilbakemeldingene helt like som gruppe 1 mottok i den aller første treningsfase.

Fase 5 (baseline og generalisering)

I fase 5 gjennomgikk gruppe 1 en ny baseline-test som fungerte som en post-test. Deretter ble det utført generaliseringsoppgaver for å undersøke om ferdigheter er blitt opprettet etter å ha deltatt i de forskjellige treningsfaser i forskjellige rekkefølger. På samme måte gjennomgikk gruppe 2 også en ny baseline-test etterfulgt av generaliseringsfaser for å undersøke eventuelle ulikheter i gruppen. Et eksempel på en påstand i en generaliseringsoppgave kunne være «Hvis det er et bilde av Eple på den ene siden, står det 'Queen' på den andre siden». (Marker kort du trenger og snu ved å klikke på kortet eller i

boksen oppe i venstre hjørne). I denne fasen fikk ingen av gruppene tilbakemelding på om svarene var riktig eller galt på de siste seks baselineoppgavene og tre generaliseringsoppgavene. Etter deltakerne anga sitt siste svar i siste generaliseringsoppgave, ble de ført til en ny side der det ble gitt informasjonen: Takk for at du deltok i dette forsøket! Det er «10» minutter siden du startet. Forsøket dreier seg blant annet om læringseffekter av ulike typer «*feedback*». Til slutt ble det gitt en score på hvor mange riktige oppgaver deltakeren hadde.

Resultater

Baseline (A-fase)

I den første baselinefasen hadde tre av fem deltakere i gruppe 1 og fire av fem i gruppe 2 ingen korrekte svar, mens to i gruppe 1 (Deltakerne 3 og 4) og én i gruppe 2 (Deltaker 8) hadde et korrekt svar. Alle deltakerne i begge gruppene tilfredsstilte kriteriet for å gå videre til første treningsfase.

Treningsfaser 1 (B for gruppe 1 og C for gruppe 2)

Gruppe 1

I B-fasen med enkle tilbakemeldinger hadde deltaker 1 henholdsvis 3 og 4 feil før korrekt respons på de to første oppgavene, og løste deretter de siste fire oppgavene på første forsøk etter mindre enn 20s. Deltaker 2 hadde 35 feil på den første oppgaven, 11 på den andre, og 9 på den tredje. Tiden gikk også ut alle disse tre gangene, som vil si at deltaker 2 ikke fikk muligheten til å få bekreftet riktig svar på de tre første oppgavene. På oppgave 4 og 5 gjorde deltaker henholdsvis 4 og 1 feil før korrekt respons, mens den siste oppgaven ble løst på første forsøk etter 12s. Deltaker 3 svarte korrekt på første forsøk på den første oppgaven, men svarte feil henholdsvis 10, 7, 1 og 15 ganger før riktig svar på andre til femte oppgave. På den 6. oppgaven svarte deltakeren riktig på første forsøk etter 4s.

Deltaker 4 gjorde 3 og 4 feil på den første og andre oppgaven, men svarte riktig på mindre enn 13s på de resterende oppgavene. Deltaker 5 gjorde 10 feil før riktig svar på første oppgave, deretter 1 feil på hver av de neste to oppgavene, men svarte korrekt etter mindre enn 11 sekunder. De siste oppgavene ble løst på første forsøk og ble gjort på under 15 sekunder.

Gruppe 2

I C-fasen med konkrete tilbakemeldinger hadde deltaker 7 henholdsvis 6 feil før korrekt svar på oppgave. Den andre og siste oppgaven krevde 1 forsøk før det ble korrekt, mens de siste oppgavene i denne fasen ble løst på første forsøk under 45s. Deltaker 8 hadde henholdsvis 1 feil på første oppgave før den ble løst. Deltakeren gjorde 3 feil på den andre og tredje oppgaven, 5 feil på fjerde oppgave og 11 feil på nest siste oppgave. På siste oppgave hadde deltakeren en feil før korrekt respons. Deltaker 9 hadde henholdsvis 1 feil før korrekt respons på de tre første oppgavene. Deltaker 10 brukte 9 forsøk på den første oppgaven før korrekt respons. Deltakeren hadde 2 feil på to neste oppgavene før korrekt svar, og 1 feil på nest siste oppgave. Fjerde og siste oppgave ble løst på første forsøk og det ble brukt 14s på hver av oppgavene. Deltaker 11 løste første, tredje og femte oppgave på første forsøk, mens de resterende oppgavene krevde 1 feil respons før korrekt svar ble avgitt.

Baseline og generalisering

Gruppe 1

Deltaker 1 hadde korrekt respons på fire oppgaver i baselinefasen og brukte til sammen 64s på alle seks oppgavene. I generaliseringsfasen brukte deltakeren 18-28s på hver oppgave, men hadde ukorrekt svar på alle oppgaver. Deltaker 2 hadde henholdsvis feil på alle oppgaver i denne baselinefasen og brukte mellom 2-14s på oppgavene. Det ble brukt 10-20 sekunder hver av oppgavene i generaliseringsfasen, men ingen var korrekte. Deltaker 3 svarte korrekt på den fjerde baselineoppgaven, og ukorrekt på de resterende oppgaver. Det ble brukt 1-125s på alle oppgavene. Ingen av generaliseringsoppgavene var korrekte og det ble brukt

10-25s på hver av oppgavene. Deltaker 4 hadde korrekt på den andre, fjerde og sjette oppgaven i baselinefasen, og ukorrekt på de resterende. Det ble brukt 26s på den første oppgaven, og mellom 2-6s på de andre oppgavene. Det ble brukt 6-14s på generaliseringsoppgavene, og ingen ble løst. Det ble registret korrekt svar respons på den tredje, femte og sjette oppgaven i baselinefasen for deltaker 5, mens de andre ble registrert som ukorrekt. Deltakeren bruke mellom 6-16s på hver av oppgavene. Deltaker 5 hadde korrekt på den første generaliseringsoppgaven, men hadde henholdsvis feil på de to siste og brukte 17-70s.

Gruppe 2

I denne baselinefasen hadde deltaker 7 henholdsvis korrekt på alle oppgaver og brukte 7-19 sekunder på å løse oppgavene. Den første og tredje generaliseringsoppgaven ble løst, men den andre oppgaven hadde ukorrekt svar, og i denne fasen ble det bruk 36-86s. Deltaker 8 hadde korrekt på den andre oppgaven i baseline og korrekt på resterende oppgaver med 3-23s responstid. Ingen av generaliseringsoppgavene ble løst og deltakeren brukte 23-59s. Deltaker 9 hadde feil svar på den første baselineoppgaven, men korrekt på resterende, og brukte mindre enn 10s på alle oppgaver. I generaliseringsfasen ble alle tre oppgaver løst, og deltakeren brukte mindre enn 20s på disse oppgavene. I denne baselinefasen hadde deltaker 10 korrekt svar på tredje oppgave og ukorrekt på resterende, og hadde 7-16s i responstid. Ingen av generaliseringsoppgavene ble løst og deltaker 10 brukte 13-19s på oppgavene. Deltaker 11 brukte mellom 3-42s på å svare på oppgavene, men hadde ingen korrekte svar i denne fasen. I likhet med denne baselinefasen, var det ingen korrekte svar i generaliseringsfasen med 32-58s i responstid.

Treningsfaser 2 (C for gruppe 1 og B for gruppe 2)

Gruppe 1

I C-fasen med konkrete tilbakemeldinger hadde deltaker 1 én feil før korrekt respons på andre oppgave. De resterende oppgavene ble løst på første forsøk på under 10s. I likhet med deltaker 1, hadde deltaker 2 én feil før korrekt respons på andre oppgave, og klarte å løse de resterende oppgavene på første forsøk, men hadde 7-46s i responstid. Deltaker 3 hadde 1 feil på første, tredje, fjerde og femte oppgave før korrekt respons. Den andre og sjette oppgaven ble løst på første forsøk under 65s. Deltaker 4 hadde korrekt respons på alle oppgaver på første forsøk, og brukte 33s på første oppgave og under 10s på de andre oppgavene. Deltaker 5 løste de første fem oppgavene på første forsøk og brukte under 20s på hver oppgave. Deltaker 5 hadde 1 feil på den siste oppgaven før riktig respons på 32s.

