

MASTEROPPGAVE
Læring i komplekse systemer
Juni 2018

En atferdsanalytisk tilnærming til kategorisering

A behavior analytic approach to categorization

Kristiane Rustad Bevolden

Fakultet for helsefag
Institutt for atferdsvitenskap

OsloMet – storbyuniversitetet

Anerkjennelse

Først og fremst vil jeg takke Professor Erik Arntzen for all veiledning i forbindelse med denne masteroppgaven. Jeg vil også virkelig takke alle og en hver som har stilt opp som forsøkspersoner, og ikke minst Anne-Lise og Ole-Henrik Ramstad som satte huset sitt til disposisjon i forbindelse med rekruttering av deltakere utenfor Oslo. Takk til alle familiemedlemmer, venner, bekjente og kollegaer som har heiet på meg, hele veien frem til målstreken. Og sist, men ikke minst, takk til Ole-Jørgen for at du har gitt meg mulighet til å prioritere oppgaven min, og for at du har støttet meg på veien.

Innhold artikkel 1

En sammenfatning av atferdsanalytisk og kognitiv tilnærming til begrepsdanning og kategorisering

Abstrakt.....	2
Innledning.....	4
Kognitiv psykologi og kategorisering.....	4
Atferdsanalyse og kategorisering.....	7
Stimuluskontroll.....	7
Klassebegrepet og funksjonelle egenskaper ved stimulusklasser.....	7
Perseptuelle stimulusklasser.....	9
Relasjonelle stimulusklasser.....	11
Funksjonelle ekvivalensklasser.....	12
Stimulusekvivalensklasser.....	14
Etablering og testing av stimulusekvivalensklasser.....	15
Treningsstrukturer.....	16
Stimulusekvivalens og begrepsdanning.....	18
Generaliserte ekvivalensklasser og kategorisering.....	19
<i>Minimal elaborated generalized equivalence classes</i>	21
<i>Partially elaborated generalized equivalence classes</i>	21
<i>Fully elaborated generalized equivalence classes</i>	21
<i>Linked perceptual classes</i>	22
Diskusjon.....	23
Avslutning.....	26
Referanser.....	27

Innhold artikkel 2

Forholdet mellom utførelse på sorteringstester og matching-to-sample tester

Abstrakt.....	40
Innledning.....	42
Metode.....	49
Deltakere.....	49
Setting.....	50
Utstyr.....	50
Design.....	51
Prosedyre.....	52
Uforutsette hendelser.....	54
Resultater.....	55
Etablering av baseline relasjoner.....	55
Sorteringsdata.....	55
Umiddelbar emergens av klasser i Gruppe 1.....	56
MTS test data.....	57
Umiddelbar emergens av klasser i Gruppe 2.....	57
Samsvar mellom sorteringsblokkene i hver testfase.....	58
Suboptimal formasjon av klasser.....	58
Intakte baseline relasjoner i test.....	60
Responshastighet.....	60
Anekdotiske data.....	62
Diskusjon.....	63
Umiddelbar emergens av klasser.....	63
Korrespondanse i prestasjoner på MTS tester og sorteringstester.....	64

Korrespondanse i prestasjoner på sorteringstester og MTS tester.....	64
Korrespondanse mellom sorteringstestblokkene.....	65
Utsatt emergens.....	67
Forskjeller mellom testformatene.....	68
Respons hastighet.....	70
Begrensninger.....	70
Oppsummering.....	71
Referanser.....	72

Tabeller og figurer artikkel 1

Figur 1. Testbatterier for ulike typer generaliserte ekvivalensklasser.....	38
---	----

Tabeller og figurer artikkel 2

Tabell 1. Rekkefølge på betingelser for deltakere i gruppe 1 og 2.....	82
Tabell 2. Trenings- og testfase i eksperimentets betingede diskriminasjonsprosedyre.....	83
Tabell 3. Oppsummerte resultater på tvers av faser for deltakerne i Gruppe 1.....	84
Tabell 4. Oppsummerte resultater på tvers av faser for deltakerne i Gruppe 2.....	85
Tabell 5. Indeks over korrekte responser per klasse i testtrials for ekvivalens.....	86
Tabell 6. Prosentvis korrekt seleksjon av hver relasjonstype i Matching-to-sample test 1a...87	87
Figur 1. Oversikt over stimuli benyttet i eksperimentet.....	88
Figur 2. Umiddelbar formasjon av tre klasser da sortering var første testformat.....	89
Figur 3. Umiddelbar emergens av tre klasser da MTS var første testformat.....	90
Figur 4. Suboptimal formasjon av klasser.....	91
Figur 5. Resultater fra sorteringstest 1.2 hos deltaker 16027.....	92
Figur 6. Antall korrekte responser per testdel på MTS 1a hos deltaker 16027.....	93
Figur 7. Invers reaksjonstid (InvRT).....	94

Løpetittel: BEGREPSDANNING OG KATEGORISERING

En sammenfatning av atferdsanalytisk og kognitiv tilnærming til begrepsdanning og kategorisering.

Behavior analytic and cognitive approaches to concept formation and categorization

Kristiane Rustad Bevolden

OsloMet – storbyuniversitetet

Master i Læring i komplekse systemer

Juni 2018

Abstrakt

Kompleks menneskelig atferd som begrepsdanning og kategorisering er mye omtalt i litteratur innen kognitiv psykologi. De samme fenomenene er mulig å forklare med henvisning til atferdsanalytisk forskning på stimulusekvivalens. I denne artikkelen presenteres en oversikt over kognitiv og atferdsanalytisk tilnærming til kategorisering og begrepsdanning. Det blir lagt særlig vekt på hvordan fenomenene kan forstås fra et atferdsanalytisk perspektiv, og hvor begrepsdannelse og kategorisering er knyttet til grunnleggende atferdsprosesser og formasjon av forskjellige typer stimulusklasser. Det blir også gjort rede for hvordan stimulusklasser kan slås sammen og hvordan nye stimulusfunksjoner kan overføres fra et klassemedlem til resten av stimulusklassen, og på den måten illustrere hvordan kompleks begrepsdanning og kategorisering forekommer i dagliglivet.

Nøkkelord: Begrep, Kategorisering, Kompleks atferd, Stimulusklasser, Stimulusekvivalens

Abstract

Complex human behavior such as concept formation and categorization is widely discussed in cognitive psychology literature. It's also possible to explain these phenomena in behavior analytic terms, and with reference to stimulus equivalence research. This article presents an overview of cognitive and behavior analytical approaches to categorization and concept formation. Particular emphasis will be placed on how the phenomena can be understood from a behavioral analytical perspective, where concept formation and categorization is linked to basic behavioral processes and the formation of different types of stimulus classes. This article offers an explanation for how stimulus classes can merge into more complex classes and how new stimulus functions can be acquired by one stimulus class member and transfer to the other class members, thus illustrating how complex concept formation and categorization can occur in everyday life.

Keywords: Category, Concept formation, Complex behavior, Stimulus class, Stimulus equivalence

Det finnes flere uavhengige forklaringer på hva begreper og kategorier er, og hva det vil si å kategorisere og bruke begreper. Men det er til dags dato ingen unison enighet om hvordan man best kan definere hverken kategorisering eller begrepsdanning (e.g. Cohen & Lefebvre, 2017), bortsett fra at det er en relativ enighet om at kategorisering er knyttet til læring (Harnad, 2017). Hva innebærer begrepsdanning og hva innebærer det å kategorisere? Hvordan kan man etablere kategorier i opplæringsammenheng? Kompleks menneskelig atferd som begrepsdannelse og kategorisering har blitt beskrevet innen kognitiv psykologi, og blitt forklart med henvisning til kognitive prosesser og mentale strukturer (e.g. Harnad, 2017). Fra et atferdsanalytisk perspektiv blir fenomener som begrepsdanning og kategorisering sett i sammenheng med grunnleggende atferdsprosesser knyttet til formasjon av stimulusklasser—tema som inngår i stimulusekvivalensforskning—og hvor etablering av begreper og ferdigheter som kategorisering til syvende og sist handler om grad av stimuluskontroll (Arntzen, 2010; Donahoe & Palmer, 1994; Green & Saunders, 1998; Sidman, 1994, 2009). I denne artikkelen sammenstilles en oversikt over hvordan begrepsdanning og kategorisering kan forklares med en atferdsanalytisk tilnærming, og med en tilnærming basert på kognitiv psykologi. Det blir lagt spesiell vekt på hvordan stimulusklasser defineres i atferdsanalytisk litteratur, samt hvordan disse kan etableres.

Kognitiv psykologi og kategorisering

Kunnskap om kategorier i psykologi er i utgangspunktet knyttet til noen overordnede teorier. Av de mest sentrale er den klassiske teorien, prototype teori, eksemplar teori, teori-teorier og neo-empirisme teori (Laurence & Margolis, 1999; Palmer, 2002). Den klassiske teorien går ut på at begreper defineres av en struktur som inneholder en kompleks mental representasjon av nødvendige og tilstrekkelige egenskaper. Å kategorisere innebærer å resonnerer seg fram til om disse egenskapene er tilstede (Laurence & Margolis, 1999). Innenfor prototypeteori inneholder begreper mentale representasjoner av statistisk typiske

egenskaper eller karakteristikk som er avgjørende for at eksemplarer tilhører en kategori. Slike typiske eksemplarer kalles prototyper. Kategorisering innebærer å resonnerer seg frem til om f.eks. et objekt har en tilstrekkelig grad av samme karakteristikk som prototypen (Laurence & Margolis, 1999; Machery, 2009). Viktige bidragsytere innenfor prototypeteorier er blant andre (Rosch, 1973; Rosch & Mervis, 1975). Innenfor eksemplarteori antas det at vi har lagret begreper i form av sett med eksemplarer i langtidshukommelsen. Å kategorisere innebærer å beregne grad av likhet og resonnerer seg fram til korrekt eksemplar blant korrekte og ikke korrekte medlemmer av kategorien (Machery, 2009). Som en variant av eksemplarteori foreslo Medin og Schaffer (1978) det de kalte *Context Theory*, og hvor kunnskap om begreper hentes fra hukommelsen med en resonneringsmekanisme hvor ”klassifisering av nye stimuli er basert på lagret informasjon om eldre eksemplarer”, s. 210, min oversettelse. Som en videreføring av *Context theory* kom *Generalized context model of categorization* (Nosofsky, 1986). Denne modellen skal kunne brukes til å beregne dimensjonal psykologisk avstand mellom det objektet som skal kategoriseres og andre objekter for å vurdere kategoritilhørighet, og har fått stor betydning for hvordan man forstår kategorisering og begrepsdannelse innen kognitiv psykologi i nyere tid (Machery, 2009). I forbindelse med teori-teoriene oppstod et skille mellom å se på begreper som teorier— dvs. at teoriene inneholder kunnskap for å beskrive begrepene— og det å forklare begreper med henvisning til teorier (Laurence & Margolis, 1999; Machery, 2009; Palmer, 2002). Neo-empiristiske teorier har som utgangspunkt at mennesker har mekanismer for konseptuell prosessering i hjernen (Barsalou, Simmons, Barbey, & Wilson, 2003). En komplett redegjørelse for alle relevante teorier går utenfor rammene for denne artikkelen, men en oversikt over hovedteoriene er viktig for å forstå utgangspunktet for definisjoner av kategorisering og begrepsdannelse i kognitiv psykologi.

Selv om det finnes veldig mange formeninger om hva begreper og kategorisering innebærer, finnes det som nevnt innledningsvis ikke overordnede definisjoner av fenomenene. Kategorisering innebærer vanligvis å avgjøre hvilke ting som hører sammen i en kategori, og kategorien inneholder ting som hører sammen (Zentall, Galizio, & Critchfield, 2002). Begrepsdanning er knyttet til mekanismer som fasiliterer konseptuell prosessering (Barsalou et al., 2003).

”Cognitive science is concerned with the kinds of knowledge that underlie human cognition, the details of human cognitive processes, and the computational modeling of these processes” (Cohen & Lefebvre, 2017s. 1). Kategorisering er det mest grunnleggende fenomenet ved kognisjon, er fundamentet for vår kunnskap om verden, og defineres som ”en mental operasjon hvor hjernen klassifiserer objekter og hendelser” (Cohen & Lefebvre, 2017, s.1, min oversettelse). I følge Harnad (2017, s. 24) er kategorisering ”...any systematic differential interaction between an autonomous, adaptive sensorimotor system and its world”. Videre forklarer Harnad (2017) at kategorisering er en sansemotorisk ferdighet som innebærer at vi bruker sensorer og detektorer, og som gjør at vi kan forholde oss til en verden rundt oss som er i stadig endring, og hvor en sentral oppgave er analyse av sanse input, og hvor resultatet er kategorisk output.

Categories are kinds, and categorization occurs when the same output occurs with the same *kind* of input, rather than the exact same input. A different output occurs with a different kind of input. That is where the differential comes from (Harnad, 2017, s.24).

Videre skriver Harnad (2017) at vi kan ha medfødte evner til kategorisering, men at det er mange beviser for at de fleste av våre kategorier er lærte. ”Learning occurs when a system samples inputs and generates outputs in response to them on the basis of trial and error; its performance is guided by corrective feedback” (Harnad, 2017, s. 27), i betydningen at våre responser til omverdenen medfører konsekvenser. Forfatteren trekker i denne

sammenhengen frem B. F. Skinner og operant atferd, men avslutter avsnittet med at Skinner ikke forklarte hvordan operant læring kunne foregå, da han utelot å presentere en indre mekanisme for vår sansemotoriske kapasitet.

Atferdsanalyse og kategorisering

Stimuluskontroll

Atferdsanalyse tilbyr en nok så annerledes tilnærming til de samme fenomenene (Arntzen, 2010; Hayes, Barnes-Holmes, & Roche, 2001). Stimuluskontroll er en betegnelse på at en tilsiktet eller utilsiktet forsterkningsbetingelse er etablert, det vil si at det er en kontrollerende relasjon mellom spesifikke stimuli i miljøet og atferd (Urcuioli, 2013). Stimuluskontroll kan etableres ved diskriminasjonstrening, som innebærer at man differensielt forsterker noen spesifikke responser i nærvær av spesifikke stimuli. Dermed vil de spesifikke stimuliene sette anledning for og kontrollere atferd når de er tilstede. Enkel og betinget diskriminasjonstrening beskrives og analyseres som henholdsvis tre-termkontingenser og fire-termkontingenser. Enkel diskriminasjonstrening innebærer differensiell forsterkning av en respons i nærvær av en spesifikk stimulus slik at stimulusen etableres som diskriminativ stimulus (S^D). Diskriminasjon betyr i praksis at en respons forekommer med økt sannsynlighet i nærvær av—eller kontrolleres av—den diskriminative stimulusen, men ikke andre stimuli (McIlvane, 2013; Urcuioli, 2013). Betinget diskriminasjon innebærer at tre-termkontingensen kommer under kontroll av en fjerde stimulus, stimulus kondisjonale (S^C) (Sidman, 1986). Det innebærer at den diskriminative funksjonen til en stimulus er betinget av at en spesifikk S^C også er tilstede. E.g. hvilket bilde det er korrekt å peke på (hund eller katt) er betinget av hvilket dyr man blir bedt om å peke på (Urcuioli, 2013). Betinget diskriminasjon omtales ofte som hvis-så relasjoner (Sidman & Tailby, 1982) eller kontekstuell stimuluskontroll (Sidman, 1986).

Klassebegrepet og funksjonelle egenskaper ved stimulusklasser

Mens noen bruker begrepene konsepter og stimulusklasser om hverandre (e.g. Zentall et al., 2002), foretrekker andre å skille mellom stimulusklasser og konsepter slik de omtales i henholdsvis forskning og dagliglivet, og hvor kunnskap om stimulusklasser gjør det mulig å forklare ferdigheter som kategorisering og begrepsdanning (e.g. Donahoe & Palmer, 1994). Forskning på formasjon av stimulusklasser kommer opprinnelig fra to ulike forskningstradisjoner, med Clark L. Hull og Edgar C. Tolman på den ene siden, og Burrhus F. Skinner og Murray Sidman på den andre siden (Zentall & Smeets, 1996, s. vi). Både Skinner (1938), Sidman (1971, 1994), Arntzen (2012), samt Fields og Verhave (1987) har vært nøkkelpersoner når det kommer til kunnskap om henholdsvis operant atferd og stimulusekvivalens, og har bidratt til å gjøre det mulig å studere kompleks menneskelig atferd innenfor en atferdsanalytisk forskningstradisjon. Keller og Schoenfeld (1950) har fått mye oppmerksomhet i atferdsanalytisk litteratur for sin tilnærming til begrepsdanning. De understreket at i stedet for å anse konsepter som en ting, er det hensiktsmessig å se på hvilken interaksjon vi har med konsepter. "When a group of objects gets the same response, when they form a class the members of which are reacted to similarly, we speak of a concept" (Keller & Schoenfeld, 1950, s. 148). Dersom en gruppe objekter fungerer som klasse når vi reagerer likt til dem, så må vi nødvendigvis reagere på en annen måte til objekter som ikke tilhører klassen. Konseptuell atferd innebærer dermed "generalisering innen klasser og diskriminasjon mellom klasser" (Keller & Schoenfeld, 1950, s. 148, min oversettelse). For å forstå hvordan vi interagerer med konsepter må vi studere stimulusklasser og responsklasser, det vil si interaksjon mellom atferd og miljø.

Responser som inngår i responsklasser er sjelden topografisk identiske, men klassemedlemsskap avgjøres ved felles funksjon som betyr at alle responser i klassen har den samme effekten på omgivelsene. Responsklasser som kontrolleres av spesifikke stimulusklasser omtales som diskriminerte operanter (Catania, 2013; Skinner, 1938). Når

stimuli er medlemmer av en stimulusklasse betyr det at alle klassemedlemmene—minimum to stimuli—kontrollerer samme respons, det vil si at de har samme funksjon (Skinner, 1938). En respons vil dermed forekomme i nærvær av alle stimuli i samme klasse, men ikke i nærvær stimuli som ikke hører til klassen (Fields & Reeve, 2001; Keller & Schoenfeld, 1950). Videre vil alle stimuli som tilhører samme stimulusklasse sette anledning for seleksjon av de andre stimuliene i klassen, selv om det bare er etablert relasjoner mellom noen få av stimulusklassemedlemmene (Fields & Reeve, 2001; Sidman, 1994). Når stimuli inngår i stimulusklasser, er det fordi de har felles perseptuelle og/eller funksjonelle egenskaper (Donahoe & Palmer, 1994), og dermed omtales som perseptuelle, relasjonelle, funksjonelle, ekvivalente, generaliserte klasser (e.g. Fields & Reeve, 2000). Stimulusklasser kan omtales som *open-ended* eller *close-ended*. I *open-ended* klasser er medlemskap basert på perseptuelle fellestrekk blant stimuliene, og antallet medlemmer er ubegrenset (Herrnstein, 1990). *Close-ended* klasser inneholder derimot et begrenset antall stimuli i hver klasse, og hvor medlemskap ikke er basert på perseptuelle likheter mellom stimuli. Medlemskap i *close-ended* klasser kan etableres ved spesifikk trening (Fields & Reeve, 2001).