Gruppe 2

I B-fasen med enkle tilbakemeldinger brukte deltaker 7 bare én respons på hver oppgave og hadde en responstid på under 40s. Deltaker 8 svarte henholdsvis feil 17, 3, 5, 13 og 8 på første til femte oppgave. Tiden gikk også ut på alle de fem oppgavene. Den sjette oppgaven ble løst på første forsøk på under 90s. Deltaker 9 hadde korrekt på alle seks oppgaver på første forsøk og hadde 4-6s i responstid på alle oppgaver. Deltaker 10 løste alle seks oppgaver på et forsøk og hadde 4-9s i responstid. Deltaker 11 hadde tre og fem feil på første og tredje oppgave før riktig respons. Deltakeren løste andre, fjerde, femte og sjette oppgave på første forsøk og brukte under 20s på å løse hver av oppgavene.

Baseline og generaliseringsfase

Gruppe 1

I den siste baselinefasen, hadde deltaker 1 korrekt svar på første og sjette oppgave, og ukorrekt på de resterende oppgaver. Deltakeren brukte mellom 20 og 40s på de korrekte oppgavene. Deltaker 1 hadde ingen korrekte generaliseringsoppgaver, og brukte 11-17s på hver av oppgavene. I den siste baselinefasen hadde deltaker 2 ingen korrekte responser og brukte mindre enn 15 sekunder på hver oppgave. I den siste generaliseringsfasen hadde

deltaker 2 ingen korrekte responser og brukte under 5s på hver av oppgavene. Deltaker 3 hadde ingen korrekte svar på noen oppgaver i den siste baselinefasen, og brukte 0-2s på å svare. Deltaker 3 hadde null korrekte svar på de siste tre generaliseringsoppgavene og hadde mindre enn 1 sekund i responstid. Deltaker 4 hadde ingen korrekte svar i den siste baselinefasen og hadde under 35s i responstid på hver av oppgavene. Deltaker 5 hadde henholdsvis korrekt på alle baselineoppgaver bortsett fra den tredje oppgaven. Deltakeren brukte mindre enn 10s på å svare på hver oppgave. I den siste generaliseringsfasen hadde deltaker 5 korrekt på alle oppgaver. Deltakeren brukte 59s på den første oppgaven, men de siste to ble svart på under 10s.

Gruppe 2

I den siste baselinefasen hadde deltaker 7 alle oppgaver korrekt, og brukte maks 30s per oppgave. I den siste generaliseringsfasen hadde deltaker 7 korrekt på alle oppgaver og brukte 15-20s på hver oppgave. Deltaker 8 hadde ingen korrekte responser i den siste baselinefasen og brukte 17s på den første oppgaven, mens de resterende oppgavene ble svart på under 6s. I generaliseringsfasen hadde deltaker 8 ukorrekt på alle oppgaver med en responstid på 15-59s. Deltaker 9 hadde korrekt på alle siste baselineoppgaver og brukte mindre enn 12s på hver av oppgavene. I generaliseringsfasen hadde deltaker 9 henholdsvis korrekt på alle oppgaver, med en responstid på 5-17s per oppgave. Deltaker 10 hadde korrekt på alle oppgaver i den siste baselinefasen. Deltakeren brukte 21s på den tredje oppgaven, men under 6s på alle de andre oppgavene. I den siste generaliseringsfasen hadde deltaker 10 korrekt på andre og tredje oppgave og hadde under 10s i responstid på hver av oppgavene. Første oppgave ble ukorrekt, og deltakeren hadde 18s i responstid. Deltaker 11 hadde korrekt svar på første og sjette oppgave og brukte under 25s. De resterende oppgavene var ukorrekt. I den siste generaliseringsfasen hadde deltaker 11 henholdsvis ukorrekt på alle oppgaver og brukte under 20s på hver oppgave.

En visuell analyse av figurene viser at deltakerne i gruppe 2 hadde flere korrekte oppgaver enn gruppe 1 i den siste baselinefasen. Deltakerne i gruppe 1 hadde til sammen 10 korrekte oppgaver, mens deltakerne i gruppe 2 hadde til sammen 21 korrekte oppgaver. I generaliseringsfasene hadde deltakerne i gruppe 1 til sammen korrekt på 7 oppgaver, og deltakerne i gruppe 2 hadde til sammen korrekt på 13 oppgaver.

Diskusjon

Dette eksperimentet hadde som mål å sammenlikne effekten av to ulike typer tilbakemelding på korrekte og feil svar på oppgaver med Wason kort. Hovedspørsmålet var om trening med en forklarende tilbakemelding (C-fasen) var mer effektiv enn trening med en enkel korrekt/inkorrekt-tilbakemelding (B-fasen) når det gjaldt å redusere bekreftelsestendensen. For å undersøke om treningen ledet til en begrenset korrekt respondering på oppgaver av samme type som treningsoppgavene, der korrekt respons alltid var å snu to kort, ble deltakerne også testet på generaliseringsoppgaver med åtte kort, hvor korrekt respons ville være å snu fire av kortene. Deltakerne ble fordelt i to ulike grupper som ble eksponert for samme betingelser, bortsett fra at intervensjonsfasene ble byttet om. Gruppe 1 ble utsatt for ABACA og gruppe 2 ble utsatt for ACABA. Resultatene viste en reduksjon i forekomst av bekreftelsestendensen etter alle deltakere har vært gjennom ulike tilbakemeldinger

I den første baselinefasen skulle deltakerne svare på seks oppgaver, og det ble ikke registrert noen store forskjeller. Det var kun deltaker 3,4 og 8 som svarte korrekt på én av seks oppgaver. Resterende deltakere svarte feil på alle oppgaver i denne fasen. Det alle deltakere hadde til felles var at de selekterte ulike kombinasjoner av kort for hver eneste oppgave i den første A-fasen. Ingen av deltakerne valgte samme kombinasjon på alle seks oppgaver. Resultatene fra den første baselinefasen er helt i overensstemmelse med forskningen til Wason's four card problem (1966, gjengitt i 1968). I denne studien hadde alle deltakerne en

tendens til å velge enten P-kortet eller kombinasjon av P og Q i den første baselinefasen og den første generaliseringsfasen, likt som funnene i Wason (1966, gjengitt i 1968). I den første A-fasen hadde alle deltakere veldig likt responsmønster. De fleste deltakerne brukte like lang tid (0-80 sekunder) på den første baselinefasen, med unntak av deltaker 3 og 4 som brukte mellom 140-180 sekunder på noen av oppgavene i A-fasen. I forsøk gjort av andre, har begrepet bekreftelsesfellen også blitt nevnt som «*matching bias*» eller «tilpasningsbias». I oppgave 1 under baselinefasen ble deltakerne presentert for en regel. «Hvis det er en sirkel på den ene siden, er det rødt på den andre siden. Det ble vist 4 ulike kort på skjermen og det første var en sirkel (P), en firkant med fargen grønn (\bar{Q}), fargen rød (Q) og til slutt en triangel (\bar{P}). Thompson et al., (2013) skriver at en tilpasser responsen i forhold til hvilken regel som blir fremstilt. Deltakerne valgte kortene P og Q flest ganger, og disse kortene ble nevnt i regelen. Deltaker 3 og 5 var de eneste i A-fasen som valgte å snu \bar{P} -kortet. Wason (1968) skriver at \bar{P} -kortet er et kort som ikke er nødvendig å snu, grunnet at det ikke har noe å si på hva som befinner seg på andre siden av kortet. De fleste deltakere valgte gjennom hele denne fasen ikke å snu \bar{P} -kortet, i tråd med hva Wason (1968) sier om dette kortet. I denne fasen ble derimot P-kortet valgt 49/60 ganger blant alle deltakerne. I lys av den gjeldende studien er det mulig at deltakerne ikke hadde egne hypoteser som de siktet på å bekrefte. I stedet kan de bare ha reagert på kort P og Q som var de to som var nevnt i regelen. Valg av P og Q kan forstås som et resultat av en tendens til å velge basert på likheter (*matching bias*) istedenfor bekreftelsesfellen.

Den amerikanske forskningspsykologen Gary Klein (2019a) har uttrykt en kritikk mot seleksjonsoppgavene til Wason som har som mål å demonstrere tendenser i å søke bekreftelse. Klein (2019a) påstår at det ikke er noe poeng å forsøke å «omprogrammere» personer, og at det ikke er viktig å «de-biase». Han skriver også at tendenser til bekreftelsesfellen blir eliminert når oppgavene blir satt i en kontekst der folk ikke kan relatere

seg til dem, noe som resulterer i at de svarer annerledes eller feil. Klein (2019a; Klein, 2019b) postulerte at det som deltakerne opplevde i oppgavene til Wason, ikke kunne klassifiseres som bekreftelsestendensen. Han brukte ordet «fiksering» istedenfor, og argumenterte for at i stedet for å velge kort som deltakerne anså som korrekte eller ukorrekte, hadde de søkelys på bestemte kort. Ifølge Klein (2019a) handler begrepet fiksering om at folk blir låst fast på en spesifikk beskrivelse. Beskrivelsen kan være korrekt og nøyaktig, men når det ikke stemmer, innser vi senere at vi klamret oss til denne forklaringen i lang tid, noe som står i motsetning til våre primære hypoteser. Begrepet «fikseringsfeil» er ofte anvendt om hendelser der individer retter søkelyset på noe, samtidig som de ignorerer noe annet. Så en kan si at fiksering dreier seg om hvorfor vi holder på våre forklaringer. En kan oppsummere med å si at fiksering er når en blir sterkt bundet til en spesifikk tanke eller tilnærming. På den annen side handler bekreftelsesfellen om tendensen vi har til å lete etter og godkjenne informasjon som bekrefter troen en har.

I den første treningsfasen (b) kan det tolkes at deltakernes respondering i gruppe 1 var et eksempel på atferd som er kontingensformet, siden de ikke fikk klare instruksjoner om hvilke responser som var korrekte eller ukorrekte, før de gikk videre til den neste fasen. I stedet ble det samme kortet vist repeterende til deltakerne, helt til de ga korrekt respons. Som et resultat kan atferden hos deltakerne ha blitt påvirket av konsekvensene av deres handlinger, uten å være kontrollert av en regel om hva slags kort som skulle snus. På den annen side, kan en se på figurene for deltakerne i gruppe 1, at det i den andre treningsfasen (c) ble registrert en del korrekte svar. Det kan være mulig at deltakernes atferd etter å ha sett på de konkrete tilbakemeldingene ble regelstyrt, ettersom de studerte tilbakemeldingen som ga informasjon om hva som var nødvendig for å få en riktig respons. Det kan også tenkes at enkle tilbakemeldinger som riktig/galt fører til større sannsynlighet for «personlig» regelstyring, der flere teorier blir testet ut.

Det er blitt registrert store ulikheter blant deltakerne, og en visuell analyse av deltaker 2, som sett på figur 2, viser at deltakeren ikke klarte å svare riktig på oppgave 51, 52 og 53 innen tidsfristen. Det ble også registrert 35 forsøk på oppgave 51. Hos deltaker 8 ble veldig få oppgaver registrert som korrekt, der tiden gikk ut store deler av studien. En kan tydelig se i figur 8 at treningen ikke hadde så stor suksess. Det virker som at flere deltakere ga litt opp på noen oppgaver da de bare trykket fort og mange ganger helt tilfeldig, med litt «lært hjelpeløshet» for å komme videre til neste oppgave. Etter å ha analysert dataene så er det konkludert med at det ble registrert mindre forsøk på fasen med konkrete tilbakemeldinger i motsetning til fasen med enkle tilbakemeldinger hos deltakerne. Dette kan gi mening, da de konkrete tilbakemeldingene inneholder mer informasjon enn de enkle tilbakemeldingene. En visuell analyse av figurene viser at deltakerne i gruppe 2 hadde flere korrekte oppgaver i den siste baselinefasen. Deltakerne i gruppe 1 hadde henholdsvis 10 korrekte responser til sammen i den siste baselinefasen. Gruppe 2 hadde 21 korrekte responser i den sistebaseline fasen, som er over dobbelt så mye som gruppe 1. Det kan tyde på at det å få konkrete tilbakemeldinger tidlig i eksperimentet sammenlignet med enkle tilbakemeldinger har utgjort en vesentlig forskjell. Den ulike rekkefølgen kan ha påvirket læringsprogresjonen for hver av gruppene. En visuell analyse av figurene viser at til sammen så hadde deltakerne i gruppe 1 syv riktige generaliseringsoppgaver. På den annen side hadde deltakerne i gruppe 2 til sammen 13 riktige generaliseringsoppgaver.

I henhold til en studie utført av Kruger og Dunning (199), har over halvparten av befolkningen en tro på at de er bedre enn gjennomsnittet når det gjelder kompetanse. Når en har en så overdrevet tro på sin egen kompetanse, er det kanskje vanlig å undervurdere utfordringene knyttet til kritisk tenkning, rasjonalitet og beslutningstaking. Når en har dannet seg en oppfatning om noe, er det mulig at en vil ha større tillit til ens egne vurderinger i forhold til hva andres vurderinger er. Deltakerne ble gitt forklarende tilbakemeldinger som er

konkrete, og de tilbakemeldingene gir de begrunnelser for hva som er korrekt/ ukorrekt, med mål om å gi kompetanse til å løse oppgavene. En rekke av deltakerne i studien stilte spørsmål om dette er «sunn fornuft» når det var snakk om løsningene på oppgaver som ble tildelt. De var ikke klar over at det manglet kompetanse til å motvirke egen bekreftelsestendens. Det ble forsøkt å forme deltakernes atferd gjennom forsterkningsprosedyrer. Tilbakemeldingen «riktig» kunne ha funksjon som positiv forsterkning, mens tilbakemeldingen «galt» kunne ha funksjon som negativ forsterker. Fraværet av tilbakemeldingen «galt» kunne virke forsterkende. Når deltakerne fikk tilbakemeldingen «feil» kunne det virke positivt straffende, og fraværet av «riktig» kunne virke negativt straffende (Catania, 2013). Alt dette i mål om å forsterke ønsket responsmønster som er egnet for oppgaver i forsøket. Som påpekt av Catania (1998), kan bekreftelse ha en funksjon som forsterkning. Å gi både bekreftende og avkreftende forklarende tilbakemeldinger, vil fungere som en belønnet responsforsterkning for passende avkreftelse, og dermed unngå å illustrere bekreftelsestendensen. Dette kan være den nødvendige tilnærmingen for å endre slike responsmønstre som viser bekreftelsesbiasen. Funnene i studien til Dinsmoor et al. (1972) viste at når bare den negative diskriminative stimulusen var betinget av observasjonsresponsene, så sluttet duene å respondere. Dette kan indikere at avkreftelse fungerte som en stimulus som var straffende. Dersom vi antar at avkreftelse hadde en effekt som var straffende på deltakerne i dette eksperimentet, kan treningen ha motvirket avkreftelse som straffende stimuli. Samtidig ble de konkrete og enkle tilbakemeldinger etablert som forsterkere.