Perseptuelle stimulusklasser

Perseptuelle klasser inneholder et begrenset antall stimuli og omtales som *close-ended* (Fields, 2009; Herrnstein, 1990). Medlemskap i perseptuelle klasser er basert på fysiske likheter mellom stimuli—der stimuliene oppfattes som varianter av hverandre—og hvor likhetene kan rangeres langs fysiske, psykometriske dimensjoner eller plasseres langs et kontinuum (Fields, 2009; Fields et al., 2002; Fields & Moss, 2008; Fields & Reeve, 2000). Fields og Reeve (2000) skilte mellom tre varianter perseptuelle klasser: Dimensjonale, *fuzzy* og *polymorphous*. Dimensjonale stimulusklasser inneholder stimuli som kan rangeres langs en dimensjon (e.g. vekt, høyde, frekvens, temperatur) eller ved å kombinere egenskaper langs flere dimensjoner. I *fuzzy* klasser baseres medlemskap ikke bare på én spesifikk, men flere

undefinerte egenskaper. Det betyr at hvert medlem har noen av de undefinerte egenskapene eller oppfattet likhet, men ingen spesifikk egenskap er tilstede hos alle eksemplarene (Fields & Reeve, 2000). *Fuzzy* klasser har også blitt omtalt som *ill-defined* (Homa & Little, 1985), *natural* (Rosch, 1973; Rosch & Mervis, 1975) *basic level* og *probabilistic* (Medin & Smith, 1984). Både dimensjonale og *fuzzy* stimulusklasser inneholder et ubegrenset antall stimuli, og er dermed *open-ended* (Herrnstein, 1990). Stimuli som inngår klasser som betegnes *polymorphous* innehar et minimum antall (M) av et totalt antall (N) definerende egenskaper. Et eksempel er diagnoser som settes basert på (M) antall observerte av (N) antall spesifiserte kriterier. *Polymorphous* stimulusklasser består av et begrenset antall definerende egenskaper og hvor størrelsen på klassen avhenger av M og N (Fields & Reeve, 2000). Perseptuelle klasser har også blitt omtalt som likhetsbaserte klasser (Fields, 2016; Wasserman & Devolder, 1993; Zentall, 1996) og perseptuelle konsepter (Zentall et al., 2002)

Perseptuelle klasser har blitt etablert ved bruk av identitets matching (Fields, Reeve, Adams, Brown, & Verhave, 1997; Reeve & Fields, 2001), enkel (e.g. Cook, Wright, & Kendrick, 1990; Herrnstein, 1990) eller betinget diskriminasjonstrening (e.g. Fields, Matneja, Varelas, & Belanich, 2003). Av validitetshensyn er det viktig å presentere mange positive og negative stimuluseksemplarer ved etablering og testing av perseptuelle klasser (Fields & Reeve, 2000). Etablering av perseptuelle stimulusklasser ved enkel diskriminasjonstrening innebærer at man forsterker en respons i nærvær av noen varianter av stimulus A, men ikke i nærvær av varianter av stimulus B. Etter trening presenteres tovalgs (*forced-choice*) eller trevalgs primær generaliseringstester. *Forced choice* innebærer at en utvalgsstimulus presenteres med to sammenligningsstimuli, og hvor forsøkspersonen må velge mellom en av dem, uavhengig av om forsøkspersonen oppfatter at utvalgsstimulus og sammenligningsstimulus har noen fysisk likhet. Dermed vet man ikke med sikkerhet om seleksjonen skyldes tvungent valg eller klassetilhørighet i en perseptuell klasse. Det har derfor

blitt argumentert for å benytte et tredje valg i form av en såkalt *neither* stimulus, og som innebærer at ingen av sammenligningsstimuliene er relatert til utvalgsstimulus (Reeve & Fields, 2001). Dataene fra generaliseringstestene kan presenteres som generaliseringsgradienter som illustrerer hvilke stimuli som kontrollerer en spesifikk respons, og dermed inngår i den perseptuelle klassen (Catania, 2013).

Klassemedlemsskap i perseptuelle klasser er basert på primær stimulusgeneralisering, som betyr at nye stimuli får kontroll over responser fordi de nye stimuliene oppfattes som perseptuelt like de andre medlemmene i en stimulusklasse (e.g. Arntzen, 2010; Fields, 2009; Fields & Reeve, 2000; Green & Saunders, 1998). For å kunne konkludere med at stimuli hører til en perseptuell klasse må det for det første demonstreres at en respons som vekkes i nærvær av noen få stimuli i den perseptuelle klassen, også vekkes i nærvær av de andre perseptuelle variantene i klassen. De perseptuelt like stimuliene må ha samme diskriminative funksjon (S^D). For det andre må man vise at responsen forekommer med redusert sannsynlighet i nærvær av stimuli som ikke har de samme perseptuelle likhetene, slik at stimuli utenfor klassen har funksjon som stimulus delta (S^Δ) (Reeve & Fields, 2001). Igjen er det snakk om ”generalisering innen klasser og diskriminasjon mellom klasser” (Keller & Schoenfeld, 1950, s. 155, min oversettelse). For det tredje må det være mulig å diskriminere mellom stimuliene innen klassen slik at man vet at responderingen er basert på en relasjon til en stimulusklasse, og ikke bare på en relasjon til kun én stimulus (Fields & Reeve, 2000; Reeve & Fields, 2001). Formasjon av en perseptuell klasse er dokumentert når responsen som var trent til noen positive eksemplarer også forekommer i nærvær av de nye positive eksemplarene, eller når nye klassemedlemmer blir selektert utelukkende i nærvær av andre klassemedlemmer (Fields, 2009).

Relasjonelle stimulusklasser

I stedet for at medlemskap i stimulusklasser er basert på perseptuelle egenskaper, kan medlemskapet være basert på abstrakte egenskaper og relasjoner mellom medlemmene (Zentall et al., 2002). Relasjonelle stimulusklasser kan være elementspesifikke eller elementuavhengige. Elementspesifikke stimulusklasser kan for eksempel bestå av stimuli i form av varianter av sirkel i firkant (Fields & Reeve, 2000). I så fall vil stimuli som har sirkel i firkant inngå i klassen, mens trekant i firkant eller firkant i sirkel tilhører ikke klassen.

Medlemskap i elementuavhengige klasser defineres av en spesifikk type relasjon mellom stimuliene i klassen. Eksempler på slike relasjoner kan være større enn og mindre enn, refleksivitet, symmetri og transitivitet m.m. (Fields & Reeve, 2000; Zentall et al., 2002). Egenskapene refleksivitet, symmetri og transitivitet kan etableres som relasjonelle klasser ved multipel eksemplar trening og testing, i en *Matching-to-sample* (MTS) prosedyre. I motsetning til eksempelvis tester for stimulusekvivalensklasser, kan relasjonene i relasjonelle klasser trenes direkte dersom de ikke fremkommer ved testing. Dette skyldes at tester for relasjonelle klasser innebærer generalisering av respondering til nye, utrente stimulussett (Fields & Reeve, 2000).

Funksjonelle ekvivalensklasser

En mye omtalt studie i forbindelse med funksjonelle ekvivalensklasser er en studie gjennomført av Vaughan (1988). I eksperimentet organiserte Vaughan (1988) en enkel diskriminasjonsprosedyre med trening av mange eksemplarer, og hvor hensikten var å undersøke om duer kunne etablere stimulusekvivalensklasser. Totalt 40 bilder av trær ble fordelt på to sett, – sett 1 og sett 2. Bildene ble presentert suksessivt, dvs. at et og et bilde ble presentert av gangen, og i blandet rekkefølge, to ganger per dag i et repetert reverseringsdesign. Hakking på en tast i nærvær av bilder fra sett 1 ble etterfulgt av tilgang på forsterker (mat), mens hakking i nærvær av bilder fra sett 2 ble ikke etterfulgt av tilgang på mat. Dette førte til at duene lærte å hakke i nærvær av bilder fra sett 1, men ikke i nærvær av

bilder fra sett 2. Da stabil respondering var etablert til bilder i sett 1, ble forsterkningsbetingelsene byttet om, slik at hakking kun i nærvær av bilder fra sett 2 ble etterfulgt av mat. Da stabil respondering var oppnådd til bilder i sett 2 ble forsterkningsbetingelsene igjen byttet om på. Forsterkningsbetingelsene ble endret frem og tilbake mellom de to bildesettene gjentatte ganger, og Vaughan (1988) oppdaget at da forsterkningsbetingelsene endret seg, så ble også sannsynligheten for hvilke stimuli duene hakket i nærvær av endret. Det betyr at de første stimuliene i ett sett endret den diskriminative funksjonen for hakking i nærvær av resten av bildene i settet. Vaughan (1988) tolket funnene som at duene hadde etablert to stimulusekvivalensklasser med arbitrære stimuli. Det ble i ettertid påpekt at det var problematisk å kalle stimulusklassene i eksperimentet for stimulusekvivalensklasser, uten at tester for egenskapene refleksivitet, symmetri og transitivitet—som definerer stimulusekvivalens—var gjennomført (Dougher & Markham, 1996; Sidman, 1994; Sidman & Tailby, 1982). Derfor har Vaughan (1988) studien ofte blitt trukket frem som et eksempel på etablering av funksjonelle ekvivalensklasser eller kontingensklasser (Sidman, 1994).

Hva som skiller funksjonelle ekvivalensklasser og stimulusekvivalensklasser er ikke helt avklart (Eilertsen & Arntzen, 2015), men har blitt mye diskutert (e.g. Sidman, 1994; Sidman, Wynne, Maguire, & Barnes, 1989; Tonneau, 2001). Diskusjonene har i grove trekk handlet om man kan skille mellom funksjonelle ekvivalensklasser og stimulusekvivalensklasser basert på hva slags læring som ligger til grunn for klassene, e.g. operant eller respondent læring (Tonneau, 2001), basert på ulike funksjoner hos klassemedlemmene og/eller basert på forskjeller med hensyn til overføring av stimulusfunksjoner. Sidman (1994) skrev at medlemsskap i funksjonelle klasser avhenger av at medlemmene setter anledning for samme respons, og at ikke-medlemmer ikke har den samme funksjonen (Keller & Schoenfeld, 1950). I tillegg er det avgjørende at helt nye

stimulusfunksjoner trent til et medlem, overføres til de andre medlemmene av klassen. Testen for klassemedlemskap i funksjonelle ekvivalensklasser innebærer dermed å undersøke om nye ulike stimuli frembringer samme respons, og på den måten er gjensidig utskiftbare i funksjon (Dougher & Markham, 1996; Tonneau, 2001). Til sammenligning er testen for klassemedlemskap i stimulusekvivalensklasser knyttet til demonstrasjon av egenskapene refleksivitet, symmetri og transitivitet (Sidman & Tailby, 1982). Når det gjelder overføring av stimulusfunksjoner som ikke er direkte trent, så har diskusjonen handlet om overføringen skjer via eller uavhengig av stimulusekvivalensklasser (Eilertsen & Arntzen, 2015, 2017; Sidman, 1994, 2000; Tonneau, 2001).

Stimulusekvivalensklasser

Stimulusekvivalensklasser består av et begrenset antall (*close-ended*) fysisk eller psykometrisk ulike stimuli, som har blitt ekvivalente—i betydningen gjensidig utskiftbare i funksjon—etter betinget diskriminasjonstrening hvor det har blitt trent eksplisitt på relasjoner mellom kun noen av klassemedlemmene (Fields & Arntzen, 2017; Green & Saunders, 1998; Sidman, 1994). Stimulusekvivalensklasser har og blitt omtalt som assosiative konsepter (Zentall et al., 2002), *superordinate semantic categories* (Astley, Peissig, & Wasserman, 2001; Rosch & Mervis, 1975) og som konseptuelle kategorier (Fields & Arntzen, 2017).

Hvorvidt stimuli i en stimulusekvivalensklasse har blitt gjensidig utskiftbare i funksjon kan i følge Sidman og Tailby (1982) bare avgjøres ved å teste for egenskapene som definerer stimulusekvivalens: Refleksivitet, symmetri og transitivitet. Refleksivitet innebærer at hver stimulus er relatert til seg selv. Det betyr at det er etablert en betinget relasjon (R) hvor hver stimulus selekteres i nærvær av seg selv som utvalgsstimulus. E.g. hvis A så A, hvis B så B og hvis C så C osv. Symmetri innebærer at det er en bidireksjonal betinget relasjon (R) mellom to forskjellige stimuli, slik at hvis det er en betinget relasjon mellom stimuliene A og B (aRb), som henholdsvis utvalgs- og sammenligningsstimulus, så er det også en betinget

relasjon mellom B og A (bRa) som henholdsvis utvalgs- og sammenligningsstimulus.

Transitivitet innebærer at det fremkommer en betinget relasjon mellom A og C (aRc) på bakgrunn av at det er trent en betinget relasjon mellom A og B (aRb), samt mellom B og C (bRc) (Sidman & Tailby, 1982).

Et mye omtalt eksempel på formasjon av stimulusekvivalensklasser er basert på en studie gjennomført av Sidman (1971), og hvor hensikten var å lære en 17 år gammel utviklingshemmet gutt leseforståelse med trening i et *matching-to-sample* format. Gutten kunne fra før matche bilde (visuell stimulus) til ordlyd (auditiv stimulus), samt uttale ordlyd (vokal stimulus) til bilde (visuell stimulus). Sidman (1971) og kollegaer trente gutten i å matche ordlyd til skrevet ord. Da gutten ble presentert for tester viste det seg at han uten å ha direkte trent på det, hadde lært seg å matche skrevet ord til bilde, og da bilder ble presentert kunne han si de riktige navnene (Sidman, 1971). Skrevet ord, ordlyd og bilder hadde blitt substitutter eller gjensidig utskiftbare i funksjon (Green & Saunders, 1998).

Etablering og testing av stimulusekvivalensklasser. Formasjon av stimulusekvivalensklasser trenes og testes i form av betinget diskriminasjon mellom stimuli i et MTS format¹ (Sidman, 1992, 1994), men har også blitt presentert som en form for respondent trening med suksessiv presentasjon av stimuluspar, etterfulgt av en stimulusekvivalenstest (Leader, Barnes, & Smeets, 1996; Leader, Barnes-Holmes, & Smeets, 2000; Smeets, Leader, & Barnes, 1997). For å etablere betinget diskriminasjon i MTS format må relasjoner trenes mellom minimum to utvalgsstimuli og to sammenligningsstimuli (R. R. Saunders & Green, 1999), men for å kunne teste for stimulusekvivalens må man bruke minimum to klasser med tre medlemmer i hver klasse (Fields & Verhave, 1987; Sidman & Tailby, 1982).

¹ Både betinget diskriminasjon og MTS brukes om hverandre som prosedyrer, men Sidman (1994); Sidman et al. (1982) presiserte at kun MTS prosedyrer kan demonstrere *sann* MTS, som innebærer ekvivalensrelasjoner, og som er noe mer enn betingede hvis/så relasjoner.

MTS prosedyrer kan innebære arbitrær matching, det vil si at det ikke er fysiske likheter mellom utvalgs- og sammenligningsstimuli, eller identisk matching hvor utvalgs- og sammenligningsstimuli er helt like (Urcuioli, 2013). Noen studier har vist at arbitrær MTS medfører økt sannsynlighet for emergens av ekvivalensklasser, sammenlignet med identitets MTS (e.g. Arntzen, Nartey, & Fields, 2018). Det skiller også mellom simultane (SMTS) og utsatte (DMTS) MTS prosedyrer. SMTS prosedyrer innebærer at utvalgsstimulus er til stede samtidig som sammenligningsstimuli presenteres, og DMTS innebærer at det er et gitt tidsintervall fra utvalgsstimulus fjernes til sammenligningsstimuli presenteres (Arntzen, 2006). Flere studier har funnet en økning i respondering i henhold til stimulusekvivalens som funksjon av at tidsintervallet øker, noe som kan indikere at en form for støtteatferd (*precurrent behavior*) forekommer i dette tidsrommet, og som kan være en viktig variabel ved formasjon av ekvivalensklasser (Arntzen, 2006; Arntzen, Galaen, & Halvorsen, 2007; Arntzen et al., 2018; Lian & Arntzen, 2013; Vie & Arntzen, 2017). DMTS prosedyrer har også blitt benyttet i studier av hukommelse (e.g. Steingrimsdottir & Arntzen, 2014).

Treningsstrukturer. De mest brukte treningsstrukturene i stimulusekvivalensforskning er lineær serie (LS), *many-to-one* (MTO) og *one-to-many* (OTM) (e.g. Fields, Hobbie-Reeve, Adams, & Reeve, 1999; K. J. Saunders, Saunders, Williams, & Spradlin, 1993). Ved LS trener man $A \rightarrow B \rightarrow C$, hvor relasjonen mellom A og C ikke trenes direkte. B stimuli har en dobbel funksjon ved å være sammenligningsstimuli for medlemmer av A, og utvalgsstimuli for medlemmer av C under trening. Stimuli som er relatert til to stimuli under trening, slik som B, går under betegnelsen noder. Ved MTO trener man noen få sammenligningsstimuli til mange utvalgsstimuli, e.g. AB og CB, og ved bruk av OTM trener man mange sammenligningsstimuli til noen få utvalgsstimuli, e.g. AB og AC (Fields & Verhave, 1987; K. J. Saunders et al., 1993). Ved bruk av MTO kan man teste for emergens av relasjon mellom AC og CA, og med OTM kan man teste for emergens av

relasjon mellom BC og CB. Det bare mulig å teste for sann matching eller global ekvivalens ved bruk av LS treningsstruktur (Sidman, 1992, 1994). Vanligvis benyttes MTS tester for å demonstrere emergens av stimulusekvivalensklasser, men det har også blitt vanligere å presentere sorteringstester før, etter, alene eller i kombinasjon med MTS tester i stimulusekvivalensforskning. Fordelen er at sorteringstester er betydelig raskere å administrere, og at man får dokumentert at ekvivalensklasser etablert i MTS format, ikke er begrenset til det formatet (e.g. Arntzen, Granmo, & Fields, 2017; Fields & Arntzen, 2017).

En del studier har funnet få forskjeller med hensyn til antall deltakere som responderer i henhold til stimulusekvivalens når man bruker MTO og OTM treningsstruktur (Arntzen, Lian, & Halstadtrø, 2011; Eilifsen & Arntzen, 2015; Vie & Arntzen, 2017). LS treningsstruktur i kombinasjon med en simultan protokoll—hvor alle baselinereelasjoner trenes samtidig i blandet rekkefølge (e.g. Imam, 2006)—blir benyttet mye i forskning på stimulusekvivalens fordi kombinasjonen reduserer sannsynligheten for formasjon av ekvivalensklasser, og vil derfor ikke fungere som en potensiell tredjevariabel (Arntzen et al., 2018).

I MTS prosedyrer må forsøkspersonen både diskriminere mellom suksessivt presenterte utvalgsstimuli på tvers av trials² (suksessiv diskriminasjon), og mellom samtidig presenterte sammenligningsstimuli innen hver trial (simultan diskriminasjon). Dersom sammenligningsstimuli presenteres suksessivt, det vil si en og en, kalles det suksessiv *matching* (R. R. Saunders & Green, 1999; Urcuioli, 2013). Det har blitt diskutert om forskjeller i prestasjoner ved bruk av forskjellige treningsstrukturer skyldes forskjeller i antall trente enkle diskriminasjoner mellom treningsstrukturene (R. R. Saunders & Green, 1999), eller om forskjellene er knyttet til antall suksessive og simultane diskriminasjoner, og hvor

² Trial er en beskrivelse av en presentasjonsrunde med utvalgs- og sammenligningsstimuli, og benyttes som et godkjent ord på norsk da det ikke finnes noe norsk ord med samme betydning.

vidt det spiller en rolle om enkelte stimuli har funksjon som både utvalgs- og sammenligningsstimuli i form av noder (Arntzen, 2012). En mengde forskning har ført til kunnskap om hvilke variabler som påvirker formasjon av stimulusekvivalensklasser (se e.g. Arntzen, 2012; Arntzen et al., 2018).