Noen svakheter med denne studien var hvordan plasseringen av «avgi svar» og «fortsett til neste oppgave» var på samme plass på hver eneste oppgave. Dette kunne føre til at når deltakerne skulle trykke på «avgi svar» så kunne de trykke slik at de kom borti på neste oppgave, som kunne føre til at det ble registrert et forsøk uten at de hadde lest oppgaven. Som sett i figur 2, kan en se at deltaker 2 i den første treningsfasen (b) hadde 35 responser i den

første oppgaven. En stor del av de responsene har bare et par sekunder mellom seg, som kan tyde på at deltaker 2 har kommet borti «avgi svar» utilsiktet. Det at programvaren var på et digitalt brett, innebar at det kunne være *bugs*/feilmeldinger i programvaren, som kan føre til ukorrekte registreringer. En annen ulempe var at noen av deltakerne i denne studien gjennomførte studien i en offentlig kantine som var støyete. Deltaker 8 gjennomførte eksperimentet i en bråkete kantine, og som sett i figur 8, gikk tiden ut hele 9 ganger i løpet av hele eksperimentet. Deltaker 8 nevnte til eksperimentator at «det er veldig bråkete», «jeg sliter med å konsentrere meg». En visuell analyse av figur 8 kan tyde på at både mange responser og antall sekunder brukt, kan henge sammen med det støyete miljøet under gjennomføringen.

Begge typene trening som ble presentert til deltakerne i denne studien hadde funksjon som «avbiasing», til en viss grad på kort sikt. Likevel kan en si at begge typene trening førte til forbedringer i antall korrekte svar. Gjennom en visuell analyse av figurene, kan en se at treningsfasen med konkrete tilbakemeldinger krevde mindre forsøk enn treningsfasen med enkle tilbakemeldinger. Dette kan forklares ved at treningsfasen med konkrete tilbakemeldinger inneholder begrunnelse på hva som er riktig, mens fasen med enkle tilbakemeldinger bare forteller om det er korrekt/ukorrekt.

Relabiliteten i denne studien anses være av god grad. Relabilitet handler om å vurdere hvor stabile og pålitelige resultatene er over tid (Shadish et al., 2002). En kan også si at relabilitet handler om å undersøke om en ville ha oppnådd samme resultater dersom en gjentar observasjoner eller målinger. I denne studien ble det anvendt et «mixed design» med hovedfokus på innen-deltaker-design. Dette kan forklares med at målingene av deltakerens atferd under baselinefase ble anvendt som kontroll og grunnlag for å sammenligne effektene av intervensjoner i forskjellige faser i denne datainnsamlingen. Datainnsamlingen bestod av tre baselinefaser, der informasjonen om en avhengig variabel ble samlet inn, samt to intervensjonsfaser, der ulike uavhengige stimuli ble presentert i hver fase. Samlet sett

indikerer resultatene fra denne studien at tilbakemeldinger kan være et kraftig verktøy for å redusere bekreftelsestendensen hos individer. Ved å demonstrere effekten av konkrete tilbakemeldinger sammenlignet med enkle tilbakemeldinger, kan denne studien bidra til vår forståelse av hvordan vi kan motvirke bekreftelsesfellen i ulike kontekster utenfor laboratoriet.

I midlertidig kan en stille seg undrende til om denne effekten har en begrensning til oppgaver som er spesifikke, og om en er under kontroll av regler, samt om det kan generaliseres til bekreftelsesfelle-utfordringer i verden som befinner seg utenfor laboratoriet. I fremtidig forskning innen dette området, bør det være fokus på å utforske diverse tilfeller av bekreftelsesfellen, og skape varierte tilnærminger for å minske forekomsten.

Referanser

Bacon, F. (1620). *Novum organum*. Clarendon Press.

Baer, D. (2016, November 9). The “Filter Bubble” explains why Trump won and you didn’t see it coming. *New York Magazine*. Retrieved from <http://nymag.com/scienceofus/2016/11/how-facebook-and-the-filter-bubble-pushed-trump-to-victory.html>

Catania, A. C. (1998). *Learning* (4. utg.). Prentice Hall.

Catania, A. C. (2013). *Learning* (5. utg.). Sloan Publishing.

Cooper, J. O., Heron, T. E. & Heward, W. L. (2020). *Applied behavior analysis* (3. utg.). Pearson Education.

Dinsmoor, J. A., Browne, M. P., & Lawrence, C. E. (1972). A test of the negative discriminative stimulus as a reinforcer for observing. *Journal of the Experimental Analysis of Behavior*, 18, 79-85. <https://doi.org/10.1901/jeab.1972.18-79>

Festinger, L. A. (1957). *A theory of cognitive dissonance*. Stanford University Press.

Hogg, M. A., & Vaughan, G. M. (2017). *Social Psychology* (8. utg.). Pearson Education Limited.

Klein, G. (2019a, 5. Mai). *The curious case of confirmation bias*. *Psychology today*. <https://www.psychologytoday.com/us/blog/seeing-what-others-dont/201905/the-curious-case-confirmation-bias>

Klein, G. (2019b, 11. juni). *Escaping from fixation*. *Psychology today*. <https://www.psychologytoday.com/us/blog/seeing-what-others-dont/201906/escaping-fixation>

Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, 77(6), 1121-1134. <https://doi.org/10.1037/0022-3514.77.6.1121>

- Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, 2(2), 175–220. <https://doi.org/10.1037/1089-2680.2.2.175>
- Olsen, R. A. (2008). Cognitive dissonance: the problem facing behavioral finance. *Journal of behavioral finance*, 9(1), 1-4. <https://doi.org/10.1080/15427560801896552>
- Popper, K. (2002). *The logic of scientific discovery* (2. Utg). Routledge. (Originalt verk publisert i 1959).
- Ryle, G. (1949). *The concept of mind*. Barnes.
- Shadish, W. R., Thomas D. Cook, & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Skinner, B. F. (1938). *The behavior of organisms*. Appleton-Century-Crofts.
- Skinner, B. F. (1974). *About behaviorism*. Vintage Books / Random House.
- Taleb, N. N. (2010). *The Black Swan*. (2. utg.). Random House Publishing Group.
- Thompson, V. A., Evans, J. S. B., & Campbell, J. I. (2013). Matching bias on the selection task: It's fast and feels good. *Thinking & Reasoning*, 19(3-4), 431-452. <https://doi.org/10.1080/13546783.2013.820220>
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly journal of experimental psychology*, 12(3), 129-140. <https://doi.org/10.1080%2F17470216008416717>
- Wason, P. C. (1968). Reasoning about a rule. *Quarterly journal of experimental psychology*, 20(3), 273-281. <https://doi.org/10.1080%2F14640746808400161>

Tabell 1

Oversikt over hvilke responsalternativer deltakerne ble presentert for

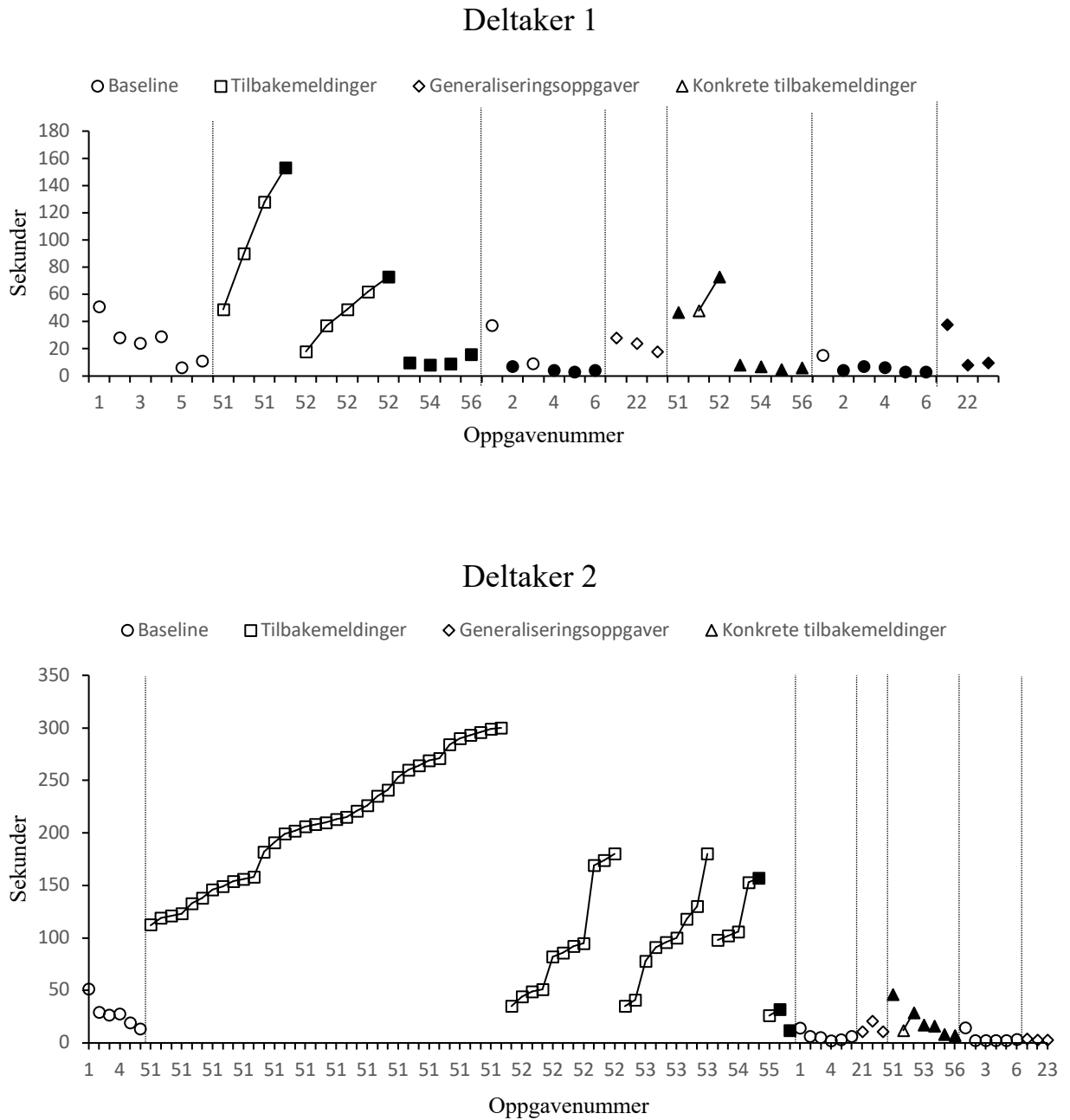
Oppgavenummer	P	\bar{P}	Q	\bar{Q}
1	Sirkel	Rød	Trekant	Grønn
2	Volvo	Glass	Toyota	Kopp
3	A	4	D	7
4	Galdhøpiggen	♂	Glittertind	♀
5	Ost	Blyant	Syltetøy	Penn
6	Gorilla	Undulat	Sjimpanse	Arapapapegøye
51	Joker	Selger ukeblader	Rema 1000	Selger ikke ukeblader
52	Gitar	Eple	Trompet	Pære
53	Valmuefrø	Elefant	Tulipan	Kanin
54	Snekker	Fiskeboller	Politi	Spaghetti
55	“Skrik” av Munch	New York	Mona Lisa	Moskva
56	Øl	25 år	Cola	16 år

	P	Q	\bar{P}	\bar{P}	\bar{P}	\bar{Q}	\bar{Q}	\bar{Q}
21	Eple	Queen	Pære	Druer	Kiwi	The Rolling Stones	The Beatles	Abba
22	Yatzy	Bipolar lidelse	Sjakk	Monopol	Stigespillet	ADHD	Autisme	Schizofreni
23	Saturn	Italia	Mars	Uranus	Neptun	Ghana	Peru	Mongolia

Merknad. Baselineoppgaver 1-6. Treningsfaser 51-56. Generaliseringsoppgaver 21-23.

Figur 1 og 2

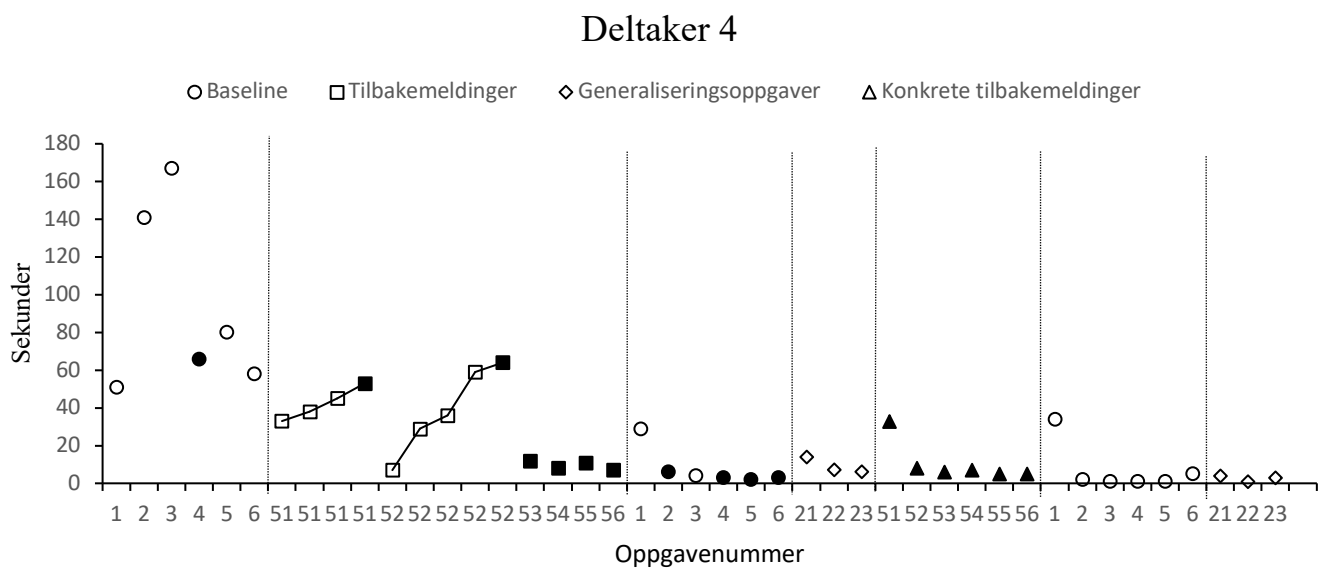
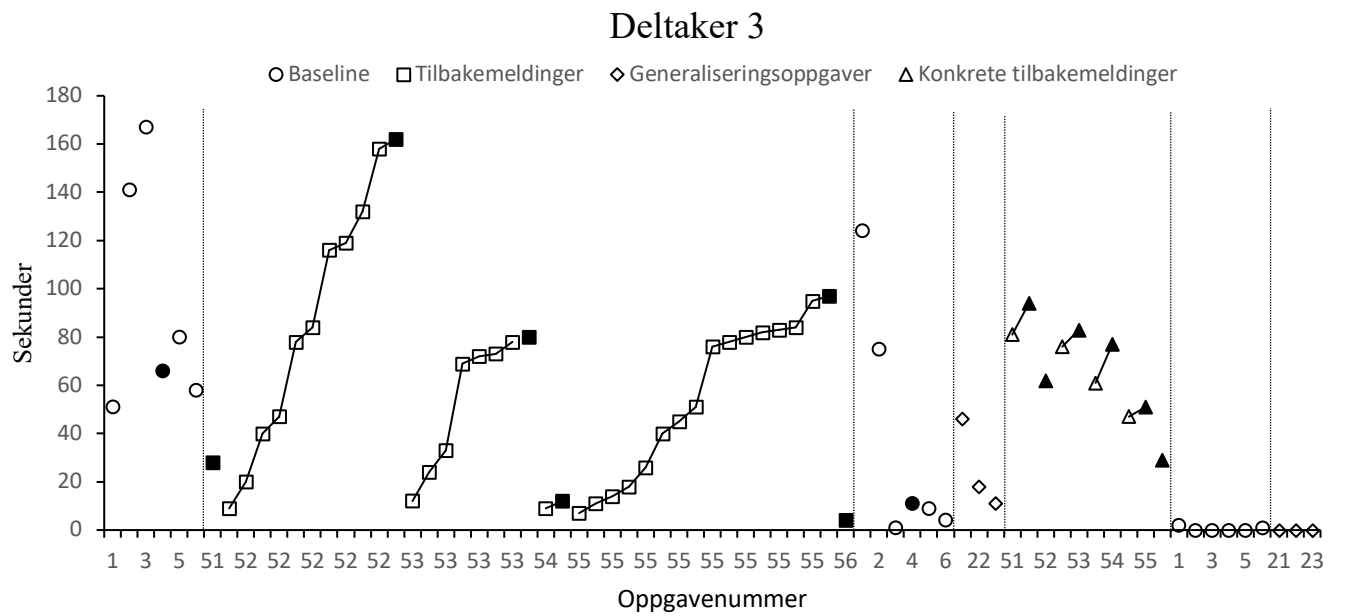
Resultater fra deltaker 1 og 2