Stimulusekvivalens og begrepsdanning. Stimulusekvivalens er relevant i forbindelse med fenomener som bl.a. blir omtalt som språk og kognisjon (e.g. Arntzen, 2010; Sidman, 1994; Zentall et al., 2002). Stimulusekvivalensforskning gir mulighet til å studere atferd som ikke har vært direkte eksponert for forsterkningsbetingelser—det vil si emergent eller generativ atferd—og er spesielt viktig for å forstå begrepsdanning og kategorisering. Begreper eller konseptuelle kategorier kan forstås som en klasse som inneholder ekvivalente stimuli, det vil si at de har samme funksjon (Arntzen, 2010; Keller & Schoenfeld, 1950). En type funksjon kan være betydning (*meaningfulness*), som og er beskrevet som semantikk (Sidman, 1994). Forskning på stimulusekvivalensklasser har bl.a. vist at inkludering av en stimulus som har betydning utenfor eksperimentet, det vil si fra før, øker sannsynligheten for emergens av ekvivalensklasser (e.g. Arntzen et al., 2018). I tillegg kan noen få betydningsfulle (definert som mer eller mindre attraktive) stimuli i en klasse med ellers nøytrale stimuli overføres til de andre klassemedlemmene og fasilitere klasseformasjon (Fields & Arntzen, 2017). Andre betydningsfulle funksjoner kan være symbolske. Begreper kan inneholde ord som refererer til ting, og vi kan begynne å oppføre oss mot ord eller symboler som om de *er* det vi oppfatter at de symboliserer. Sidman (1994) omtalte dette som ord-referent ekvivalens.

Adams, Fields, og Verhave (1993) påpekte at det er to kriterier som må oppfylles for å kunne si at en stimulusklasse fungerer som en kategori: For det første må man etablere en ekvivalensklasse og teste for emergente relasjoner, e.g. A1B1C1 etablert med en LS treningsstruktur. For det andre må man se at en ny funksjon trent til ett av medlemmene i ekvivalensklassen (e.g. ny respons i nærvær av A1) også overføres til de andre medlemmene i

ekvivalensklassen uten spesifikk trening. Overføring via symmetri vil i denne sammenhengen innebære at den nye responsen også utvises i nærvær av B1. Overføring via transitivitet eller ekvivalens innebærer at den nye responsen også forekommer i nærvær av C1 (Eilertsen & Arntzen, 2015). Dersom den nye funksjonen overføres til alle medlemmene av klassen, har man etablert en kategori (Dougher & Markham, 1996; Dymond & Rehfeldt, 2000; Eilertsen & Arntzen, 2017). Dougher og Markham (1996) presiserte her et skille mellom de emergente stimulusfunksjonene som definerer ekvivalensklassen og overføring av en ny stimulusfunksjon via ekvivalensklassen som kan skje i tillegg.

Flere atferds funksjoner har blitt dokumentert overført via ekvivalensklasser. F.eks. viste Dougher, Augustson, Markham, Greenway, og Wulfert (1994) at det via ekvivalensklasser var mulig å overføre en betinget fryktrespons i form av fuktighet målt på huden, og ekstinksjon av den samme fryktresponsen. Hayes, Kohlenberg, og Hayes (1991) demonstrerte overføring av betinget forsterkning og straff via ekvivalensklasser ved å presentere sammenligningsstimuli sammen med verbal tilbakemelding.

Generaliserte ekvivalensklasser og kategorisering

Kunnskap om generaliserte ekvivalensklasser tilbyr en måte å forstå etablering av komplekse kategorier slik de forekommer utenfor forskning i dagliglivet, blant annet fordi kategorimedlemmene kan bestå av både like og ulike stimulusmodaliteter (Fields & Reeve, 2001). De fleste stimuli inngår ofte i flere ekvivalensklasser samtidig (Sidman, 1994), og både ekvivalensklasser og kategorier i dagliglivet inneholder ofte stimuli som oppfattes perseptuelt like og ulike. E.g. kan en person sett fra flere vinkler, stemmen til den samme personen hørt fra forskjellige avstander og navnet til personen skrevet digitalt og med forskjellig håndskrift inngå i samme klasse (Fields & Moss, 2008).

Generaliserte ekvivalensklasser inneholder noen perseptuelt ulike stimuli som inngår i en ekvivalensklasse, samt noen stimuli som er perseptuelt like og som inngår i en *open-ended*

klasse. Fordi én stimulus fungerer som medlem i begge klassene, blir samtlige stimuli relaterte til hverandre i en generalisert ekvivalensklasse, og alle medlemmene av den generaliserte ekvivalensklassen vil føre til seleksjon av medlemmer i samme klasse (Adams et al., 1993; Fields & Reeve, 2001). Sammenslåingen av den basale ekvivalensklassen og den perseptuelle klassen kan foregå uten trening, eller etableres ved å arrangere en betinget diskriminasjonsprosedyre (Fields & Reeve, 2001). Ved etablering av generaliserte ekvivalensklasser er det to atferdsprosesser som virker sammen: formasjon av ekvivalensklasser og primær stimulusgeneralisering (Adams et al., 1993). For det første etableres en stimulusekvivalensklasse med tre medlemmer, eg. bilde av jente (A1), ordet jente skrevet på norsk (B1) og ordet *mädchen* skrevet på tysk (C1). Dersom en primær generaliseringstest presenteres og viser at varianter av A1 (A1') også selekteres i nærvær av B1 og C1, har man påvist at en generalisert ekvivalensklasse er etablert. Klassen vil da inneholde stimuli både fra den basale (opprinnelige) ekvivalensklassen, og medlemmer fra den perseptuelle klassen med varianter av bilder (A'). Stimuliene som inngår i den basale ekvivalensklassen omtales som ankerstimuli, og den varianten av hver ankerstimulus som befinner seg perseptuelt lengst unna ankerstimulusen, men som hører til den perseptuelle klassen, kalles grense eller *boundary (b)* stimulus. Grensestimulus representerer dermed ytterpunktet for klassetilhørighet. Den stimulusen som befinner seg perseptuelt midt mellom anker og grensestimulus kalles midtpunkt eller *midpoint (m)* (Fields et al., 2002; Fields & Reeve, 2001). I denne artikkelen blir alle perseptuelle varianter av stimuli markert med en apostrof (').

Fields og Reeve (2001) skilte mellom tre typer generaliserte ekvivalensklasser, avhengig av hvor mange perseptuelle stimulusklasser som inngår i de generaliserte ekvivalensklassene: *Minimal-, partially- og fully elaborated generalized equivalence classes*. De samme forfatterne beskrev også tre typer teststrukturer eller testbatterier for å kunne

demonstrere emergens av de ulike typene generaliserte ekvivalensklassene: *Variant-to-base*, *base-to-variant* og *variant-to-variant* tester. Med *variant-to-base* testes det for emergente relasjoner mellom samtlige ankerstimuli i ekvivalensklassen til varianter av en gitt utvalgsstimulus. Med *base-to-variant* testes det for emergente relasjoner mellom varianter av en gitt sammenligningsstimulus til samtlige ankerstimuli som utvalgsstimuli. Og med *variant-to-variant* testes det for emergente relasjoner mellom varianter av både utvalgsstimuli og varianter av sammenligningsstimuli.

Minimal elaborated generalized equivalence classes (MEG). En MEG ekvivalensklasse inneholder varianter av et medlem av den basale ekvivalensklassen, e.g. A1, B1 og C1 i form av en *close-ended* ekvivalensklasse, i tillegg til varianter av e.g. C1' som medlemmer av en *open-ended* perseptuell klasse. For å undersøke om varianter av C1 fungerer som medlemmer kan man bruke en *variant-to-base* test og en *base-to-variant* test. Se Figur 1 for en oversikt over testtrials for ulike generaliserte egenskaper. Dersom det demonstreres at samtlige C1' stimuli setter anledning for seleksjon av alle andre medlemmer av både den opprinnelige ekvivalensklassen og medlemmer av *open-ended* klassen, heter det at alle medlemmene inngår i en MEG ekvivalensklasse (Fields & Reeve, 2001).

Partially elaborated generalized equivalence classes (PEG). En PEG ekvivalensklasse inneholder medlemmer av en basal *close-ended* ekvivalensklasse (e.g. A1, B1 og C1) og varianter av to eller tre basale klassemedlemmer, i dette tilfellet eksempelvis varianter av A1 og C1 som i tillegg inngår i *open-ended* perseptuelle klasser. Ved testing av PEG klasser kan man bruke alle testbatteriene. For en full oversikt se Figur 1.

Fully elaborated generalized equivalence classes (FEG). En FEG ekvivalensklasse inneholder varianter av alle medlemmene av en basal ekvivalensklasse. Det vil si at i tillegg til ankerstimuliene A1, B1 og C1 inngår varianter av dem (A1', B1' og C1') i klassen. Når man skal teste for en FEG klasse, gitt at det er varianter av både A, B og C, må også alle

testbatteriene benyttes. Med *variant-to-base* testes variantene av både A', B' og C' som utvalgsstimuli mot anker som sammenligningsstimuli. Med *base-to-variant* testes anker som utvalgsstimuli mot variantene av A', B' og C' som sammenligningsstimuli, og med *variant-to-variant* testes variantene av A', B' og C' som utvalgsstimuli mot variantene av A', B' og C' som sammenligningsstimuli (Fields, 2009; Fields & Reeve, 2001).

Linked perceptual classes (LP). LP klasser innebærer at to forskjellige perseptuelle klasser—av samme eller forskjellig sansemodalitet—inngår i en klasse. For eksempel kan mange bilder av en mors ansikt kan inngå i en felles klasse med mange lyder av den samme morens stemme (Fields et al., 2002). LP klasser som inneholder stimuli av forskjellige sansemodaliteter, har også blitt kalt kryssmodale klasser. Det har blitt foreslått at man ved å studere atferdsprosessene ved etablering av LP klasser kan forklare fenomener som kryssmodal persepsjon, intersensorisk integrasjon og intermodal persepsjon (Fields, 2009; Fields et al., 2002).

LP klasser kan slås sammen i større klasser ved å trene betinget diskriminasjon mellom medlemmer av to perseptuelle klasser, på samme måte som en ekvivalensklasse og en perseptuell klasse kan slås sammen til generaliserte ekvivalensklasser. Fields (2009) beskrev LP klasser som en form for generaliserte ekvivalensklasser, hvor ankerstimuli i to forskjellige perseptuelle klasser fungerer som noder. Fields et al. (2005) undersøkte fire forskjellige testskjema og fant at å benytte et såkalt 18/1-PRGM skjema, som innebærer å organisere 18 blokker med kun en type test *probe* i hver blokk, presentert i en systematisk rekkefølge—i motsetning til å presentere alle test *probes* blandet og tilfeldig fra starten—fasiliterte emergens av LP klasser. Fields et al. (2007) introduserte flere måter å trene betinget diskriminasjon mellom perseptuelle klasser på, enten ved å trene en eller to relasjoner mellom to perseptuelle klasser, og fant at å trene to relasjoner økte sannsynligheten for sammenslåing av klassene. For eksempel ved å trene anker fra en klasse som utvalgsstimulus til anker fra

den andre klassen som sammenligningsstimulus, og tilsvarende trene relasjon mellom grensestimuli i de to klassene. I en studie av Fields og Garruto (2009) viste det seg at å trene en betinget relasjon mellom anker fra en klasse (Aa) som utvalgsstimulus og grensestimulus fra en annen klasse (Bb) som sammenligningsstimulus, og å teste med et 18/1-PRGM skjema fasiliterte formasjon av LP klasser.

Generaliserte ekvivalensklasser kan også fungere som funksjonsoverførende nettverk ved at en ny funksjon trent til noen få av medlemmene i klassen, overføres til de andre medlemmene uten at det behøver å trenes inn. Belanich og Fields (2003) og Fields og Garruto (2009) fant at henholdsvis MEG klasser og LP klasser kan fungere som nettverk for overføring av stimulusfunksjoner.

Diskusjon

Hovedutfordringen med kunnskap om begrepsdanning og kategorisering er at det eksisterer veldig mange ulike tilnærminger innen psykologi, basert på flere overordnede teorier, modeller og ulike definisjoner (Laurence & Margolis, 1999; Zentall et al., 2002). Det har fra flere hold vært pekt på utfordringer ved samtlige av dem (Laurence & Margolis, 1999; Palmer, 2002), blant annet fordi teoriene ikke inngår i et helhetlig rammeverk med en konsistent terminologi (Dougher & Markham, 1996; Machery, 2010). Frem til tidlig 2000-tallet var det ikke vært uvanlig at en type stimulusklasse ble omtalt under flere navn, eller at samme navn ble brukt på ulike typer stimulusklasser (Dougher & Markham, 1996). Det var heller ikke noen klare skiller mellom kategorier og begreper slik de er definert i prototype teorier, og slik de ble definert i atferdsanalytisk litteratur i form av perseptuelle klasser. Det har blitt hevdet at noe av årsaken til problemene med terminologien er at forskere har behandlet beskrivelser av stimulusklasser slik de tradisjonelt er definert i psykologi, og stimulusekvivalensklasser som fremkommer under trening og testing, som to uavhengige fenomener i stedet for som aspekter ved ett fenomen (Fields & Reeve, 2000). Forskning på

stimulusekvivalens har bidratt til å etablere kunnskap om flere ulike stimulusklasser, og på bakgrunn av atferdsanalytisk forskningsmetodikk og et sammenhengende begrepsapparat, er man nå i stand til å bidra med forklaringer på flere fenomener innen kompleks atferd som i lengre tid først og fremst har vært omtalt innen fagfeltet kognitiv psykologi (Arntzen et al., 2018; Sidman, 1994).

Forskjellene mellom den kognitive og atferdsanalytiske tilnærmingen til atferd generelt, handler om ulike perspektiver på hva som kan kalles vitenskap eller kunnskap. Kognitiv psykologi er basert på et mekanistisk rammeverk, mens atferdsanalyse er basert på et pragmatisk kontekstuellet rammeverk. Av dette følger noen forskjeller i både forskningsspørsmål, forskningsmetodikk og hvor man plasserer årsaken til atferd (Dougher, 1995; Palmer, 2002). I kognitiv psykologi er sannhetskriteriet for forklaring av atferd knyttet til hvor vidt en beskrivelse og prediksjon av atferd korresponderer med observerte hendelser. I atferdsanalyse er målet beskrivelse og påvirkning/kontroll, og hvor en gyldig årsaksforklaring innebærer at man må kunne demonstrere reliabelt at endringer i én variabel er årsaken til endringer i en annen variabel (Dougher, 1995). Det overordnede synet i atferdsanalyse er at atferd etableres, opprettholdes og endres som ledd i seleksjonsprosesser (e.g. Donahoe, 2017; Donahoe & Palmer, 1994). Atferd forklares med henvisning til funksjonelle relasjoner mellom klasser av stimuli og klasser av responser, og hvor fokuset utelukkende er på observerbare hendelser. (Arntzen, 2010; Donahoe & Palmer, 1994; Skinner, 1938).

Harnad (2017) etterlyste henvisning til en indre mekanisme for sansemotorisk kapasitet i forklaring av læring, noe som er problematisk fra et atferdsanalytisk perspektiv på atferd. Dersom kontrollerende betingelser blir plassert på innsiden av individet i form av mentale strukturer, blir muligheten for vitenskapelige undersøkelser basert på prediksjon og kontroll umulig (Skinner 1938; Dougher 1995). Likevel er det ikke uforståelig fra et atferdsanalytisk perspektiv at man ønsker å trekke inn sanseceller og nervesystemet i

sammenheng med atferd (e.g. Sidman, 2008). Skinner (1950) erkjente at det forekommer fysiologiske endringer i nervesystemet når vi lærer, eksempelvis endringer i synaptiske forbindelser og overføring av signaler mellom nervecellene, men han stilte spørsmålsteget ved om det var *nødvendig* å henvise til nervesystemet i en analyse av atferd, spesielt når målet er å predikere og kontrollere atferd. Skinner (1950) argumenterte med at variablene vi inkluderer i en årsaksforklaring av atferd må være på samme observasjonsnivå som atferden vi analyserer. Det er med andre ord viktig at vi ikke henviser til en teori på et annet (ofte utilgjengelig) observasjonsnivå, og som vi igjen må forklare for å kunne påvirke og kontrollere atferd (Holth, 2001). I stedet for å definere læring med henvisning til teorier om eksempelvis hjerneaktivitet—og spesielt hvis teoriene omhandler hypotetiske konstrukter—argumenterte Skinner (1950) for å anse læring som endring i responsrate som funksjon av endring i forsterkningsbetingelser. Kort oppsummert kan man si at læring er den atferdsendringen som kan observeres når vi endrer miljøbetingelser, og som påvirker sannsynlighet for fremtidig forekomst av tilsvarende atferd i lignende situasjoner. Se for eksempel Zilio (2013) for en oversikt over noe av det Skinner har skrevet om atferdsanalyse og nevrovitenskap.

Kunnskap om begrepsdanning og kategorisering fra stimulusekvivalensforskning er spesielt nyttig i forbindelse med opplæringssituasjoner. I den siste tiden har noen forskere begynt å undersøke ekvivalensbasert instruksjon (EBI) som opplæringsmetode (Barron, Leslie, & Smyth, 2018; Varelas & Fields, 2017). Barron et al. (2018) presenterte et EBI program på nettbrett med berøringsskjerm for å lære fem normalt fungerende 4 åringer aldersadekvate kategorier som barna ikke kunne på forhånd. Siden barn stadig tidligere blir introdusert for elektronisk utstyr i form av mobiltelefoner og nettbrett, og flere skoler legger opp til bruk av nettbrett i sin pedagogikk, bør det være gode muligheter for å undersøke effektivitet av EBI som undervisningsmetode videre. Under navn som *Zero-shot translation* og begrep som *transfer learning*, ser det ut til at datateknologi som for eksempel Google

Translate er basert på treningsstrukturer som ligner de vi finner i stimulusekvivalensforskning (Johnson et al., 2016). Det er grunn til å tro at kunnskap om blant annet begrepsdannelse og kategorisering fra stimulusekvivalensforskning kan bli etterspurt i forbindelse med utvikling av ny teknologi som innebærer kunstig intelligens. Det kan få både praktiske og etiske implikasjoner.

Avslutning

Kategorisering og begrepsdannelse i kognitiv psykologi blir forklart med henvisning til ulike strukturelle aspekter og egenskaper hos medlemmene i kategoriene (Rosch, 1973; Rosch & Mervis, 1975), og hvor menneskets evne til å etablere og utvikle begreper skjer via sansemotorikk og informasjon bearbeidet og lagret i langtidshukommelsen (Barsalou et al., 2003; Harnad, 2017). Innen det atferdsanalytiske fagfeltet er man opptatt av hva man kan observere at individer gjør når man sier at de kategoriserer eller etablerer begreper (Keller & Schoenfeld, 1950). Etablering av samtlige stimulusklasser kan forklares med henvisning til grunnleggende atferdsprosesser som stimulusgeneralisering, enkel eller betinget diskriminasjon og generalisering av responsrepertoarer (Catania, 2013; Donahoe & Palmer, 1994; Fields & Reeve, 2000).

Noe av budskapet til Fields og Reeve (2000) var at man ved å benytte multipl eksemplar trening og testing, kan etablere stimulusklasser som inneholder medlemmer som både er perseptuelt like og ulike, og at flere stimulusklasser kan kombineres til større stimulusklasser. Komplekse, sammensatte stimulusklasser kan modellere komplekse kategorier slik vi kjenner dem fra verden utenfor forskningslaboratoriet. Forskning på stimulusekvivalensklasser tilbyr også en måte å forklare menneskers evne til tilpasning til nye situasjoner: Når vi demonstrerer hvordan nye stimulusfunksjoner kan fremkomme uten at de er direkte trent på, viser vi også hvordan mennesker slipper å trene eksplisitt på alle begreper for å lære dem (Arntzen, 2010; Fields & Moss, 2008; Fields & Reeve, 2001).