Merknad: X-aksen illustrerer de reelle oppgavenumrene fra programvaren, og Y-aksen er sekunder. Sort indikator er korrekte responser og hvit indikator er ukorrekte responser.

Figur 3 og 4

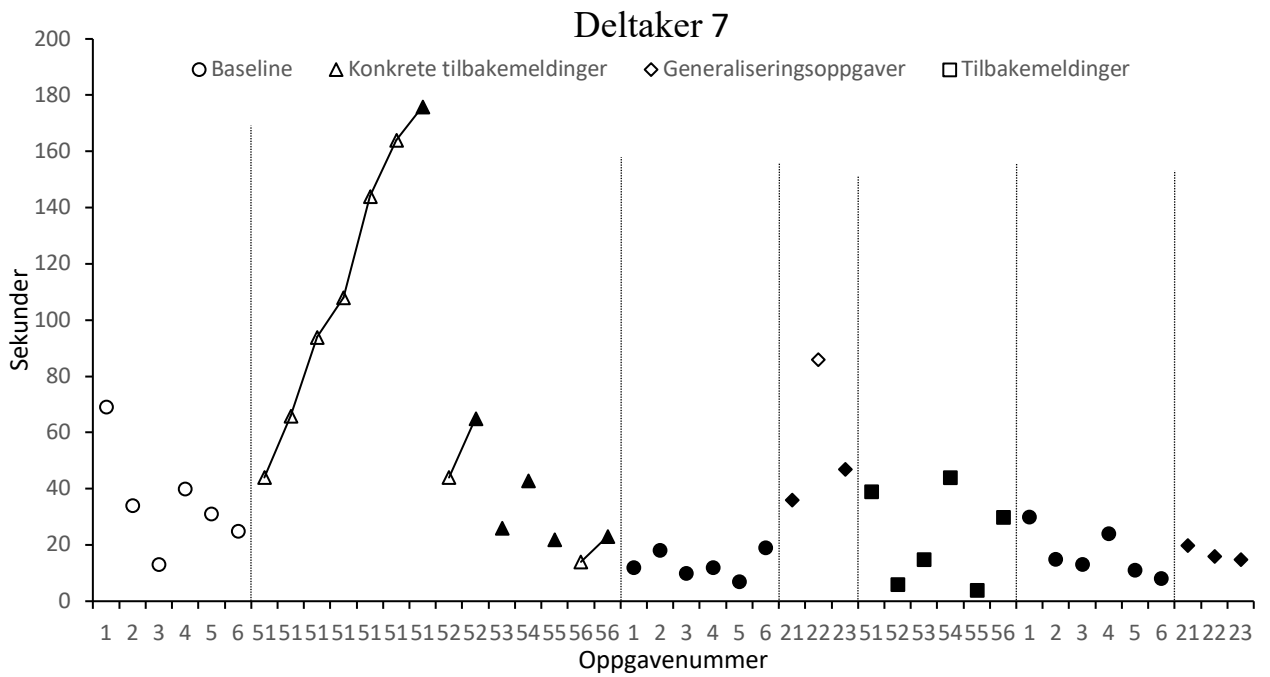
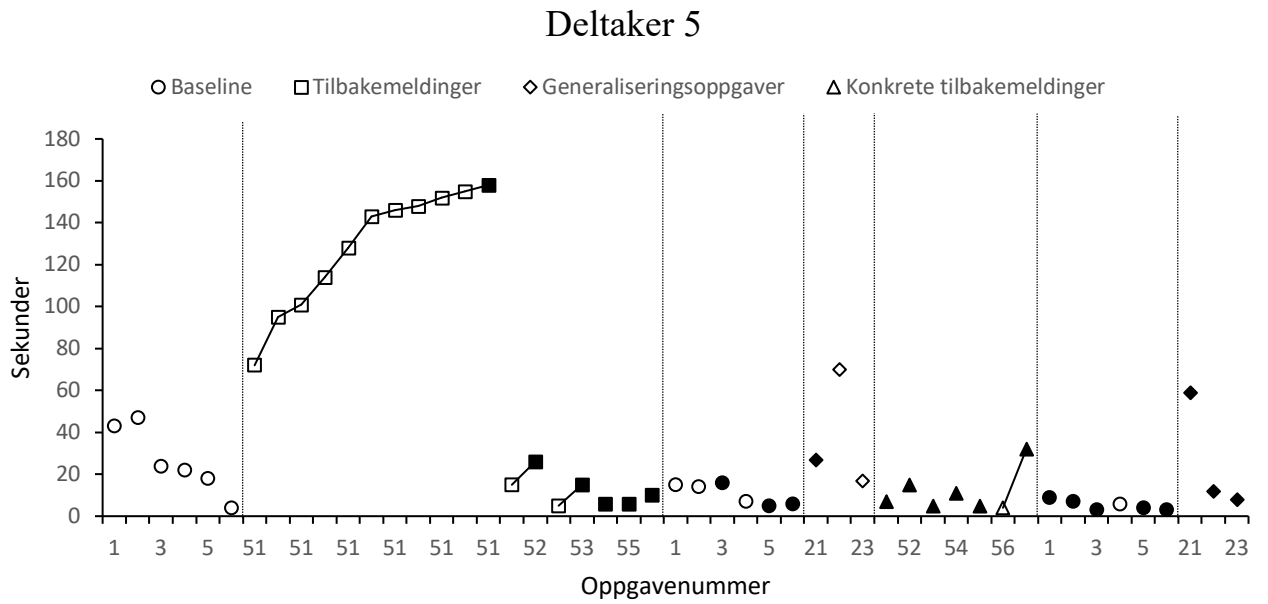
Resultater fra deltaker 3 og 4



Merknad: X-aksen illustrerer de reelle oppgavenumrene fra programvaren, og Y-aksen er sekunder. Sort indikator er korrekte responser og hvit indikator er ukorrekte responser.

Figur 5 og 6

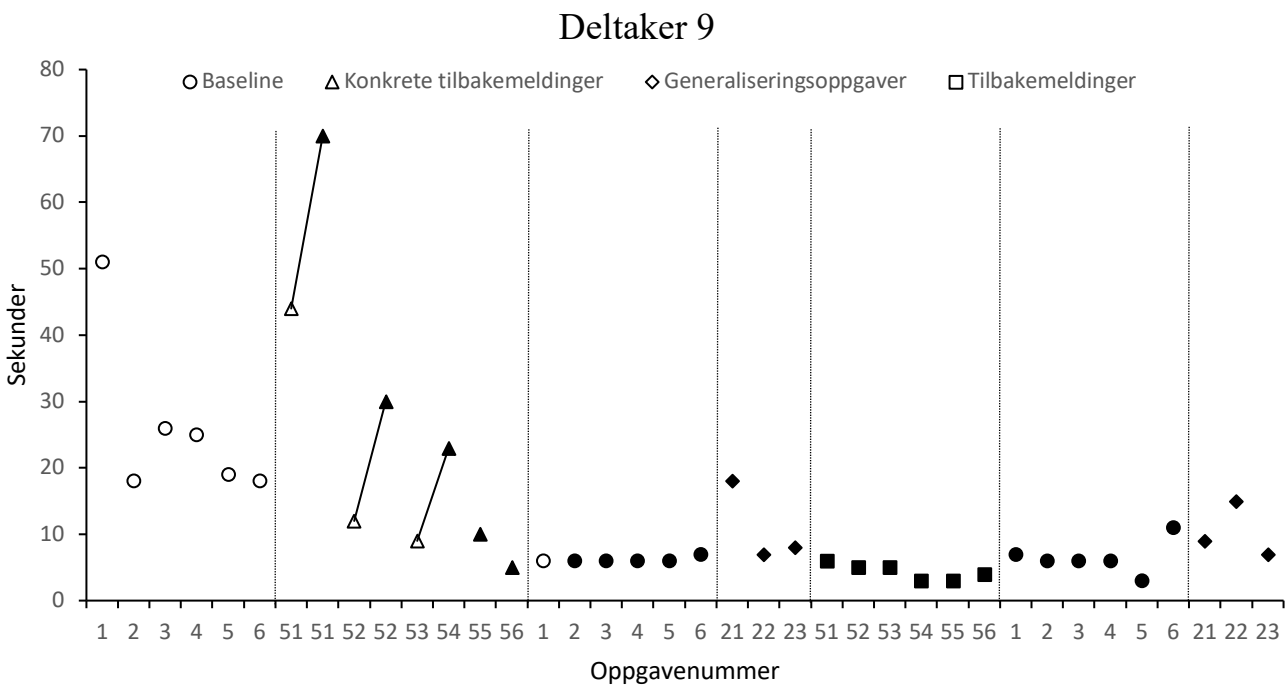
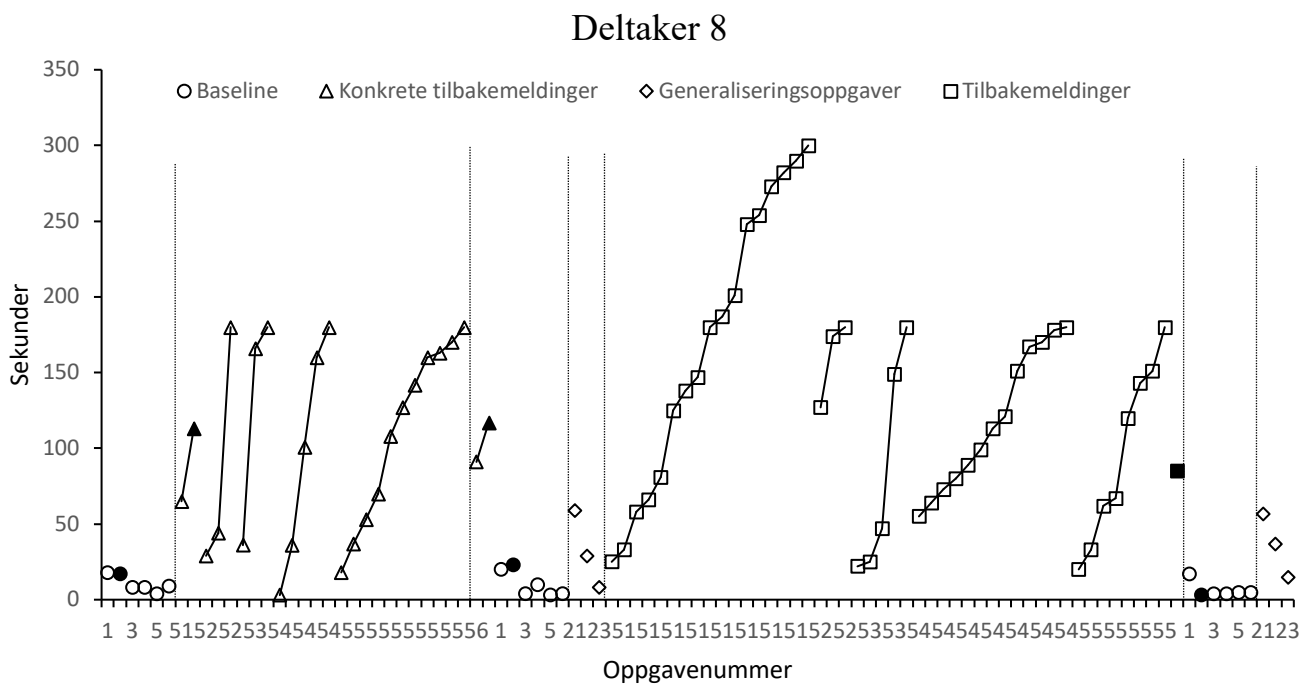
Resultater fra deltaker 5 og 7



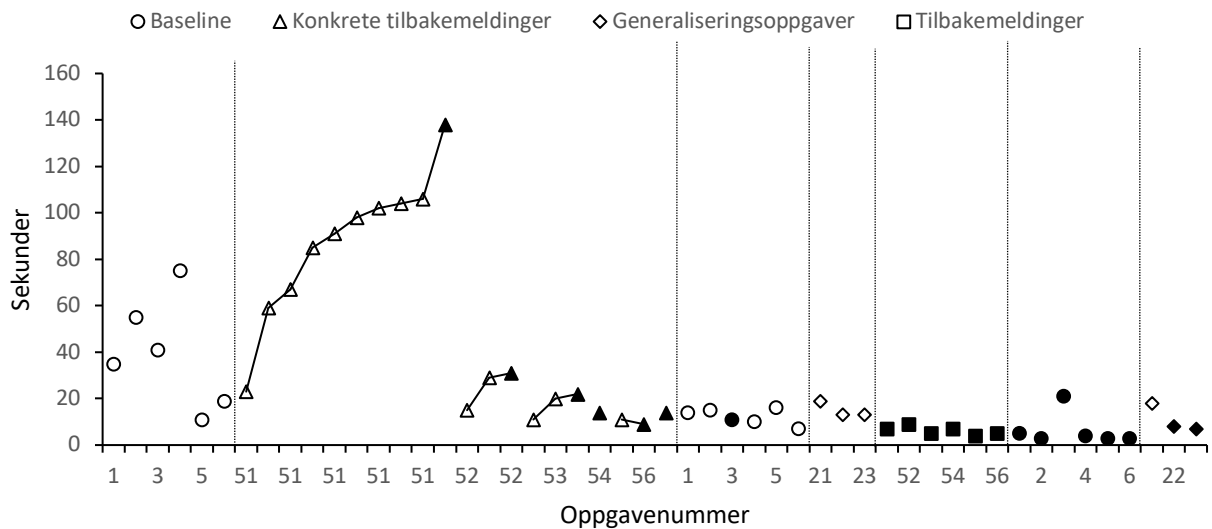
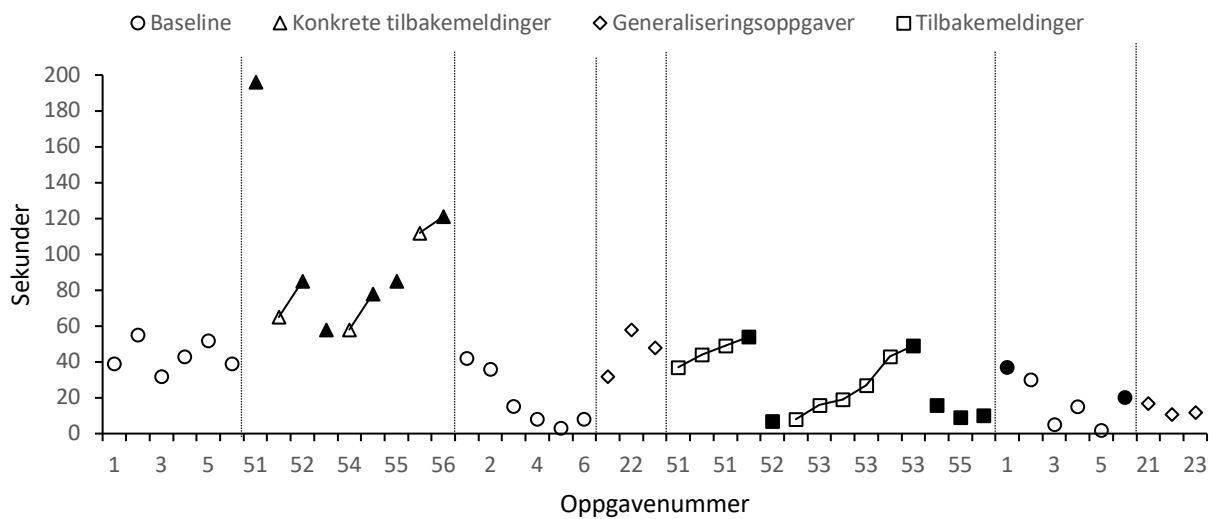
Merknad: X-aksen illustrerer de reelle oppgavenumrene fra programvaren, og Y-aksen er sekunder. Sort indikator er korrekte responser og hvit indikator er ukorrekte responser.