Referanser

- Adams, B. J., Fields, L., & Verhave, T. (1993). Formation of generalized equivalence classes. *The Psychological Record*, *43*, 553–566. doi:10.1007/BF03395899
- Arntzen, E. (2006). Delayed matching to sample: Probability of responding in accord with equivalence as a function of different delays. *The Psychological Record*, *56*(1), 135–167. doi:10.1007/BF03395541
- Arntzen, E. (2010). Om stimulusekvivalens. In S. Eikeseth & F. Svartdal (Eds.), *Anvendt atferdsanalyse: Teori og praksis* (2 ed., pp. 100–138). Oslo: Gyldendal Norsk Forlag AS.
- Arntzen, E. (2012). Training and testing parameters in formation of stimulus equivalence: Methodological issues. *European Journal of Behavior Analysis*, *13*(1), 123–135. doi:10.1080/15021149.2012.11434412
- Arntzen, E., Galaen, T., & Halvorsen, L. R. (2007). Different retention intervals in delayed matching-to-sample: Effects of responding in accord with equivalence. *European Journal of Behavior Analysis*, *8*(2), 177–191. doi:10.1080/15021149.2007.11434281
- Arntzen, E., Granmo, S., & Fields, L. (2017). The relation between sorting test and matching-to-sample tests in the formation of equivalence classes. *The Psychological Record*, *67*(1), 81–96. doi:10.1007/s40732-016-0209-9
- Arntzen, E., Lian, T., & Halstadtrø, L. (2011). Anvendelse av matching-to-sample prosedyrer i etableringen av akademiske ferdigheter *Norsk Tidsskrift for Atferdsanalyse*, *38*(1), 1–26. Hentet fra <http://www.nta.atferd.no/journalissue.aspx?IdDocument=207>
- Arntzen, E., Nartey, R. K., & Fields, L. (2018). Graded delay, enhanced equivalence class formation, and meaning. *The Psychological Record*, *68*(2), 123–140. doi:10.1007/s40732-018-0271-6

- Astley, S. L., Peissig, J. J., & Wasserman, E. A. (2001). Superordinate categorization via learned stimulus equivalence: Quantity of reinforcement, hedonic value, and the nature of the mediator. *Journal of experimental psychology: Animal behavior processes*, 27(3), 252–268. doi:10.1037//0097-7403.27.3.252
- Barron, R., Leslie, J. C., & Smyth, S. (2018). Teaching real-world categories using touchscreen equivalence-based instruction. *The Psychological Record*, 68(1), 89–101. doi:10.1007/s40732-018-0277-0
- Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in cognitive sciences*, 7(2), 84–91. doi:10.1016/S1364-6613(02)00029-3
- Belanich, J., & Fields, L. (2003). Generalized equivalence classes as response transfer networks. *The Psychological Record*, 53, 373–413. Hentet fra <http://opensiuc.lib.siu.edu/tpr/vol53/iss3/3/>
- Catania, A. C. (2013). *Learning* (5 ed.). New York, NY: Sloan Publishing.
- Cohen, H., & Lefebvre, C. (2017). Bridging the category divide: Introduction to the first edition. In H. Cohen & C. Lefebvre (Eds.), *Handbook of Categorization in Cognitive Science* (2 ed., pp. 1–17). San Diego: Elsevier.
- Cook, R. G., Wright, A. A., & Kendrick, D. F. (1990). Visual categorization by pigeons. In M. J. Commons, R. J. Herrnstein, S. M. Kosslyn, & D. B. Mumford (Eds.), *Quantitative analysis of behavior: Behavioral approaches to pattern recognition and concept formation* (Vol. VIII, pp. 187–214). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Donahoe, J. W. (2017). Behavior analysis and neuroscience: Complementary disciplines. *Journal of the experimental analysis of behavior*, 107(3), 301–320. doi:10.1002/jeab.251

- Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior*. Boston: Allyn & Bacon.
- Dougher, M. J. (1995). A bigger picture: Cause and cognition in relation to differing scientific frameworks. *Journal of Behavior Therapy and Experimental Psychiatry*, 26(3), 215–219. doi:10.1016/0005-7916(95)00021-Q
- Dougher, M. J., Augustson, E., Markham, M. R., Greenway, D. E., & Wulfert, E. (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of experimental analysis of behavior*, 62(3), 331–351. doi:10.1901/jeab.1994.62-331
- Dougher, M. J., & Markham, M. R. (1996). Stimulus classes and the untrained acquisition of stimulus functions. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus Class Formation in Humans and Animals* (pp. 137–152). Amsterdam, The Netherlands: Elsevier Science B.V.
- Dymond, S., & Rehfeldt, R. A. (2000). Understanding complex behavior: The transformation of stimulus functions. *The Behavior Analyst*, 23(2), 239–254. Hentet fra <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2731337/>
- Eilertsen, J. M., & Arntzen, E. (2015). Aspekter ved forskning på stimulusklasser i og utenfor laboratoriet. *Norsk Tidsskrift for Atferdsanalyse*, 42(2), 65–78. Hentet fra <http://hdl.handle.net/10642/3138>
- Eilertsen, J. M., & Arntzen, E. (2017). Overføring av stimulusfunksjoner *Norsk Tidsskrift for Atferdsanalyse*, 44(2), 33–49. Hentet fra <http://www.nta.atferd.no/journalissue.aspx?IdDocument=678>
- Eilifsen, C., & Arntzen, E. (2015). Effects of training structure and the passage of time on trained and derived performance. *The Psychological Record*, 65(1), 1–12. doi:10.1007/s40732-014-0067-2

- Fields, L. (2009). The synthesis of complex categories from perceptual and equivalence classes: Effects of training and testing parameters. *European Journal of Behavior Analysis, 10*(2), 205–227. doi:10.1080/15021149.2009.11434314
- Fields, L. (2016). Stimulus relatedness in equivalence classes, perceptual categories, and semantic memory networks. *European Journal of Behavior Analysis, 17*(1), 2–18. doi:10.1080/15021149.2015.1084713
- Fields, L., & Arntzen, E. (2017). Meaningful stimuli and the enhancement of equivalence class formation. *Perspectives on Behavior Science*. doi:10.1007/s40614-017-0134-5
- Fields, L., Fitzer, A., Shamoun, K., Matneja, P., Watanabe, M., & Tittelbach, D. (2005). The effect of test schedules on the formation of linked perceptual classes. *Journal of the experimental analysis of behavior, 84*(2), 243–267. doi:10.1901/jeab.2005.45-03
- Fields, L., & Garruto, M. (2009). Optimizing linked perceptual class formation and transfer of function *Journal of the experimental analysis of behavior, 91*(2), 225–251. doi:10.1901/jeab.2009.91-225
- Fields, L., Hobbie-Reeve, S. A., Adams, B. J., & Reeve, K. F. (1999). Effects of training directionality and class size on equivalence class formation by adults. *The Psychological Record, 49*(4), 703–724. doi:10.1007/BF03395336
- Fields, L., Matneja, P., Varelas, A., & Belanich, J. (2003). Mutual selection and membership in open-ended classes: Variant-to-base and base-to-variant testing. *The Psychological Record, 53*(2), 287–311. doi:10.1007/BF03395446
- Fields, L., Matneja, P., Varelas, A., Belanich, J., Fitzer, A., & Shamoun, K. (2002). The formation of linked perceptual classes. *Journal of the experimental analysis of behavior, 78*(3), 271–290. doi:10.1901/jeab.2002.78-271

- Fields, L., & Moss, P. (2008). Formation of partially and fully elaborated generalized equivalence classes. *Journal of the experimental analysis of behavior*, *90*(2), 135–168. doi:10.1901/jeab.2008.90-135
- Fields, L., & Reeve, K. F. (2000). Synthesizing equivalence classes and natural categories from perceptual and relational classes. In J. C. Leslie & D. Blackman (Eds.), *Experimental and applied analysis of human behavior* (pp. 59–83). Reno, NV: Context press.
- Fields, L., & Reeve, K. F. (2001). A methodological integration of generalized equivalence classes, natural categories, and cross-modal perception. *The Psychological Record*, *51*(1), 67–87. doi:10.1007/BF03395387
- Fields, L., Reeve, K. F., Adams, B. J., Brown, J. L., & Verhave, T. (1997). Predicting the extension of equivalence classes from primary generalization gradients: The merger of equivalence classes and perceptual classes. *Journal of the experimental analysis of behavior*, *68*(1), 67–91. doi:10.1901/jeab.1997.68-67
- Fields, L., Tittelbach, D., Shamoun, K., Watanabe, M., Fitzer, A., & Matneja, P. (2007). The effects of differential training procedures on linked perceptual class formation. *Journal of the experimental analysis of behavior*, *87*(1), 97–119. doi:10.1901/jeab.2007.10-06
- Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of experimental analysis of behavior*, *48*(2), 317–332. doi:10.1901/jeab.1987.48-317
- Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 229–262). New York, NY: Plenum Press.

- Harnad, S. (2017). To cognize is to categorize: Cognition is categorization. In H. Cohen & C. Lefebvre (Eds.), *Handbook of Categorization in Cognitive Science* (2 ed., pp. 21–54). San Diego: Elsevier.
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). Relational Frame Theory: A précis. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 141–155). New York, NY: Kluwer Academic/Plenum.
- Hayes, S. C., Kohlenberg, B. S., & Hayes, L. J. (1991). The transfer of specific and general consequential functions through simple and conditional equivalence relations. *Journal of the experimental analysis of behavior*, 56(1), 119–137. doi:10.1901/jeab.1991.56-119
- Herrnstein, R. J. (1990). Levels of stimulus control: A functional approach. *Cognition*, 37, 133–166. doi:10.1016/0010-0277(90)90021-B
- Holth, P. (2001). The Persistence of category mistakes in psychology. *Behavior and Philosophy*, 29, 203–219. Hentet fra <http://www.jstor.org/stable/27759428>
- Homa, D., & Little, J. (1985). The abstraction and long-term retention of ill-defined categories by children. *Bulletin of the Psychonomic Society*, 23(4), 325–328. doi:10.3758/BF03330172
- Imam, A. A. (2006). Experimental control of nodality via equal presentations of conditional discriminations in different equivalence protocols under speed and no-speed conditions. *Journal of the experimental analysis of behavior*, 85(1), 107–124. doi:10.1901/jeab.2006.58-04
- Johnson, M., Schuster, M., Quoc, V. L., Krikun, M., Wu, Y., Chen, Z., & Thorat, N. (2016). Google's multilingual neural machine translation system: Enabling zero-shot translation. Hentet fra <https://arxiv.org/abs/1611.04558>

- Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of psychology; A systematic text in the science of behavior*. New York: Appleton–Century–Crofts.
- Laurence, S., & Margolis, E. (1999). Concepts and cognitive science. In S. Laurence & E. Margolis (Eds.), *Concepts: Core readings*. Cambridge, MA, US: The MIT Press.
- Leader, G., Barnes, D., & Smeets, P. M. (1996). Establishing equivalence relations using a respondent-type training procedure. *The Psychological Record*, *46*(4), 685–706.
doi:10.1007/BF03395192
- Leader, G., Barnes-Holmes, D., & Smeets, P. M. (2000). Establishing equivalence relations using a respondent-type training procedure III. *The Psychological Record*, *50*(1), 63–78. doi:10.1007/BF03395343
- Lian, T., & Arntzen, E. (2013). Delayed matching-to-sample and linear series training structures. *The Psychological Record*, *63*(3), 545–561.
doi:10.11133/j.tpr.2013.63.3.010
- Machery, E. (2009). *Doing without concepts: Why cognition is not unified*. USA - OSO: Oxford University Press.
- Machery, E. (2010). Précis of doing without concepts. *Behavioral and brain sciences*, *33*, 195–244. doi:10.1017/S0140525X09991531
- McIlvane, W. J. (2013). Simple and complex discrimination learning. In G. J. Madden, W. V. Dube, T. D. Hackenberg, G. P. Hanley, & K. A. Lattal (Eds.), *APA handbook of behavior analysis, Vol. 2. Translating principles into practice* (pp. 129–163). Washington, DC, US: American Psychological Association.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning *Psychological Review*, *85*(3), 207–238. doi:10.1037/0033-295X.85.3.207
- Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology*, *35*, 113–138. doi:10.1146/annurev.ps.35.020184.000553.

- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of experimental psychology: General*, *115*(1), 39–57.
doi:10.1037/0096-3445.115.1.39
- Palmer, D. C. (2002). Psychological essentialism: A review of E. Margolis and S. Laurence (eds.), *Concepts: Core readings*. *Journal of experimental analysis of behavior*, *78*(3), 597–607. doi:10.1901/jeab.2002.78-597
- Reeve, K. F., & Fields, L. (2001). Perceptual classes established with forced-choice primary generalization tests and transfer of function. *Journal of the experimental analysis of behavior*, *76*(1), 95–114. doi:10.1901/jeab.2001.76-95
- Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, *4*(3), 328–350.
doi:10.1016/0010-0285(73)90017-0
- Rosch, E. H., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. doi:10.1016/0010-0285(75)90024-9
- Saunders, K. J., Saunders, R. R., Williams, D. C., & Spradlin, J. E. (1993). An interaction of instructions and training design on stimulus class formation: Extending the analysis of equivalence *The Psychological Record*, *43*(4), 725–744. doi:10.1007/BF03395909
- Saunders, R. R., & Green, G. (1999). A discrimination analysis of training-structure effects on stimulus equivalence outcomes. *Journal of the experimental analysis of behavior*, *72*(1), 117–137. doi: 10.1901/jeab.1999.72-117
- Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, *14*, 5–13. doi:10.1044/jshr.1401.05
- Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and Integration of Behavioral Units* (pp. 213–245). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Sidman, M. (1992). Equivalence relations: Some basic considerations. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 15–27). Reno, NV: Context Press.
- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative, Inc.
- Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the experimental analysis of behavior*, 74(1), 127–146. doi:10.1901/jeab.2000.74-127
- Sidman, M. (2008). Reflections on stimulus control. *The Behavior Analyst*, 31(2), 127–135. Hentet fra <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2591753/>
- Sidman, M. (2009). Equivalence relations and behavior: An introductory tutorial. *The Analysis of Verbal Behavior*, 25(1), 5–17. doi:10.1007/BF03393066
- Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *Journal of the experimental analysis of behavior*, 37(1), 23–44. doi:10.1901/jeab.1982.37-23
- Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the experimental analysis of behavior*, 37(1), 5–22. doi:10.1901/jeab.1982.37-5
- Sidman, M., Wynne, C. K., Maguire, W., & Barnes, T. (1989). Functional classes and equivalence relations. *Journal of the experimental analysis of behavior*, 52(3), 261–274. doi:10.1901/jeab.1989.52-261
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis* (R. M. Elliot Ed.). New York: Appleton-Century-Crofts, Inc.
- Skinner, B. F. (1950). Are theories of learning necessary? *The Psychological Review*, 57(4), 193–216. doi:10.1037/h0054367

- Smeets, P. M., Leader, G., & Barnes, D. (1997). Establishing stimulus classes in adults and children using a respondent-type training procedure: A follow-up study. *The Psychological Record*, *47*(2), 285–308. doi:10.1007/BF03395226
- Steingrimsdottir, H. S., & Arntzen, E. (2014). Performance by older adults on identity and arbitrary matching-to-sample tasks. *The Psychological Record*, *64*(4), 827–839. doi:10.1007/s40732-014-0053-8
- Tonneau, F. (2001). Equivalence relations: A critical analysis. *European Journal of Behavior Analysis*, *2*(1), 1–33. doi:10.1080/15021149.2001.11434165
- Urcuioli, P. J. (2013). Stimulus control and stimulus class formation. In G. J. Madden, W. V. Dube, T. D. Hackenberg, G. P. Hanley, & K. A. Lattal (Eds.), *APA handbook of behavior analysis, Vol. 1. Methods and principles* (pp. 361–386). Washington, DC, US: American Psychological Association.
- Varelas, A., & Fields, L. (2017). Equivalence based instruction by group based clicker training and sorting tests. *The Psychological Record*, *67*(1), 71–80. doi:10.1007/s40732-016-0208-x
- Vaughan, W. (1988). Formation of equivalence sets in pigeons. *Journal of experimental psychology: Animal behavior processes*, *14*(1), 36–42. doi:10.1037/0097-7403.14.1.36
- Vie, A., & Arntzen, E. (2017). Talk-aloud protocols during conditional discrimination training and equivalence class formation. *The Analysis of Verbal Behavior*, *33*(1), 80–97. doi:10.1007/s40616-017-0081-y
- Wasserman, E. A., & Devolder, C. L. (1993). Similarity- and nonsimilarity-based conceptualization in children and pigeons. *The Psychological Record*, *43*(4), 779–793. doi:10.1007/BF03395912

- Zentall, T. R. (1996). An analysis of stimulus class formation in animals. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus Class Formation in Humans and Animals* (pp. 15–34). Amsterdam, The Netherlands: Elsevier Science
- Zentall, T. R., Galizio, M., & Critchfield, T. S. (2002). Categorization, concept learning, and behavior analysis: An introduction. *Journal of the experimental analysis of behavior*, 78(3), 237–248. doi:10.1901/jeab.2002.78-237
- Zentall, T. R., & Smeets, P. M. (Eds.). (1996). *Stimulus class formation in humans and animals*. Amsterdam, The Netherlands: Elsevier Science B.V.
- Zilio, D. (2013). Filling the gaps: Skinner on the role of neuroscience in the explanation of behavior. *Behavior and Philosophy*, 41, 33–59. Hentet fra <http://www.jstor.org/stable/behaphil.41.33>

Minimal elaborated generalized equivalence class

<i>Variant-to-base</i>	<i>Base-to-variant</i>	
	B1/ <u>C1'</u> C2 C3	Generalisert baseline
C1'/ <u>C1</u> C2 C3	C1/ <u>C1'</u> C2 C3	Primær generalisering
C1'/ <u>B1</u> B2 B3		Generalisert symmetri
	A1/ <u>C1'</u> C2 C3	Generalisert transitivitet
C1'/ <u>A1</u> A2 A3		Generalisert ekvivalens

Partially elaborated generalized equivalence class

<i>Variant-to-base</i>	<i>Base-to-variant</i>	<i>Variant-to-variant</i>	
A1'/ <u>B1</u> B2 B3	B1/ <u>C1'</u> C2 C3		Generalisert baseline
A1'/ <u>A1</u> A2 A3	C1/ <u>C1'</u> C2 C3	A1'/ <u>A1'</u> A2 A3	Primær generalisering
C1'/ <u>C1</u> C2 C3	A1/ <u>A1'</u> A2 A3	C1'/ <u>C1'</u> C2 C3	Primær generalisering
C1'/ <u>B1</u> B2 B3	B1/ <u>A1'</u> A2 A3		Generalisert symmetri
A1'/ <u>C1</u> C2 C3	A1/ <u>C1'</u> C2 C3	A1'/ <u>C1'</u> C2 C3	Generalisert transitivitet
C1'/ <u>A1</u> B2 A3	C1/ <u>A1'</u> A2 A3	C1'/ <u>A1'</u> A2 A3	Generalisert ekvivalens

Fully elaborated generalized equivalence class

<i>Variant-to-base</i>	<i>Base-to-variant</i>	<i>Variant-to-variant</i>	
A1'/ <u>B1</u> B2 B3	A1 / <u>B1'</u> B2 B3	A1'/ <u>B1'</u> B2 B3	Generalisert baseline
B1'/ <u>C1</u> C2 C3	B1 / <u>C1'</u> C2 C3	B1'/ <u>C1'</u> C2 C3	Generalisert baseline
A1'/ <u>A1</u> A2 A3	A1 / <u>A1'</u> A2 A3	A1'/ <u>A1'</u> A2 A3	Primær generalisering
B1'/ <u>B1</u> B2 B3	B1 / <u>B1'</u> B2 B3	B1'/ <u>B1'</u> B2 B3	Primær generalisering
C1'/ <u>C1</u> C2 C3	C1 / <u>C1'</u> C2 C3	C1'/ <u>C1'</u> C2 C3	Primær generalisering
B1'/ <u>A1</u> A2 A3	B1 / <u>A1'</u> A2 A3	B1'/ <u>A1'</u> A2 A3	Generalisert symmetri
C1'/ <u>B1</u> B2 B3	C1 / <u>B1'</u> B2 B3	C1'/ <u>B1'</u> B2 B3	Generalisert symmetri
A1'/ <u>C1</u> C2 C3	A1 / <u>C1'</u> C2 C3	A1'/ <u>C1'</u> C2 C3	Generalisert transitivitet
C1'/ <u>A1</u> A2 A3	C1 / <u>A1'</u> A2 A3	C1'/ <u>A1'</u> A2 A3	Generalisert ekvivalens

Figur 1. Testbatterier for ulike typer generaliserte ekvivalensklasser ved sammenslåing av en stimulusekvivalensklasse bestående av ankerstimuliene A1, B1 og C1, og perseptuelle varianter markert med enkel apostrof ('). Utvalgsstimulus presenteres før skråstrek, og sammenligningsstimulus som er korrekt å selektere er markert med understrek. Type testtrial er beskrevet i høyre kolonne. Merk at testtrial *probes* for baseline relasjoner i forbindelse med emergens av basale ekvivalensklasser er utelatt fra denne oversikten. Figuren er basert på Fields og Reeve (2001).