Figur 8 og 9

Resultater fra deltaker 8 og 9



Merknad: X-aksen illustrerer de reelle oppgavenumrene fra programvaren, og Y-aksen er sekunder. Sort indikator er korrekte responser og hvit indikator er ukorrekte responser.

Figur 10 og 11*Resultater fra deltaker 10 og 11***Deltaker 10****Deltaker 11**

Merknad: X-aksen illustrerer de reelle oppgavenumrene fra programvaren, og Y-aksen er sekunder. Sort indikator er korrekte responser og hvit indikator er ukorrekte responser.

Appendiks

Personvernerklæring- Behandling av kontaktinformasjon i forskningsprosjekter på Institutt for atferdsvitenskap

Effekten av tilbakemeldinger for å redusere bekreftelsestendensen

Formål

Innsamling og bruk av kontaktinformasjon i forbindelse med ulike forskningsprosjekter ved Institutt for atferdsvitenskap. Kontaktinformasjonen brukes til å rekruttere informanter, avtale og veilede underveis og i etterkant av prosjektene.

Spørsmål om personopplysninger

Vi registrerer og bruker disse personopplysningene: navn.

Navnet ditt blir oppbevart separert fra data som er innsamlet. Navnet blir erstattet med deltakernummer.

Ditt personvern – hvordan oppbevarer vi og bruker dine opplysninger

Vi vil bare bruke kontaktinformasjonen om deg til formålene vi her har opplyst om. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

Prosjektleder og studenter eller assistenter som deltar i forskningen har tilgang til kontaktinformasjonen.

Kontaktinformasjonen slettes når eksperimentet eller forskningsprosjektet er gjennomført for den enkelte deltaker eller formålet er oppnådd.

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg, og å få utlevert en kopi av opplysningene,
- å få rettet personopplysninger om deg,
- å få slettet personopplysninger om deg, og
- å sende klage til Datatilsynet om behandlingen av dine personopplysninger.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler kontaktinformasjon om deg basert på personvernforordningen artikkel 6 f) berettiget interesse.

Kontaktinformasjonen trengs for å kunne rekruttere, avtale og veilede underveis og i etterkant av prosjekter. Prosjektene avdekker hvordan miljøvariabler påvirker atferd, og det er nødvendig å kunne rekruttere deltakere i prosjektene. Det vil ikke medføre noen personvernulempe å delta. Kontaktinformasjonen brukes kun til dette formålet og det er mulig

å be om at kontaktinformasjonen slettes. Det vil ikke få innvirkning på relasjonen/forholdet til forsker eller utdanningen dersom du ikke vil delta i studien eller senere velger å trekke deg. Dersom du blir bedt om å svare på spørsmål i forbindelse med forskningsprosjektet, og det framgår personopplysninger om deg i svarene eller det innsamlede materialet, vil dette være et nytt formål og en annen bruk av personopplysningene. Informantene vil da bli forespurt og informert spesielt om dette i det enkelte forskningsprosjekt. Hvis prosjektene er anonyme og personopplysningene ikke framkommer i forskningsmaterialet, vil dette informasjonsskrivet om bruk av kontaktinformasjon for rekrutteringen til prosjektet være tilstrekkelig.

Ansvarlig

OsloMet- storbyuniversitetet er ansvarlig for denne behandlingen av dine personopplysninger.

Kontaktpersoner

Hvis du har spørsmål eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

Student: Svein Kevin Loe: S361561@oslomet.no

Veileder: Per Holth: Pholth@oslomet.no

Hvis du har spørsmål til personvernombudet kan du ta kontakt på

personvernombudet@oslomet.no

Med vennlig hilsen Johannes Gjerstad.

(Forsker/veileder) -----

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet og har fått anledning til å stille spørsmål.

Jeg samtykker til:

◆ å delta i et prosjekt hvor navnet mitt erstattes med et deltakernummer Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet

(Signert av prosjektdeltaker, dato)

Etisk refleksjonsnotat

ROS-vurderinger er gjennomført, referansenummer 20/10901-148.

Jeg deltok på obligatoriske forskningsetiske seminarer i forbindelse med mastergraden min, og de inneholdt nyttig informasjon om hvordan jeg kunne gjennomføre denne studien etter etiske retningslinjer. REK (Regionale komiteer for medisinsk og helsefaglig forskningsetikk) har noen retningslinjer for forskning innen medisin og helse, som innebærer pliktig rapportering i innsamlingen av personopplysninger, eller når man søker om unntak fra taushetsplikt. Dette forskningsprosjektet innebar ingen relevans for de nevnte elementer, og derfor var det unødvendig å melde denne studien til REK.

Informasjonen som samles inn i denne studien skjer gjennom bruk av et dataprogram. Dataene som samles inn inkluderer deltakernes kjønn og aldersgruppe, og de er koblet til et spesifikt øktnummer. I tilfeller som dette, er det obligatorisk for eksperimentatorer å sende inn en melding til Norsk senter for forskningsdata (NSD). I henhold til NSDs sjekkliste, innebar den nåværende studien ingen innsamling av direkte eller indirekte identifiserbare data. Det ble ikke samlet inn bilder, lydopptak eller videoer.

Eksperimentator tok kontakt med bekjente og lurte på om de ville delta i min masteroppgave. De potensielle deltakerne fikk beskjed om at det handlet om problemløsning, og at deltakelsen vil bli gjennomført på et nettbrett, og ville ta rundt 15-20 minutter.

Deltakerne måtte skrive under på et samtykkeskjema og de ble informert om at de kunne avslutte eksperimentet på hvilket som helst tidspunkt. Når deltakerne skulle starte eksperimentet måtte de fylle inn hvilken aldersklasse de tilhørte i, og kjønn.

Eksperimentatoren var til stede gjennom hele eksperimentet, dersom det skulle oppstå spørsmål underveis. Dataene som er blitt samlet inn, vil bli destruert når masteroppgaven er innlevert.