Løpetittel: FORHOLDET MELLOM SORTERINGSTESTER OG MTS TESTER

Forholdet mellom utførelse på sorteringstester og matching-to-sample tester

The relation between performance on sorting tests versus matching-to-sample
tests

Kristiane Rustad Bevolden

Oslo Met – storbyuniversitetet

Mastergrad i Læring i komplekse systemer

Juni 2018

Abstrakt

Denne studien er en systematisk replikasjon av Arntzen, Granmo, og Fields (2017). Ved bruk av en *Many-to-one* treningsstruktur og simultan protokoll i et *matching-to-sample* format, ble 12 betingede diskriminasjoner trent hos 22 forsøkspersoner i alderen 24–60 år. Det ble så testet for om tre ekvivalensklasser med fem medlemmer framkom. Deltakere i Gruppe 1 ble eksponert for en fase med trening av baseline relasjoner, to blokker med sorteringstester, to blokker med Matching-to-sample (MTS) tester, og to blokker med sorteringstester.

Deltakerne i Gruppe 2 ble eksponert for en fase med trening av baseline relasjoner, to blokker med MTS tester, to blokker med sorteringstester, og to blokker med MTS tester. I Gruppe 1 viste syv deltakere umiddelbar emergens av tre stimulusklasser på begge sorteringsblokkene, og fire deltakere viste andre responsmønstre. I Gruppe 2 viste ni deltakere umiddelbar formasjon av tre klasser på de to blokkene med MTS tester, og en deltaker viste et annet responsmønster. Tjue av 22 deltakere viste formasjon av de tre eksperimentatordefinerte klassene i den siste testfasen, uavhengig av type testformat. Det var 100% samsvar i prestasjoner mellom de to sorteringstestblokkene i hver sorteringstestfase på 28 av totalt 32 presenterte sorteringstester. Det ble påvist statistisk signifikant reduksjon i hastighet eller invers reaksjonstid (InvRT) for trials fra treningsfasen til baseline, symmetri og ekvivalens trials i testfasen, uavhengig av om deltakerne responderte i henhold til stimulusekvivalens eller ikke. Samvariasjonen i resultater på de to testformatene bekrefter at sorteringstester er et egnet testformat i forbindelse med etablering og opprettholdelse av stimulusklasser. Disse resultatene åpnet for en diskusjon av definisjon av emergente stimulusklasser i sorteringstestene.

Nøkkelord: Invers reaksjonstid, Sortering, Stimulusklasser, Stimulusekvivalens,

Umiddelbar emergens

Abstract

In an attempt to systematically replicate earlier findings from the Arntzen et al. (2017) study, 22 participants aged 24–60 years, were taught 12 conditional discriminations in a matching-to-sample format using a Many-to-one training structure and the simultaneous protocol. Then derived relations tests were used to assess the emergence of three equivalence classes with five members. In Group 1, baseline training was followed by two sorting test blocks, two MTS test blocks and finally two sorting test blocks. In Group 2, baseline training was followed by two MTS test blocks, two sorting test blocks and finally two MTS test blocks. In Group 1, initial sorting test blocks showed immediate emergence of three experimenter-defined classes for 7 participants. Four participants showed other patterns of responding. In Group 2, initial MTS test blocks showed immediate emergence of three equivalence classes for nine participants. One participant showed an other pattern of responding. Twenty of 22 participants showed emergence of the three experimenter-defined classes on the last test block, independent of test format type. Sorting test data showed a 100% concordance in performance between test blocks on 28 of a total of 32 presented sorting tests. Inverse reaction time data (InvRT) showed a statistical significant reduction in response speed from training to test trials for baseline relations, symmetry and equivalence, independent of response accuracy. The degree of concordance in responding on the two types of test formats supports the notion about the usefulness of sorting tests in testing for emergence and maintenance of stimulus classes. In light of these findings, the definition of emergent behavior on sorting tests was discussed.

Keywords: Inverse reaction time, Immediate emergence, Sorting, Stimulus classes, Stimulus equivalence

Stimulusekvivalens studeres som regel ved bruk av *matching-to-sample* (MTS) prosedyrer hvor man trener systematisk på noen betingede (hvis-så) relasjoner mellom stimuli, og tester for om nye, emergente relasjoner mellom stimuliene har oppstått (e.g., Sidman, 1994; Sidman, 2009; Sidman & Tailby, 1982). Dersom den betingede relasjonstreningen har ført til etablering av betinget diskriminasjon og partisjonering av stimulusklasser, samt at alle stimuliene innad i hver stimulusklasse har blitt gjensidig utskiftbare i funksjon, inngår stimuliene i ekvivalensklasser (Green & Saunders, 1998; Sidman, 1994).

Stimuliene som benyttes kan være abstrakte eller kjente for forsøkspersonen, og være av felles eller forskjellige modaliteter (Green & Saunders, 1998; Sidman, 1994). Ved betinget diskriminasjonstrening brukes bokstaver som betegnelse på ulike klasser, og tall for å indikere antall medlemmer i hver klasse (Fields & Verhave, 1987). I stimulusekvivalensstudier benyttes minimum to klasser med tre medlemmer, men det anbefales å bruke tre klasser på grunn av fare for seleksjon av sammenligningsstimuli basert på avvisningskontroll (*rejection control*) (Carrigan & Sidman, 1992; Green & Saunders, 1998; Johnson & Sidman, 1993; Sidman, 1987). I hver trial presenteres en utvalgsstimulus, f.eks. A1, og forsøkspersonen får anledning til å velge mellom sammenligningsstimuliene B1, B2 og B3. Valg av stimulus B1 medfører programmert konsekvens i form av tilbakemelding om at valget var korrekt, mens valg av B2 eller B3 medfører tilbakemelding om at valget ikke var korrekt. På tilsvarende måte trenes betingede relasjoner eksempelvis mellom A1/ B1B2B3, A2/B1B2B3, A3/B1B2B3 osv. (Fields & Verhave, 1987; Sidman, 2009; Sidman & Tailby, 1982).

Under ekstinksjonsbetingelser, dvs. uten programmerte konsekvenser, testes det for emergens av utrente relasjoner, og om de emergente relasjonene innehar egenskapene som definerer ekvivalens: refleksivitet, symmetri og transitivitet (Fields & Verhave, 1987; Sidman, 1994; Sidman & Tailby, 1982). Refleksivitet testes ved å undersøke om hver stimulus er relatert til seg selv, e.g. hvis A2 er utvalgsstimulus så selekteres A2 som

sammenligningsstimulus. Videre hvis B1 så B1. Symmetri testes ved å undersøke om relasjonen mellom stimuliene forblir intakte dersom utvalgsstimulus og sammenligningsstimulus bytter funksjon. Det betyr at dersom relasjonen A1B1 er etablert under trening, så vil deltakeren under testbetingelse selektere A1 som sammenligningsstimulus når B1 er utvalgsstimulus (B1A1). For å undersøke om en relasjon innehar egenskapen transitivitet må minst to betingede relasjoner mellom stimuli i en klasse på totalt tre stimuli ha blitt lært (e.g. A1B1 og B1C1). Transitivitetstesten går ut på å undersøke om forsøkspersonen matcher sammenligningsstimulus fra den andre betingede relasjonen til utvalgsstimulusen i den første betingede relasjonen (A1C1). Egenskapen kombinert symmetri og transitivitet —global ekvivalens eller sann matching—testes ved å se om forsøkspersonen matcher sammenligningsstimulus fra den første betingede relasjonen til utvalgsstimulusen fra den andre betingede relasjonen (C1A1). Dersom testen viser at de emergente betingede relasjonene innehar egenskapene refleksivitet, symmetri og transitivitet, heter det at forsøkspersonen responderer i henhold til stimulusekvivalens (Sidman, 1986, 1994; Sidman & Tailby, 1982).

Forskning har avdekket at flere variabler knyttet til metodevalg påvirker formasjon av stimulusekvivalensklasser (Arntzen, 2012). Valg av treningsstruktur har fått mye oppmerksomhet i denne sammenhengen (Arntzen, 2012; Arntzen, Grondahl, & Eilifsen, 2010; Arntzen & Hansen, 2011). Treningsstrukturer er systematiske oppsett for i hvilken retning de betingede relasjoner trenes i MTS prosedyrer, og er —sammen med bl.a. størrelse på stimulusklassene som trenes og valg av protokoll— avgjørende for hvilke egenskaper, og hvor mange emergente relasjoner det er mulig å påvise (Arntzen et al., 2010; Fields & Verhave, 1987). De tre vanligste treningsstrukturene er *Linear Series* (LS), *One-to-many* (OTM) og *Many-to-one* (MTO) (e.g. K. J. Saunders, Saunders, Williams, & Spradlin, 1993). MTO og OTM kalles også *comparison-as-node* og *sample-as-node*, hvor en node er en

stimulus som er relatert til minst to andre stimuli i trening. Å benytte MTO treningsstruktur innebærer å trene AC og BC, mens man ved å benytte en OTM treningsstruktur trener AB og AC. Ved LS treningsstruktur trenes $A \rightarrow B \rightarrow C$, og hvor B stimulus er en node som er relatert til A som sammenligningsstimulus, og relatert til C som utvalgsstimulus i trening. De forskjellige treningsstrukturene tillater dermed forskjellige egenskaper å bli testet for ut fra i hvilken retning stimuliene har blitt relatert i trening (Fields & Verhave, 1987; Fields, Verhave, & Fath, 1984). Ved LS testes bl.a. AC og CA, og er den eneste strukturen som tillater å teste for kombinert symmetri og transitivitet slik egenskapene defineres av Sidman og Tailby (1982). Ved bruk av MTO testes relasjonene AB og BA, og ved bruk av OTM testes BC og CB. Disse relasjonene refereres til som ekvivalenstrials i stimulusekvivalensforskning (Eilifsen & Arntzen, 2015). En del studier har vist at MTO er mest effektiv med hensyn til antall deltakere som responderer i henhold til ekvivalens (Fields, Hobbie-Reeve, Adams, & Reeve, 1999; Hove, 2003; K. J. Saunders et al., 1993; R. R. Saunders, Chaney, & Marquis, 2005; R. R. Saunders, Drake, & Spradlin, 1999; R. R. Saunders & McEntee, 2004; R. R. Saunders, Wachter, & Spradlin, 1988), mens noen har rapportert at flest deltakere responderer i henhold til ekvivalens når OTM benyttes (Arntzen et al., 2010; Arntzen & Hansen, 2011; Arntzen & Holth, 1997; Holth & Arntzen, 2000). Andre fant ingen forskjeller mellom bruk av MTO og OTM (Arntzen, Lian, & Halstadtrø, 2011; Eilifsen & Arntzen, 2015; Vie & Arntzen, 2017). LS blir ansett for å være minst effektiv med hensyn til antall deltakere som responderer i henhold til stimulusekvivalens, og spesielt i kombinasjon med simultan protokoll (Arntzen et al., 2010; Arntzen & Hansen, 2011; Arntzen & Holth, 1997; Eilifsen & Arntzen, 2015; Holth & Arntzen, 2000) som vil si at alle baseline relasjoner trenes før testtrials introduseres (Imam, 2006). I senere tid har data vist at effekten av kombinasjonen LS treningsstruktur og simultan protokoll i trening kan reduseres ved å kombinere andre variabler i tillegg (Arntzen, Nartey, & Fields, 2018). Poenget er at det i

stimulusekvivalensforskning er nyttig å arrangere en kombinasjon av variabler knyttet til metodevalg som ikke direkte fasiliterer de variablene en ønsker å studere (Arntzen et al., 2017; Arntzen et al., 2018; Lian & Arntzen, 2013). Ved å benytte programvare for MTS prosedyrer kan man supplere med målemetoder som kan avdekke variabler som er vanskelige å observere med det blotte øyet, men som kan være viktige av validitetshensyn (Dymond & Rehfeldt, 2001; Palmer, 2010). Aktuelle supplerende målemetoder i ekvivalensforskning er *eye-tracking* (e.g. Dube et al., 2006; Huziwara, De Souza, & Tomanari, 2016; Steingrimsdottir & Arntzen, 2016), elektroencefalografi/EEG (Arntzen & Steingrimsdottir, 2017), verbal rapportering (Vie & Arntzen, 2017), stimulus *recall* og sortering (Arntzen, Norbom, & Fields, 2015; Dymond & Rehfeldt, 2001; Pilgrim & Galizio, 1996) og reaksjonstidsmålinger (e.g. Spencer & Chase, 1996).

Reaksjonstidsmålinger har blitt målt både i form av latenstid og responshastighet (Baron, 1985; Spencer & Chase, 1996; Whelan, 2008), og kan gjennomføres ved å måle tiden fra sammenligningsstimuli vises på skjermen, til forsøkspersonen har selektert en sammenligningsstimulus. Målinger av responshastighet kan gi informasjon om forhold ved formasjon av ekvivalensklasser ut over det vi kan observere av korrekte eller ikke korrekte responser (Dymond & Rehfeldt, 2001; Spencer & Chase, 1996), og har blant annet blitt benyttet for å undersøke om alle stimuli i en stimulusklasse er like sterkt relaterte til hverandre (Fields, 2016). Måling av responstid kan være spesielt nyttig når man benytter treningsstrukturer hvor det er stor sannsynlighet for klassebasert respondering, eller når det ikke er mulig eller relevant å studere forskjeller i klassebasert respondering som funksjon av nodal avstand. Flere studier har funnet en reduksjon i InvRT fra trening til testtrials, og hvor deltakerne har vist lavere InvRT på ekvivalenstrials enn på symmetri-trials (e.g. Arntzen, Braaten, Lian, & Eilifsen, 2011; Arntzen et al., 2010; Eilifsen & Arntzen, 2009; Holth & Arntzen, 2000; Spencer & Chase, 1996).

Sorteringstester har blitt benyttet i ulike grener av psykologi, bl.a. for å studere konseptformasjon (Ludvigson & Caul, 1964), forsterkningsbetingelser i problemløsningsituasjoner (Grant & Berg, 1948), i studier av frontallapp-funksjon ved ervervede hjerneskader, og i studier av problemløsning hos barn (se f.eks. Chase-Carmichael, Ris, Weber, & Schefft, 1999). På 90-tallet ble sorteringstester tatt i bruk i stimulusekvivalensforskning (Cowley, Green, & Braunling-McMorrow, 1992; Dymond & Rehfeldt, 2001; Green, 1990), og har blitt benyttet til å undersøke sammenslåing og inndeling av stimulusklasser i forbindelse med overføring av kontekstuell kontroll (Mackay, Wilkinson, Farrell, & Serna, 2011; Sigurdardottir, Mackay, & Green, 2012), overføring av betinget forsterkning og straff (Hayes, Kohlenberg, & Hayes, 1991), i forbindelse med *tact* trening (Lowe, Horne, Harris, & Randle, 2002) og i forbindelse med emergens av transitive relasjoner etter trening med MTS prosedyrer versus trening med *Serial Learning* prosedyrer (Dickins, 2011). Sorteringstester har også blitt benyttet i forbindelse med etablering av akademiske ferdigheter og ekvivalensbaserte instruksjoner (Arntzen, Lian, et al., 2011; Barron, Leslie, & Smyth, 2018; Varelas & Fields, 2017), samt blitt benyttet som et alternativ til tradisjonelle MTS tester for å måle etablering og opprettholdelse av ekvivalensklasser (Arntzen, 2004; Arntzen et al., 2017; Arntzen, Nartey, & Fields, 2014; Arntzen et al., 2018; Arntzen et al., 2015; Arntzen & Steingrimsdottir, 2017; Cowley et al., 1992; Dickins, 2015; Eilifsen & Arntzen, 2009, 2011; Fields, Arntzen, & Moksness, 2014; Fields, Arntzen, Nartey, & Eilifsen, 2012; Fienup & Dixon, 2006; Green, 1990; Hove, 2003; Mensah & Arntzen, 2017; Nartey, Fields, & Arntzen, 2015; Nedelcu, Fields, & Arntzen, 2015; Pilgrim & Galizio, 1996; Smeets, Dymond, & Barnes-Holmes, 2000; Steingrimsdottir & Arntzen, 2016).

Flere varianter av sorteringstester har blitt benyttet i ekvivalensforskning. Mackay et al. (2011) benyttet sorteringstester i form av en type flervalgs MTS test som innebar simultan presentasjon av alle utvalgs- og sammenligningsstimuli. Dickins (2015) benyttet

sorteringstester hvor stimuli skulle plasseres i forhold til hverandre i et rutenett, mens Lowe et al. (2002) ba deltakere om å sortere fysiske gjenstander. Smeets et al. (2000) benyttet sorteringsoppgaver på ark, hvor sorteringen skulle markeres ved å ringe rundt bilder av stimuli med penn. Noen sorteringsoppgaver har gått ut på å sortere laminerte kort med stimuli som er identiske med stimuli benyttet i MTS trening og testing (Arntzen, 2004; Arntzen et al., 2018; Arntzen & Steingrimsdottir, 2017; Cowley et al., 1992; Eilifsen & Arntzen, 2009, 2011; Fields et al., 2014; Fields et al., 2012; Hove, 2003; Mensah & Arntzen, 2017; Steingrimsdottir & Arntzen, 2016). Tilsvarende sorteringsoppgave har også blitt presentert i dataprogram (Arntzen et al., 2017; Arntzen et al., 2015). Sorteringstester har nytteverdi både av generaliserings- og generalitetshensyn (Dymond & Rehfeldt, 2001; Pilgrim & Galizio, 1996), og tar betydelig kortere tid å administrere og gjennomføre sammenlignet med tradisjonelle MTS tester (Arntzen et al., 2017; Arntzen et al., 2015; Fields et al., 2014).

Opprinnelig har sorteringstester blitt presentert etter MTS tester for å undersøke om ekvivalensklassene kan generaliseres til et annet testformat. Det har gjentatte ganger blitt dokumentert korrespondanse mellom formasjon av klasser i MTS tester og sorteringstester, og suksess på MTS tester har dermed blitt forstått som en indikator for suksess på sorteringstester (Arntzen et al., 2017; Arntzen et al., 2015; Fields et al., 2014; Fields et al., 2012). Arntzen et al. (2015) presenterte sorteringstester etter trening av baselinereelasjoner for å undersøke om det var mulig å dokumentere umiddelbar emergens av ekvivalensklasser i andre formater enn MTS format. Tre deltakere viste umiddelbar formasjon av klassene, men Arntzen et al. (2015) påpekte faren for interaksjonseffekter fordi disse deltakerne hadde erfaring med formasjon av andre ekvivalensklasser tidligere i den samme studien. Arntzen et al. (2017) kontrollerte for interaksjonseffekter ved å benytte en mellomgruppe design, og hvor deltakerne i den ene gruppen ble eksponert for sorteringstest rett etter trening av baseline relasjoner, uten å ha erfaring med MTS tester. Ti forsøkspersoner i Gruppe 1 ble eksponert for

en blokk med sorteringstester, to blokker med MTS tester og til slutt en blokk med sorteringstester. I Gruppe 2 ble 10 forsøkspersoner eksponert for to blokker med MTS tester, en blokk med sorteringstester, og til slutt to blokker med MTS tester. Fem av deltakerne i Gruppe 1 viste umiddelbar formasjon av de tre eksperimentatordefinerte klassene da sortering var første testformat, og tre av de samme deltakerne viste opprettholdelse av klassene i påfølgende MTS test. Tre av 10 deltakere i Gruppe 2 viste emergens av de tre klassene da MTS test var første testformat, og seks deltakere viste formasjon av de tre eksperimentatordefinerte klassene da sortering var andre testformat. To deltakere i hver gruppe viste formasjon av de tre eksperimentatordefinerte klassene på sorteringstesten, men viste ikke opprettholdelse av klassene i påfølgende MTS test (Arntzen et al., 2017). Data viste at noen deltakere som ikke nådde mestringskriteriet for stimulusekvivalens likevel viste formasjon av en, to eller tre klasser i påfølgende sorteringstest (Arntzen et al., 2017; Arntzen et al., 2015; Arntzen & Steingrimsdottir, 2017; Fields et al., 2014; Fields et al., 2012; Mensah & Arntzen, 2017; Nartey et al., 2015). Det har blant annet blitt diskutert om prestasjonsforskjeller på de to testformatene er et uttrykk for utsatt emergens av ekvivalensklasser ved repetert testing, eller om forskjellene skyldes at klassene i sorteringstestene sorteres under en løsere form for stimuluskontroll, og at sorteringstestene dermed er mer sensitive for å avdekke klassetilhørighet (Arntzen et al., 2017; Arntzen et al., 2015; Fields et al., 2014; Fields et al., 2012). Plassering av stimuli ved siden av hverandre i sorteringstestene kan skje under diskriminativ kontroll fra to eller flere stimuli samtidig, og kan føre til at klassetilhørigheten blir svakere definert sammenlignet med ekvivalensklassene i MTS testene (Arntzen et al., 2017). Arntzen et al. (2017) foreslo å undersøke robustheten ved stimulusklassene ved å ta flere målinger av klassene i sorteringstestene, der samsvar mellom flere sorteringstester etter hverandre vil kunne indikere sterkere klassebasert kontroll.

Denne studien er en systematisk replikasjon av (Arntzen et al., 2017). Hensikten med

studien var å undersøke om det var mulig å replikere dokumentasjon av umiddelbar emergens av ekvivalensklasser ved bruk av sorteringstest. Flere studier hvor det har blitt benyttet sorteringstester i forbindelse med undersøkelser av emergens og opprettholdelse av ekvivalensklasser, har benyttet LS treningsstruktur (Arntzen et al., 2017; Arntzen et al., 2015; Fields et al., 2014; Fields et al., 2012). Siden det kan være nyttig å undersøke sortering etter bruk av andre treningsstrukturer, ble MTO benyttet i denne studien. Siden det ikke er aktuelt å måle effekter av nodal avstand med en MTO treningsstruktur, og siden MTO er en treningsstruktur som erfaringsmessig fører til høy grad av klassebasert respondering, ble det foretatt målinger av responshastighet/ invers reaksjonstid. For å undersøke om stimulusklassene i sorteringstesten var løsere definert enn ekvivalensklassene i MTS testene, ble det administrert to blokker med sorteringstester i hver testfase for sortering.

Metode

Deltakere

Totalt 22 deltakere ble rekruttert fortløpende blant personlige kontakter. De fleste ble kontaktet skriftlig via forespørsel på tekstmelding, og andre tok selv kontakt etter å ha sett et oppslag på en oppslagstavle. Åtte kvinner og 14 menn fra 24 til 60 år, med en gjennomsnittsalder på 36 år deltok i studien. Blant deltakerne var det variasjon i yrker og utdanningslengde, og noen var studenter. Felles for deltakerne var at de oppga at de ikke hadde kjennskap til stimulusekvivalens fra tidligere.

Ved ankomst til eksperimentlokalet ble deltakerne bedt om å lese og skrive under på en informert samtykkeerklæring. Det informerte samtykket inneholdt blant annet informasjon om at deltakelsen var anonym, at eksperimentet ikke kom til å påføre deltakeren risiko eller ubehag, og at deltakerne hadde rett til å trekke seg når som helst uten at dette ville medføre negative konsekvenser. Videre ble det gitt informasjon om forventet varighet på eksperimentet og at deltakerne ville få en full gjennomgang av egne resultater og

forskningsprosjektet i sin helhet, samt muligheter til å stille spørsmål etter at eksperimentet var ferdig.

Etter at eksperimentet var avsluttet fikk deltakerne en gjennomgang av egne resultater, en innføring i stimulusekvivalens, og ble vist praktiske eksempler på hvordan MTS prosedyrer kan brukes i opplæring av språkferdigheter. Deltakerne fikk anledning til å stille spørsmål og komme med egne innspill. Samtlige fikk med seg et eksemplar av en relevant artikkel om stimulusekvivalens på norsk.

Setting

Det ble benyttet to eksperimentlokaler i studien: Et forskningslaboratorium på OsloMet – storbyuniversitetet, og et eksperimentlokale i en privat bolig i Østfold. Femten forsøkspersoner ble rekruttert i Oslo og syv forsøkspersoner ble rekruttert i Østfold. Forskningslaboratoriet på OsloMet inneholdt to avgrensede båser med lettvegger som rommet 130 x 190 cm i henholdsvis bredde og dybde. Båsene var møblert med et bord på 120 (b) x 45 cm (d) og en stol. Eksperimentlokalet i den private boligen rommet 380 x 240 cm og var møblert med et skrivebord på 155 (b) x 85 (d) cm og en stol, og inneholdt i tillegg en bokhylle og en safe. Det var under tiden eksperimentene pågikk ingen kjente potensielt forstyrrende stimuli i, eller i nærheten av eksperimentlokalene. Deltakerne brukte fra 1,5 til 3 timer på å gjennomføre eksperimentøkten. De som ikke var ferdige med trening av baseline relasjoner innen det hadde gått 2,5 timer fikk tilbud om å avbryte. Det ble sørget for at alle deltakerne hadde drikkevann tilgjengelig i umiddelbar nærhet under hele forsøket. Eksperimentator tok i mot kun en deltaker av gangen og satt alltid utenfor døren til eksperimentlokalet mens eksperimentet pågikk.

Utstyr

Det ble benyttet en 15 tommer Hewlett Packard bærbar datamaskin datamaskin med operativsystem Windows 10 (64-bit). I tillegg ble en Dell ikke-trådløs mus benyttet, samt

mousematte. To dataprogram ble benyttet i studien. Det ene var en spesiallaget programvare som styrte presentasjoner av stimuliene i et MTS format. En utvalgsstimulus ble presentert midt på skjermen, og tre sammenligningsstimuli ble tilfeldig plassert i tre av fire hjørner på skjermen. Da deltaker klikket med musepekeren på utvalgsstimulus kom sammenligningsstimuliene til syne, og en av sammenligningsstimuliene kunne velges ved å bruke musepekeren. Programmet sørget for automatisk registrering av antall korrekte og ikke-korrekte responser per trial og testblokk, reaksjonstider til utvalgs- og sammenligningsstimuli, og forekomst av programmerte konsekvenser under trening. I slutten av hver testfase summerte programmet antall korrekte responser til baseline relasjoner, symmetri- og ekvivalensrelasjoner, samt utregnet en prosentvis totalskåre basert på antall riktige responser i testfasen. MTS programvaren ble brukt til trening av baseline relasjoner (TBR) og til testing av emergente relasjoner (MTS 1a, 1b, 2a og 2b).

Det andre dataprogrammet som ble benyttet var en spesiallaget programvare for sorteringstester (SRT 1 og SRT 2). Sorteringsprogrammet styrte presentasjon av stimuliene i en stablet bunke midt på skjermen, og lagret data i form av skjermbilder automatisk.

Det ble i begge programvarene benyttet en hvit bakgrunn og det samme abstrakte stimulussettet som vist i Figur 1. I MTS programmet målte stimuliene 3 cm (h) x 2 til 3 cm (b) på skjermen, og var sorte på et 4x4 cm hvitt lerret som tilsvarte klikksensitivt område. I sorteringsprogrammet målte de samme stimuliene 2 cm (h) x 1,5 til 2 cm (b), med tilsvarende klikksensitivt område på 3x3 cm.

Design

Det ble benyttet et mellom-gruppe design med randomisert fordeling av deltakere til en av to grupper (se Tabell 1). I gruppe 1 ($n = 11$) var betingelsene ABCB, og i gruppe 2 ($n = 11$) var betingelsene ACBC. Det vil si at alle deltakerne ble eksponert for samme fase med trening av baseline relasjoner (TBR). Deltakerne i gruppe 1 gikk videre igjennom en fase med

to tester i sorteringsformat (SRT1.1. og SRT 1.2), før en fase med to tester i *matching-to-sample* format (MTS 1a og MTS 1b) og til slutt en fase med to tester i sorteringsformat (SRT 2.1 og SRT 2.2). Deltakerne i gruppe 2 gikk videre fra TBR til en fase med to tester i *matching-to-sample* format (MTS 1a og MTS 1b), før en fase med to tester i sorteringsformat (SRT 1.1 og SRT 1.2) og avslutningsvis en fase med to tester i *matching-to-sample* format (MTS 2a og MTS 2b).

Prosedyre

Betingede relasjoner mellom stimuli i tre 5-medlemmers klasser ble etablert ved bruk av MTO treningsstruktur. I Tabell 2 vises en detaljert oversikt over trenings og testfaser i eksperimentets betingede diskriminasjonsprosedyre. Relasjoner mellom AE, BE, CE og DE - stimuli ble trent slik at det var korrekt å velge E1 blant E1, E2 og E3 som sammenligningsstimuli, gitt at A1 var utvalgsstimulus. Videre var det korrekt å velge E2 blant E1, E2 og E3 som sammenligningsstimuli, gitt at A2 var utvalgsstimulus osv. Det ble benyttet en simultan protokoll hvor alle baseline relasjonene ble trent på før testene ble introdusert. Relasjonene ble presentert samtidig (*concurrent*) og randomisert. Det betyr at alle relasjonene ble presentert i blandet og tilfeldig rekkefølge i hver blokk.

Etter å ha signert på samtykkeerklæringen fikk deltakerne informasjon om at eksperimentet bestod av flere faser, og ble bedt om å hente eksperimentator når hvert program var slutt. Etter å ha satt seg i eksperimentbåsen ble deltakerne møtt av følgende instruksjon på datamaskinskjermen:

Det vil komme en stimulus midt på skjermen. Du skal klikke på denne med musen.

Tre andre stimuli vil komme til syne. Velg en av disse ved å klikke med musen. Hvis du velger den vi har definert som korrekt vil det stå "bra", "supert" osv. på skjermen.

Hvis du trykker feil, så vil det stå "feil" på skjermen. Nederst på skjermen vil det telles opp antall korrekte responser. I løpet av eksperimentet vil datamaskinen ikke gi

tilbakemelding på om dine valg er riktige eller feil, men ut fra det du har lært kan du få alle oppgavene riktig. Gjør så godt du kan for å få mest mulig riktig. Lykke til! Trykk på Start for å sette i gang eksperimentet.

Treningsfasen (TBR) bestod av blokker á 60 trials hvor hver av de 12 betingede relasjonene ble presentert fem ganger i randomisert rekkefølge. I den første delen av treningsfasen var målet å etablere de betingede relasjonene, og det ble dermed gitt programmerte konsekvenser på hver trial per blokk. Programmerte konsekvenser ble vist på skjermen i 500 ms, etterfulgt av et *intertrial* intervall på 1000 ms. Mestringskriteriet under TBR var 60 av 60 korrekte trials i en blokk. En ikke-korrekt respons i en blokk medførte en påfølgende blokk med tilsvarende antall programmerte konsekvenser som blokken det var respondert feil i. Etter oppnådd mestringskriterium med programmerte konsekvenser på hver trial i en blokk, ble antall programmerte konsekvenser redusert til 50 % per blokk for å trene på opprettholdelse av relasjonene. Etter oppnådd mestringskriterium med kun 50 % programmerte konsekvenser per blokk, bestod den siste delen av treningen av 0 % programmerte konsekvenser per blokk.

Ingen instruksjoner ble gitt før testene i MTS format, og ingen programmerte konsekvenser ble gitt underveis i testen. Hver testblokk inneholdt 180 trials hvor av 36 var baseline trials (AE, BE, CE og DE), 36 var symmetri trials (EA, EB, EC og ED) og 108 var trials hvor det ble testet for ekvivalens (AB, AC, AD, BA, BC, BD, CA, CB, CD, DA, DB og DC). De betingede relasjonene ble presentert tre ganger hver, og det ble administrert to MTS testblokker etter hverandre. Respondering i henhold til stimulusekvivalens ble definert som minimum 95% korrekte responser per test, og med minimum 103 av 108 (95%) korrekte ekvivalenstrials.

På dataskjermen stod følgende instruksjon før sorteringstestene ble presentert:

Bildene vil nå være plassert oppå hverandre. Din oppgave er å plassere bildene i grupper slik du selv mener er korrekt. Når du har gruppert bildene, vær vennlig og marker grupperingen. For å markere holder du en museknapp nede samtidig som du flytter på musen. Dersom du vil markere på nytt kan du flytte på et av bildene og markeringen du har laget vil forsvinne.

Alle stimuliene ble presentert i randomisert rekkefølge, i en bunke med en stimulus på toppen og resten under. Deltakerne måtte selv klikke på og dra stimuliene ut fra bunken og plassere de der de ville. Nederst på skjermbildet var en tast merket med ordet *ferdig*. Da deltagerne trykket på ferdig-tasten kom følgende beskjed opp på skjermen: ”Har du sortert bildene og markert hvordan du har sortert bildene? Markeringen gjøres ved å holde inne høyre museknapp samtidig som du drar musepekeren over skjermen”. Deltakerne kunne videre velge mellom to taster merket med *ja* eller *nei*. Et klikk på nei-tasten førte til at deltakerne forble i samme blokk med sorteringstest for å ferdiggjøre oppgaven, og et klikk på ja-tasten gjorde at et automatisk skjermbilde ble tatt og deltakerne ble eksponert for en ny blokk med sorteringstest. Sorteringsprogrammet lukket seg automatisk etter at skjermbildet i den andre blokken med sorteringstest var tatt. Korrekt formasjon av stimulusklasser på sorteringstestene innebar at deltakeren hadde gruppert stimuli i samsvar med de eksperimentatordefinerte klassene og markert rundt hver klasse.

Uforutsette hendelser

Skjermbilder fra sorteringstest 1.2 ble sjekket fortløpende, mellom presentasjon av de ulike i testfasene. Hos fire deltakere (D16019, D16028, D16031 og D16036) så det ut til at stimuliene var blitt stablet oppå hverandre i bunker hvor stimuliene overlappet, i stedet for å plassere dem ved siden av hverandre i grupper. I disse tilfellene viste skjermbildet kun plassering av en stimulus fra hver av de tre klassene, og med markering rundt hver av dem. Dermed kunne ikke sorteringsdataene brukes i studien. Disse deltakerne ble eksponert for den

første testfasen med sortering på nytt, og ble bedt om å plassere stimuliene ved siden av hverandre uten at stimuliene overlappet. Skjerm bilde av sorteringstest 1.1 hos D16027 viste kun en stimulus, uten markering rundt. Dette kan skyldes en teknisk feil, men dataene er inkludert i studien sammen med data fra deltakerens øvrige testresultater.

Resultater

Tjueen av 22 deltakere fullførte eksperimentet. D16020 trakk seg under trening av baselinereelasjonene. Data fra denne deltakeren er derfor utelatt.

Etablering av baseline relasjoner

I Tabell 3 vises resultater fra trening av baselinereelasjoner, sorteringstester og MTS tester for deltakerne i Gruppe 1, i kronologisk testrekkefølge. I Tabell 4 vises tilsvarende resultater i kronologisk rekkefølge for deltakerne i Gruppe 2. Antall treningstrials varierte fra 360 til 1380, med et gjennomsnitt på 737 trials blant alle deltakerne. En analyse av data basert på gruppetilhørighet viste et gjennomsnitt på 807 treningstrials i Gruppe 1, med en variasjonsbredde fra 360 til 1380 trials. I Gruppe 2 var gjennomsnittet på 660 trials, med en variasjonsbredde fra 420 til 1320 trials. En uavhengig t-test viste at det ikke var statistisk signifikant forskjell mellom gruppene med hensyn til antall trials i trening, $t(9) = 1.325$, $SEM = 113.2$, $p > .05$ (to-halet). Det betyr at en eventuell forskjell i testresultater mellom de to gruppene ikke kan forklares med henvisning til forskjeller i treningsmengde.

Sorteringsdata

For å beskrive sorteringsresultatene kvantitativt ble det benyttet et trestringssystem bestående av klynger med tall, der hver klynge indikerte gruppene i sorteringen. Tallene i hver klynge indikerte hvor mange stimuli fra hver klasse som var plassert i gruppen. Sorteringsdata som reflekterte de eksperimentatordefinerte klassene ble notert som 500 050 005. Eksperimentator og en medobservatør evaluerte skjerm bildedata fra alle sorteringstestene på hvert sitt noteringsark, uavhengig av hverandre, i etterkant av

eksperimentet. Ved visuell inspeksjon av registreringene ble det påvist 100% enighet mellom observatørene på sorteringsdata for begge gruppene.

D16028 sorterte stimuliene i henhold til de eksperimentatordefinerte klassene både på sorteringstest 1.1 og 1.2, men manglet markering rundt klassene på sorteringstest 1.1. Begge observatørene registrerte skjermbildet fra sorteringstest 1.1 som korrekt hos D16028 selv om markering manglet. Testresultatet ble derfor besluttet godkjent som formasjon av de tre klassene.

Blant alle deltakerne viste 16 av 21 (76%) formasjon av de tre eksperimentatordefinerte klassene på alle sorteringstestene, uavhengig om sortering var første eller andre testfase (se Tabell 3 og Tabell 4). Det ble så undersøkt om det var forskjell med hensyn til antall deltakere i hver gruppe som sorterte alle klassene korrekt, avhengig av om første testfase bestod av MTS tester eller sorteringstester. Siden data fra sorteringstest 1.1 hos D16027 ikke var anvendbar som sammenligningsgrunnlag, ble sorteringsdata fra sorteringstest 1.2 i Gruppe 1 vurdert opp mot sorteringsdata fra sorteringstest 1.2 i Gruppe 2. Det ble benyttet en *Fisher's exact* test (FET) som viste at det ikke var statistisk signifikant forskjell mellom gruppene med hensyn til hvor mange i hver gruppe som sorterte alle de tre klassene korrekt, $p > .05$.

Umiddelbar emergens av klasser i Gruppe 1

Oversikt over data fra deltakere som viste formasjon av de tre eksperimentatordefinerte klassene da sortering var første testformat, vises i venstre kolonne i Figur 2. Syv av 11 deltakere viste umiddelbar formasjon av de tre eksperimentatordefinerte klassene på sorteringstest 1.1 og viste opprettholdelse av klassene på sorteringstest 1.2. Samtlige av de syv deltakerne viste emergens av tre ekvivalensklasser på de to påfølgende blokkene med MTS tester. En av deltakerne viste bortfall av to klasser på sorteringstest 2.1. Dermed viste kun seks av syv deltakere opprettholdelse av de tre eksperimentatordefinerte

klassene på sorteringstest 2.1. På den siste sorteringstesten (2.2) viste syv av syv deltakere opprettholdelse av de tre klassene. Det betyr at de emergente klassene som oppstod umiddelbart i første sorteringstest ble opprettholdt på tross av repetert testing og to endringer i testformat hos seks av syv deltakere. Fire av 11 deltakere i Gruppe 1 viste ikke umiddelbar formasjon av de eksperimentatordefinerte klassene da det første testformatet var sortering. Data fra en deltaker i Gruppe 1 som ikke viste umiddelbar formasjon av de tre eksperimentatordefinerte klassene, vises i panelet til høyre i Figur 2. Denne deltakeren viste emergens av tre ekvivalensklasser i de to påfølgende blokkene med MTS tester, og viste deretter opprettholdelse av de tre klassene i den siste fasen med sorteringstester. Tre deltakere i Gruppe 1 viste andre responsmønstre.

MTS test data

Mestringskriteriet for respondering i henhold til stimulusekvivalens forutsatte en total testskåre på 95% eller mer, samt minst 103 av 108 (95%) korrekte testtrials for ekvivalensrelasjoner hos hver deltaker. Blant alle deltakerne i begge gruppene responderte 18 av 21 (85%) i henhold til ekvivalens på MTS test 1a og 1b (se femte og sjette kolonne i Tabell 3, og tredje og fjerde kolonne i Tabell 4). Totalt to av 21 deltakere responderte ikke i henhold til ekvivalens på noen av MTS testene. En deltaker responderte i henhold til ekvivalens kun i siste testfase, etter erfaring med sorteringstester. Det ble benyttet en FET som viste at det ikke var statistisk signifikant forskjell mellom gruppene med hensyn til antall deltakere som responderte i henhold til ekvivalens på MTS test 1, $p > .05$.

Umiddelbar emergens av klasser i Gruppe 2

I Figur 3 vises data for deltakere i Gruppe 2. Totalt ni av 10 deltakere viste umiddelbar emergens av de tre ekvivalensklassene da MTS var første testformat. Samtlige av disse deltakerne viste opprettholdelse av de tre eksperimentatordefinerte klassene under påfølgende faser med sortering og MTS tester. En deltaker i Gruppe 2 viste ikke formasjon av

ekvivalensklassene i første testfase med MTS tester, men viste formasjon av tre eksperimentatordefinerte klasser i den påfølgende fasen på sorteringstest 1.1 og 1.2. Den samme deltakeren viste opprettholdelse av de tre ekvivalensklassene på siste fase med MTS tester.

Samsvar mellom de sorteringsblokkene i hver testfase

I Gruppe 1 var det 100% samsvar i prestasjoner mellom sorteringstestblokkene 1.1. og 1.2 hos ni av 11 deltakere. En deltaker (D16033) viste forbedring i prestasjoner fra formasjon av ingen korrekte klasser i testblokk 1.1, til formasjon av en klasse i testblokk 1.2. Hos en deltaker (D16027) var det som nevnt tidligere ikke mulig å sammenligne prestasjoner på de to første sorteringstestblokkene. Det 100% samsvar i prestasjoner mellom sorteringsblokkene 2.1. og 2.2. hos ni av 11 deltakere i den siste testfasen. To deltakere viste forbedring i prestasjoner fra formasjon av to korrekte klasser på sorteringstest 2.1. til tre korrekte klasser på sorteringstest 2.2. Hos samtlige deltakere i Gruppe 2 var et 100% samsvar i prestasjoner på sorteringstestene 1.1. og 1.2.

Suboptimal formasjon av klasser

D16027 og D16023 i Gruppe 1, og D16024 i Gruppe 2, responderte ikke i henhold til stimulusekvivalens på MTS 1a, og viste dermed heller ikke formasjon av de tre ekvivalensklassene. Det ble undersøkt om en eller to av ekvivalensklassene likevel var etablert i MTS testen hos disse deltakerne, og om eventuelle etablerte klasser i MTS testen samsvarte med etablerte klasser i sorteringstesten. For å avdekke klasser i MTS testen ble det hos hver deltaker regnet ut en indeks basert på antall korrekte responser til stimuli i hver klasse, dividert på antall mulige korrekte responser til stimuli i hver klasse. D16027 viste kun emergens av ekvivalensklasse 3 på MTS test 1a og 1b (se Tabell 5). D16023 viste ikke emergens av ekvivalensklasser på noen av MTS testene. D16024 viste ikke emergens av noen ekvivalensklasser på MTS test 1a, men viste utsatt emergens av klasse 1 på MTS test 1b.

En sammenstilling av etablerte klasser i MTS testene og etablerte klasser i sorteringstestene vises i Figur 4. D16024 viste, som nevnt over, utsatt emergens av klasse 1 på MTS 1b. I den påfølgende testfasen med sortering viste D16024 formasjon av de tre eksperimentatordefinerte klassene. I den siste MTS testfasen viste D16024 emergens av de tre ekvivalensklassene i begge testblokkene. Dermed viste D16024 en økning i korrekt respondering fra første til siste testfase. Dette kan være et tilfelle av utsatt emergens basert på testtype og/eller repetert testing.

D16023 viste kun formasjon av deltakerdefinerte stimulusklasser på sorteringstestene, og stimuliene var sortert i identiske grupper på tvers av testfasene. D16023 viste ingen emergens av ekvivalensklasser på MTS testene. Kun to etablerte ekvivalensrelasjoner (C2D2 og A3B3) kunne gjenspeiles i grupperingen av stimuliene i sorteringstestene. Ut over dette ble det ikke påvist samsvar mellom respondering på MTS testene og formasjon av de deltakerdefinerte klassene i sorteringstesten hos D16023.

Data fra D16027 viste suboptimal formasjon av klasse 2 og 3 på sorteringstest 1.2, men kun emergens av ekvivalensklasse 3 på MTS 1a og 1b. I den siste testfasen med sortering viste deltakeren kun opprettholdelse av klasse 3 på sorteringstest 2.1. På sorteringstest 2.2 viste D16027 formasjon av alle de tre eksperimentatordefinerte klassene. Dermed var det kun samsvar mellom respondering på de to testformatene med hensyn til klasse 3 hos denne deltakeren. I figur 5 vises resultater fra sorteringstest 1.2 hos D16027. Deltakeren viste formasjon av klasse 2 og 3, men klasse 1 var plassert i to grupper. I påfølgende MTS test 1a viste D16027 kun emergens av klasse 3. For å undersøke forskjellen i responderingen mellom de to formatene nærmere, ble det gjennomført en videre analyse av data fra denne deltakeren. Det ble først kontrollert for at distribusjonen av de ulike trial typene til hver klasse var blitt jevnt fordelt i over hele testen. Etter å ha delt testen i to, viste det seg at den første delen av testen inneholdt fem flere presentasjoner av relasjoner knyttet til klasse 3, enn til klasse 1 og

2. Med andre ord forekom det ingen store skjevheter i tidsmessig presentasjon av trial typer knyttet til de tre klassene. Videre ble data fra MTS test 1a delt i tre like deler, som illustrert i figur 6. Hver av de tre testdelene bestod av 60 trials. Den første testdelen inneholdt de 20 første trials med relasjoner knyttet til klasse 1, de 20 første trials med relasjoner knyttet til klasse 2, og de 20 første trials med relasjoner knyttet til klasse 3. Testdel 2 inneholdt de 20 neste trials med relasjoner knyttet til klasse 1, 2 og 3. Testdel 3 inneholdt de 20 siste trials med relasjoner knyttet til klasse 1, 2 og 3. Dette ble gjort for å undersøke om det var flest ikke korrekte responser i starten eller slutten av MTS testen, og om distribusjonen av ikke korrekte responser til de ulike klassene endret seg i løpet av testen.

I del 1 utviste D16027 seks ikke korrekte responser til relasjoner vedrørende klasse 1, ni ikke korrekte responser til klasse 2, og tre ikke korrekte responser til klasse 3. I løpet av del to utviste D16027 fem ikke korrekte responser til klasse 1 og tre ikke korrekte responser til klasse 2. I løpet av den siste testdelen utviste D16027 to ikke korrekte responser til klasse 1 og to ikke korrekte responser til klasse 2. Samlet viste dataene at D16027 utviste flest ikke korrekte responser i første del av MTS testen, dvs. kort tid etter sorteringstesten. På tross av at D16027 viste formasjon av klasse 2 i sorteringstesten, var de fleste av de ikke korrekte responsene i den første delen av den påfølgende MTS testen knyttet til klasse 2. Åtte feil var knyttet til ekvivalensrelasjoner, og en feil var knyttet til symmetrirelasjoner.

Intakte baseline relasjoner i test

Det ble undersøkt om manglende respondering i henhold til ekvivalens kunne skyldes at baseline relasjoner etablert i trening ikke lenger var intakte i MTS test 1a (se Tabell 5). Deltaker 16027, 16024 og 16023 hadde henholdsvis 100%, 97% og 42% korrekte baseline trials i testblokken.

Responshastighet

Som et mål på responshastighet ble invers reaksjonstid (InvRT) benyttet. Utgangspunktet for å måle InvRT var tiden fra tre utvalgsstimuli ble presentert på skjermen til deltaker valgte en utvalgsstimulus ved å klikke på den med musepeker. Beregning av InvRT ble gjort ved å dividere 1 på median reaksjonstid i sekunder, av de fem siste trials i treningsfasen og de fem første korrekte og ikke korrekte trials av de ulike relasjonstypene i MTS test 1a hos hver deltaker. Dette ble gjort for å kunne si noe om hastigheten på deltakernes respondering til hver type relasjon (baseline, symmetri og ekvivalens), og om det var forskjeller i responshastighet i trenings- og testfase. Jo høyere InvRT, desto raskere seleksjon av utvalgsstimulus.

InvRT data er illustrert i Figur 7. Deltakerne som responderte i henhold til ekvivalens viste høyere InvRT i treningsfasen enn på trials i testfasen. Det ble benyttet en *Repeated Measures ANOVA* som viste at det var statistisk signifikant forskjell i InvRT fra treningsfasen til baseline trials i testfasen: $F(1,34) = 33.50, p < .05$. Det var også statistisk signifikant forskjell i InvRT fra treningsfasen til symmetri trials i testfasen; $F(1,34) = 44.68, p < .05$, og fra treningsfasen til ekvivalens trials i testfasen; $F(1,34) = 111.22, p < .05$. I testfasen var InvRT høyest til relasjoner, men høyere på symmetri trials enn ekvivalens trials.

Deltakerne som ikke responderte i henhold til ekvivalens, viste også høyere InvRT i treningsfasen enn på trials i testfasen. *Repeated Measures ANOVA* ble benyttet og viste at det var statistisk signifikant forskjell i InvRT fra treningsfasen til trials i testfasen: $F(1,4) = 20.51, p < .05$. Det var også statistisk signifikant forskjell i InvRT fra treningsfasen til de fem første korrekte symmetri trials i testfasen; $F(1,4) = 13.12, p < .05$, og fra treningsfasen til de fem første ikke korrekte symmetri trials i testfasen; $F(1,4) = 17.06, p < .05$. Det var i tillegg statistisk signifikant forskjell i InvRT fra treningsfasen til de fem første korrekte ekvivalens trials i testfasen; $F(1,4) = 26.60, p < .05$, og fra treningsfasen til de fem første ikke korrekte ekvivalens trials i testfasen; $F(1,4) = 31.56, p < .05$.

Hos de 18 deltakerne som responderte i henhold til ekvivalens var InvRT i treningsfasen på 0,04 sekunder i gjennomsnitt, mot 0,31 sekunder hos de tre deltakerne som ikke responderte i henhold til ekvivalens.

En uavhengig t-test viste at det ikke var statistisk signifikante forskjeller med hensyn til InvRT i mellom de som responderte i henhold til ekvivalens, og de som ikke responderte i henhold til ekvivalens, verken da treningsfasene $t(19) = 1,561$, $p = 0.13$ (to-halet), relasjoner knyttet til baseline $t(20) = 0.4176$, $p = 0.680$ (to-halet), symmetri $t(19) = 1.001$, $p = 0.329$ (to-halet), eller ekvivalens $t(19) = 1.157$, $p = 0.261$ (to-halet), ble sammenlignet. Både hos de som responderte i henhold til ekvivalens og de som ikke responderte i henhold til ekvivalens var InvRT høyest til baseline relasjoner, men høyre på symmetri trials enn ekvivalenstrials.

Anekdotiske data

Debriefen startet med at hver deltaker ble presentert for en oversikt over hele stimulussettet tilsvarende Figur 1, men uten bokstaver for medlemskap og tall for klassetilhørighet. Deltakerne som ønsket det fikk anledning til å forklare hvordan de mente stimuliene var relatert i eksperimentet. Samtlige ga uttrykk for at de hadde oppfattet at noen hørte sammen og andre ikke. Eksperimentator spurte ”hvordan lærte du deg hvilke stimuli som hørte sammen?”. Det var spesielt to måter å relatere stimuliene på som gikk igjen i beskrivelsene: Den ene måten var at hver sammenligningsstimulus i trening ble gitt et navn, f.eks. ”musikk”, ”stopp” og ”opp”. Alle utvalgsstimuli fikk det samme navnet som en av sammenligningsstimuliene ettersom hver utvalgsstimulus viste seg å være en korrekt foranledning for å velge enten ”musikk”, ”stopp” eller ”opp”. På den måten ble alle stimuliene i en klasse relatert ved felles navn. Noen deltakere rapporterte dermed at de i testfasen konsekvent valgte stimuli med samme navn. Den andre måten som ble rapportert var å lage historier som knyttet stimuliene sammen. F. eks. hadde sammenligningsstimulus (E1) i trening fått navnet ”elg”, og både ”Anne” (A1), ”Pacman” (B1), ”Emil” (C1) og ”Nils” (D1)

var relatert til E1 ved at de ”ser elgen”. I en testtrial for ekvivalens hvor ”Anne” (A1) var utvalgsstimulus og B1 (Pacman), B2 (kineser) og B3 (ufo) var sammenligningsstimuli observerte deltakeren følgende: ”Anne” ser ”elgen”, men ”elgen” er ikke her. Er det andre her som ser ”elgen”? ”Pacman” er her, og han ser ”elgen”. Derfor når Anne (A1), så Pacman (B1). Det ble ikke samlet inn systematiske data på selvrapperte problemløsningsstrategier i denne studien, og de anekdotiske dataene behøver dermed ikke å være representative for andre personer.

Diskusjon

I dette eksperimentet ble 12 betingede diskriminasjoner trent i et MTS format, hos 22 forsøkspersoner. Det ble så testet for emergens av tre ekvivalensklasser med fem medlemmer. Totalt 21 forsøkspersoner fullførte eksperimentet. Umiddelbar emergens av de eksperimentatordefinerte klassene ble dokumentert med bruk av sorteringstester hos over halvparten av deltakerne i Gruppe 1, tilsvarende Arntzen et al. (2017) studien. Ni av 10 deltakere i Gruppe 2 viste umiddelbar emergens av stimulusekvivalensklassene da første testformatet innebar MTS. Data viste korrespondanse i prestasjoner mellom de to testblokkene i hver sorteringstest på totalt 28 av 32 presenterte sorteringstester. Det ble påvist en reduksjon i responshastighet fra trening til testfase, og med lavest responshastighet til ekvivalenstrials.

Umiddelbar emergens av klasser

Syv av ti deltakere i Gruppe 1 viste umiddelbar formasjon av de eksperimentatordefinerte stimulusklassene i den første sorteringstestfasen, presentert rett etter trening av baselinerelasjoner. Disse resultatene bekrefter dermed at sorteringstester kan benyttes til å avdekke umiddelbar emergens av stimulusklasser (Arntzen et al., 2017; Arntzen et al., 2015; Varelas & Fields, 2017). Ni av 10 deltakere i Gruppe 2 viste umiddelbar emergens av de tre stimulusekvivalensklassene da MTS tester ble presentert etter trening av

baseline relasjoner. Til sammenligning responderte kun tre av 10 deltakere i Gruppe 2 i Arntzen et al. (2017) studien i henhold til stimulusekvivalens i første MTS testblokk. Endring i treningsstruktur fra LS til MTO kan være årsaken til at flere deltakere viste umiddelbar emergens av ekvivalensklasser på MTS testene i denne studien (Fields et al., 1999; Hove, 2003; K. J. Saunders et al., 1993; R. R. Saunders et al., 2005; R. R. Saunders et al., 1999; R. R. Saunders & McEntee, 2004; R. R. Saunders et al., 1988). Det var ikke statistisk signifikante forskjeller mellom gruppene med hensyn til antall deltakere som viste umiddelbar formasjon av klassene avhengig av testrekkefølge. Gruppene som ble sammenlignet i denne studien innebar relativt få antall deltakere. Andre funn kan kanskje oppnås i tilsvarende studier med flere deltakere.

Korrespondanse i prestasjoner på MTS tester og sorteringstester

Da MTS tester ble presentert etter sorteringstester, responderte totalt ni deltakere i Gruppe 1 i henhold til stimulusekvivalens på MTS 1a og 1b. Åtte av deltakerne viste opprettholdelse av klassene på sorteringstest 2.1, og ni deltakere viste opprettholdelse av klassene på sorteringstest 2.2 (se Tabell 3). Samtlige av de ni deltakerne i Gruppe 2 som viste umiddelbar emergens av stimulusekvivalensklasser på MTS 1a og 1b, viste opprettholdelse av de tre eksperimentatordefinerte stimulusklassene i påfølgende sorteringstestfase (se Tabell 4). Det betyr at det var en korrespondanse mellom suksess på MTS tester og suksess i påfølgende sorteringstester. Disse resultatene er i tråd med tidligere funn (e.g. Arntzen et al., 2017; Arntzen et al., 2014; Arntzen et al., 2015; Arntzen & Steingrimsdottir, 2017; Fields et al., 2014; Fields et al., 2012; Hove, 2003; Lian & Arntzen, 2013; Narthey et al., 2015; Nedelcu et al., 2015; Steingrimsdottir & Arntzen, 2014).

Korrespondanse i prestasjoner på sorteringstester og MTS tester

De syv deltakerne i Gruppe 1 som viste umiddelbar emergens av de eksperimentatordefinerte stimulusklassene i første sorteringstestfase, responderte også i

henhold til stimulusekvivalens i påfølgende MTS test 1a og 1b. De 10 deltakerne i Gruppe 2 som viste formasjon av de tre eksperimentatordefinerte klassene i sorteringstestfasen, responderte og i henhold til stimulusekvivalens i påfølgende MTS testfase. Dermed var det også korrespondanse mellom suksess på sorteringstester og suksess i påfølgende faser med MTS tester i begge grupper. Samlet bidrar funnene til å styrke oppfatningen om at sorteringstester er et egnet testformat til å påvise emergens og opprettholdelse av stimulusklasser etter betinget diskriminasjonstrening (e.g. Arntzen et al., 2017).

Korrespondanse mellom sorteringstestblokkene

Det har blitt drøftet om forskjeller i kontroll av atferd på sorteringstester og MTS tester medfører at sorteringstestene er mer sensitive for klassebasert kontroll, og at klassene som dokumenteres med sorteringstester er løsere definert enn klassene som dokumenteres i MTS testene (Arntzen et al., 2017; Arntzen et al., 2015; Fields et al., 2014; Fields et al., 2012). I denne studien var det stor grad av samsvar i prestasjoner mellom hver testblokk i de suksessivt presenterte sorteringstestblokkene, og det var dermed ikke belegg for å si at klassene som ble påvist med sorteringstester var løsere definert enn klassene som ble påvist i MTS testene. I tillegg var det ingen deltakere som viste reduksjon i korrekt antall sorterte klasser på sorteringstestene med repetert testing. Et unntak er D16025 som viste bortfall av en klasse på sorteringstest 2.1 da en stimulus var sortert i feil klasse (Tabell 3). Siden D16025 hadde hatt suksess på de to forutgående blokkene med sorteringstester, to blokker med MTS tester, samt suksess i den siste sorteringstestblokken (2.2), er det sannsynlig at kollapsen av klasse 1 på sorteringstest 2.1 skyldes fatigue eller lignende.

Utsatt emergens

En deltaker (D16024) fra Gruppe 2 viste ikke umiddelbar emergens av de tre ekvivalensklassene da MTS var første testformat, men viste umiddelbar formasjon av de tre klassene i påfølgende sorteringstestfase. At deltakere ikke responderer i henhold til

stimulusekvivalens på MTS tester, men viser formasjon av de eksperimentatordefinerte klassene i sorteringstestene er et vanlig funn (e.g. Arntzen et al., 2017; Arntzen et al., 2015). Slike resultater har ført til spørsmål om hvor vidt det er forskjeller mellom testformatene med hensyn til kontrollbetingelser, eller om forskjellene er relatert til testrekkefølge, testrepetisjon og utsatt emergens (e.g. Dube & McIlvane, 1996).

Selv om fire deltakere i Gruppe 1 (D16033, D16037, D16027 og D16023) ikke sorterte stimuliene i henhold til de eksperimentatordefinerte klassene i første testfase, hadde samtlige plassert og markert rundt stimuliene i egendefinerte klasser (se Tabell 3). Dermed er det lite sannsynlig at mangel på umiddelbar formasjon av klassene skyldes at noen av deltakerne i Gruppe 1 hadde misforstått sorteringsoppgaven. Selv om D16033 og D16037 ikke sorterte korrekt, responderte begge deltakerne i henhold til stimulusekvivalens i de påfølgende på MTS testene, noe som bl.a. innebærer at baselinerelasjonene var intakte. D16027 viste intakte baseline relasjoner i påfølgende MTS testfase, men responderte derimot ikke i henhold til stimulusekvivalens. Dermed var data fra D16027 var i samsvar med resultater fra andre studier som har vist at det ikke er et sannsynlig årsaksforhold mellom manglende respondering i henhold til stimulusekvivalens og brudd i baselinerelasjoner fra trening til test (Eilifsen & Arntzen, 2009, 2015).

Både deltaker 16033, 16037, 16027 og 16024 viste utsatt emergens av klassene, og samtlige viste formasjon av de tre eksperimentatordefinerte klassene i løpet av siste testblokk. Slike data er interessante fordi de viser at noe skjer underveis i testing, selv uten programmerte konsekvenser. Sidman (1992, 1994) skrev at forsøkspersoners læringshistorie og forhold under eksperimentet, kan gjøre at stimuliene blir etablert som medlemmer i flere klasser enn det som er tilsiktet i gjennom den betingede forsterkningsprosedyren, og kan forstyrre emergens av eksperimentatordefinerte ekvivalensklasser. Dube og McIlvane (1996) foreslo at ukorrekte responser i testfasene kan skyldes at det er en differanse mellom de

faktiske kontrollbetingelsene for atferd og de kontrollbetingelsene eksperimentator har arrangert for. Forfatterne forklarte forskjellen med henvisning til ulike stimuluskontroll topografier—SCTs—som kan forstås som flere ulike kontrollerende relasjoner mellom stimuli og atferd. Forutsetningen for emergens av ekvivalensklasser i en test er at stimuluskontroll topografiene som ble etablert i trening må utkonkurrere de andre stimuluskontroll topografiene i testen. Utsatt emergens kan skyldes at utilsiktet klassemedlemsskap gradvis svekkes med repetert testing (Sidman, 1994), eller at den typen stimuluskontroll topografi som kontrollerte responderingen i treningsfasen, og som var korrelert med en høyere rate programmerte forsterkende konsekvenser, utkonkurrerer de andre utilsiktede stimuluskontroll topografiene (Dube & McIlvane, 1996).

En deltaker viste et annet responsmønster enn samtlige av de andre deltakerne. På sorteringstestene plasserte D16023 stimuliene i fire egendefinerte klasser som var identiske på tvers sorteringstestblokkene, men som ikke korresponderte med de eksperimentatordefinerte ekvivalensklassene. Det vil si at deltakeren responderte konsistent feil på de forskjellige testene. Holth og Arntzen (1998) hevdet at en pre eksperimentell læringshistorie med differensiell forsterkning for konsistent respondering kan føre til både konsekvent forbedring og konsekvent forverring (*nonequivalence*) i testprestasjoner. Dette vises ved at forsøkspersoner konsekvent selekterer stimuli til egendefinerte stimulusklasser basert på en blanding av kontroll fra korrekte (positive) og ukorrekte (negative) stimuli. D16023 responderte ikke i henhold til stimulusekvivalens på hverken MTS 1a eller 1b, og de fleste baseline relasjonene fra trening var ikke intakte i noen av testblokkene. Dermed så det ut til at respondering i treningsfasen, sorteringstestfasene og MTS testfasen var under forskjellig kontroll, kanskje under kontroll av ulike stimuluskontroll topografier, – eller kanskje på grunn av andre forhold relatert til tidligere læringshistorie (se Tabell 3 og 6).

Forskjeller mellom testformatene

Ulike stimuluskontrolltopografier kan ses i sammenheng med forskjeller i trenings- og testbetingelser, og potensielle forskjeller mellom de to typene testformat (Arntzen et al., 2017). Sidman, Kirk, og Willson-Morris (1985) hevdet at trening av baseline relasjoner kun legger grunnlaget for potensielle ekvivalensklasser, og at eksponering for testtrials hvor stimulusbetingelsene er endret, viser om ekvivalensklassene er etablert. I tillegg har det blitt påpekt at siden MTS testene og sorteringstestene ikke tester eksplisitt for de samme emergente relasjonene, så kan man ikke uten videre konkludere med at klassene i sorteringstestene er stimulusekvivalensklasser, slik de defineres ved demonstrasjon av egenskapene refleksivitet, symmetri og transitivitet (Arntzen et al., 2017; Sidman & Tailby, 1982). I MTS testene blir det testet for en spesifikk egenskap ved en type relasjon i løpet av hver trial, og hvor kun et begrenset antall av alle stimuliene fra den betingede diskriminasjonstreningen er til stede på skjermen av gangen. Hver type testtrial i en MTS prosedyre blir vanligvis presentert gjentatte ganger i løpet av en testblokk, og relasjonene det testes for er de samme på tvers av tester og forsøkspersoner. I sorteringstester derimot, testes vilkårlig kun noen av alle mulige relasjoner i en stimulusklasse usystematisk. Hvilke relasjoner det testes for kan avhenge av rekkefølgen på presentasjon av stimuli fra utvalgsbunken, og det vil variere hvilke relasjoner det testes for fra gang til gang. Dersom stimuli sorteres og plasseres ved siden av hverandre vil flere stimuli synes på skjermen av gangen, og flere stimuli kan utøve diskriminativ kontroll over forsøkspersonens sortering samtidig (Arntzen et al., 2017).

Det er også forskjeller med hensyn til hvordan man avgjør suksess på de to testformatene. En gitt prosentandel korrekte testtrials som indikator på respondering i henhold til stimulusekvivalens, gir rom for noen få ukorrekte responser til hver relasjonstype på MTS testene. På sorteringstestene derimot, kan en stimulus plassert feil sabotere for formasjon av

flere klasser. Både deltaker 16025 og 16027 feilsorterte en stimulus i henholdsvis sorteringstest 2.1 og 1.2, noe som ødela for etablering av to klasser hos begge deltakerne.

I en sorteringstest hvor stimuli plasseres ved siden av hverandre, er det sannsynlig at flere stimuli utøver kontroll over sortering av gangen, men til gjengjeld åpner ikke sorteringstestene for noen feil slik som MTS testene. For å utjevne forskjellene mellom testformatene kan man organisere sorteringstesten slik at stimuli plasseres i bunker i stedet for å plassere dem ved siden av hverandre. Videre kan man vurdere å åpne for en feil i sorteringstestene, eller å heve mestringskriteriet for stimulusekvivalens i MTS testene til 100% korrekte trials.

Testresultatene fra D16027 viste at det ikke var perfekt samsvar mellom prestasjoner i sorteringstestene og i MTS testene, og at deltakeren ikke gjorde helt det samme i starten av MTS 1a, som vedkommende gjorde på den forutgående sorteringstesten (1.2). I tillegg viste deltakeren utsatt emergens av de tre eksperimentatordefinerte klassene på den siste blokken med sorteringstest, uten å ha respondert i henhold til stimulusekvivalens. For det første viser resultatene fra denne deltakeren at klasseformasjon muligens fasiliteres av gjentatt eksponering for testing, og varierer med type testformat. For det andre viser resultatene at det for enkelte deltakere kan være stor forskjell i prestasjoner på de to testformatene, og hvor emergent atferd påvises i størst grad med sorteringstester. Pilgrim (2016) påpekte at definisjonen av stimulusekvivalens ikke omfatter alle typer emergent atferd som genereres etter stimulusekvivalens eksperimenter og prosedyrer, og at det derfor bør vurderes om en utvidelse av det vi kaller respondering i henhold til stimulusekvivalens i ekvivalensprosedyrer. Dette er interessant i forbindelse med sorteringstester, spesielt i tilfellene hvor man ser formasjon av alle de eksperimentatordefinerte klassene i sorteringstestene, men uten at forsøkspersonen responderer i henhold til stimulusekvivalens på MTS tester, slik som hos D16027. Fordelen ved å ikke avskrive emergent atferd som ikke

defineres av refleksivitet, symmetri og transitivitet, er at man kan oppnå en bedre forståelse for fenomenet stimulusekvivalens, samt oppnå mer effektiv anvendelse av MTS prosedyrer i opplæringsammenheng (Pilgrim, 2016).

Respons hastighet

Det ble påvist reduksjon i InvRT data fra trening til testfasen, samt at InvRT varierte systematisk med trial type, uavhengig av om deltakerne responderte i henhold til ekvivalens eller ikke. Disse funnene, i tillegg til at deltakerne i testfasen responderte raskere på baseline relasjoner enn på symmetri relasjoner, og viste lavest respons hastighet til ekvivalensrelasjoner, er i samsvar med tidligere forskning (Arntzen, Braaten, et al., 2011; Arntzen, Galaen, & Halvorsen, 2007; Arntzen et al., 2010; Eilifsen & Arntzen, 2009; Holth & Arntzen, 2000; Spencer & Chase, 1996; Steingrimsdottir & Arntzen, 2016). At deltakerne brukte lenger tid på å velge en sammenligningsstimulus i testfasen, og lengst tid på ekvivalensstimuli, åpner for muligheten for at deltakerne gjorde et eller annet før de selekterte en sammenligningsstimulus. Noen har foreslått former for hjelpeatferd (*precurrent behavior*), eksempelvis ved at forsøkspersonene gir stimuli navn (e.g. Arntzen, 2006; Arntzen & Vie, 2013; Holth & Arntzen, 2000; Vie & Arntzen, 2017). De anekdotiske dataene fra denne studien tydet på at noen av deltakerne ga hver stimulus et navn og uttalte beskrivelser av relasjoner mellom de navngitte stimuliene, muligens som en slags variant av intraverbal *naming* (Horne & Lowe, 1996), mens andre ga et felles navn til stimuli i hver klasse. Sistnevnte er ikke en så vanlig observasjon (Horne & Lowe, 1996; Vie & Arntzen, 2017), og kan være nyttig å undersøke systematisk i forbindelse med betinget diskriminasjonstrening.

Begrensninger

De uforutsette hendelsene hos deltaker 16019, 16028, 16031 og 16036 nevnt i metoddelen, medførte at disse deltakerne ble eksponert for en ekstra testfase med sorteringstester. I tillegg fikk de en ekstra instruksjon vedrørende å ikke plassere stimuliene

over hverandre. Dette kan være en trussel mot den indre validiteten i denne studien.

Halvparten av deltakerne som ble rammet av den overnevnte feilen tilhørte Gruppe 1, og resten tilhørte Gruppe 2. Samtlige av de fire deltakerne hadde suksess på begge testformatene.

Siden det ikke har blitt påvist noen sammenheng mellom formasjon av arbitrære stimulusklasser på sorteringstester før og etter MTS prosedyrer i andre studier—og det heller ikke er å forvente hos forsøkspersoner som meddeler at de ikke har kjennskap til stimulusekvivalensforskning fra tidligere—så ble pre sorteringstest utelatt fra denne studien.

Oppsummering

I denne studien ble det bekreftet at sorteringstester kan benyttes til å avdekke umiddelbar emergens av stimulusklasser etter trening av baseline relasjoner i MTS format. Det var stor grad av korrespondanse mellom prestasjoner på to suksessivt presenterte testblokker i hver sorteringstestfase, samt korrespondanse mellom prestasjoner på sorteringstestene og MTS testene. En deltaker viste formasjon av de tre eksperimentatordefinerte klassene i en sorteringstestblokk, men responderte ikke i henhold til ekvivalens på MTS testene. Emergent atferd som fremkommer etter betinget diskriminasjonstrening i opplæringssammenhenger, men som ikke defineres av egenskapene refleksivitet, symmetri og transitivitet, kan i anvendte settinger være like viktig som respondering i henhold til stimulusekvivalens. Sorteringstester er kjappe å administrere sammenlignet med MTS tester, noe som kan være en fordel for enkelte personer i opplæringssituasjoner. E.g. kan MTS tester ta relativt lang tid å gjennomføre og man kan risikere at forsøkspersoner trykker tilfeldig for å bli raskt ferdig. Sorteringstester kan være enklere å administrere hos små barn (Barron et al., 2018) eller når man skal dokumentere emergent atferd i settinger med mange deltakere (Varelas & Fields, 2017).

Referanser

- Arntzen, E. (2004). Probability of equivalence formation: Familiar stimuli and training sequence. *The Psychological Record*, *54*, 275–291. doi:10.1007/BF03395474
- Arntzen, E. (2006). Delayed matching to sample: Probability of responding in accord with equivalence as a function of different delays. *The Psychological Record*, *56*(1), 135–167. doi:10.1007/BF03395541
- Arntzen, E. (2012). Training and testing parameters in formation of stimulus equivalence: Methodological issues. *European Journal of Behavior Analysis*, *13*(1), 123–135. doi:10.1080/15021149.2012.11434412
- Arntzen, E., Braaten, L., Lian, T., & Eilifsen, C. (2011). Response-to-sample requirements in conditional discrimination procedures. *European Journal of Behavior Analysis*, *12*(2), 505–522. doi:10.1080/15021149.2011.11434398
- Arntzen, E., Galaen, T., & Halvorsen, L. R. (2007). Different retention intervals in delayed matching-to-sample: Effects of responding in accord with equivalence. *European Journal of Behavior Analysis*, *8*(2), 177–191. doi:10.1080/15021149.2007.11434281
- Arntzen, E., Granmo, S., & Fields, L. (2017). The relation between sorting test and matching-to-sample tests in the formation of equivalence classes. *The Psychological Record*, *67*(1), 81–96. doi:10.1007/s40732-016-0209-9
- Arntzen, E., Grondahl, T., & Eilifsen, C. (2010). The effects of different training structures in the establishment of conditional discriminations and subsequent performance on tests for stimulus equivalence. *The Psychological Record*, *60*(3), 437–462. doi:10.1007/BF03395720
- Arntzen, E., & Hansen, S. (2011). Training structures and the formation of equivalence classes. *European Journal of Behavior Analysis*, *12*(2), 483–503. doi:10.1080/15021149.2011.11434397

- Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record*, *47*(2), 309–320. doi:10.1007/BF03395227
- Arntzen, E., Lian, T., & Halstadtrø, L. (2011). Anvendelse av matching-to-sample prosedyrer i etableringen av akademiske ferdigheter *Norsk Tidsskrift for Atferdsanalyse*, *38*(1), 1–26. Hentet fra <http://www.nta.atferd.no/journalissue.aspx?IdDocument=207>
- Arntzen, E., Nartey, R. K., & Fields, L. (2014). Identity and delay functions of meaningful stimuli: Enhanced equivalence class formation. *The Psychological Record*, *64*, 349–360. doi:10.1007/s40732-014-0066-3
- Arntzen, E., Nartey, R. K., & Fields, L. (2018). Graded delay, enhanced equivalence class formation, and meaning. *The Psychological Record*, *68*(2), 123–140. doi:10.1007/s40732-018-0271-6
- Arntzen, E., Norbom, A., & Fields, L. (2015). Sorting: An alternative measure of class formation? *The Psychological Record*, *65*(4), 615–625. doi:10.1007/s40732-015-0132-5
- Arntzen, E., & Steingrimsdottir, H. S. (2017). Electroencephalography (EEG) in the study of equivalence class formation. An explorative study. *Frontiers in Human Neuroscience*, *11*(58). doi:10.3389/fnhum.2017.00058
- Arntzen, E., & Vie, A. (2013). The expression of equivalence classes influenced by distractors during DMTS test trials. *European Journal of Behavior Analysis*, *14*(1), 151–164. doi:10.1080/15021149.2013.11434453
- Baron, A. (1985). Measurement scales and the age-complexity hypothesis. *Experimental Aging Research*, *11*(4), 193–199. doi:10.1080/03610738508259187
- Barron, R., Leslie, J. C., & Smyth, S. (2018). Teaching real-world categories using touchscreen equivalence-based instruction. *The Psychological Record*, *68*(1), 89–101. doi:10.1007/s40732-018-0277-0

- Carrigan, P. F., & Sidman, M. (1992). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the experimental analysis of behavior*, 58, 183–204. doi:10.1901/jeab.1992.58-183
- Chase-Carmichael, C. A., Ris, M. D., Weber, A. M., & Schefft, B. K. (1999). Neurologic validity of the Wisconsin card sorting test with a pediatric population. *The Clinical Neuropsychologist*, 13(4), 405–413. doi:10.1076/1385-4046(199911)13:04;1-Y;FT405
- Cowley, B. J., Green, G., & Braunling-McMorrow, D. (1992). Using stimulus equivalence procedures to teach name–face matching to adults with brain injuries. *Journal of Applied Behavior Analysis*, 25(2), 461–475. doi: 10.1901/jaba.1992.25-461
- Dickins, D. W. (2011). Transitive inference in stimulus equivalence and serial learning. *European Journal of Behavior Analysis*, 12(2), 523–555. doi:10.1080/15021149.2011.11434399
- Dickins, D. W. (2015). A Simpler Route to Stimulus Equivalence? A Replication and Further Exploration of a “Simple Discrimination Training Procedure” (Canovas, Debert and Pilgrim 2014). *The Psychological Record*, 65(4), 637–647. doi:10.1007/s40732-015-0134-3
- Dube, W. V., Balsamo, L. M., Fowler, T. R., Dickson, C. A., Lombard, K. M., & Tomanari, G. Y. (2006). Observing behavior topography in delayed matching to multiple samples. *The Psychological Record*, 56(2), 233–244. doi:10.1007/BF03395547
- Dube, W. V., & McIlvane, W. J. (1996). Some implications of a stimulus control topography analysis for emergent behavior and stimulus classes. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 197–218). Amsterdam, The Netherlands: Elsevier Science B. V.

- Dymond, S., & Rehfeldt, R. A. (2001). Supplemental measures of derived stimulus relations. *Experimental Analysis of Human Behavior Bulletin*, *19*, 8–12. Hentet fra http://scholar.google.no/scholar_url?url=http://citeseerx.ist.psu.edu/viewdoc/download%3Fdoi%3D10.1.1.532.1081%26rep%3Drep1%26type%3Dpdf&hl=no&sa=X&scisig=AAGBfm1UK9TGmDgm7IRne7XYMD3npNb-Rg&nossl=1&oi=scholarr
- Eilifsen, C., & Arntzen, E. (2009). On the role of trial types in tests for stimulus equivalence. *European Journal of Behavior Analysis*, *10*(2), 187–202.
doi:10.1080/15021149.2009.11434318
- Eilifsen, C., & Arntzen, E. (2011). Single-subject withdrawal designs in delayed matching-to-sample procedures. *European Journal of Behavior Analysis*, *12*(1), 157–172.
doi:10.1080/15021149.2011.11434361
- Eilifsen, C., & Arntzen, E. (2015). Effects of training structure and the passage of time on trained and derived performance. *The Psychological Record*, *65*(1), 1–12.
doi:10.1007/s40732-014-0067-2
- Fields, L. (2016). Stimulus relatedness in equivalence classes, perceptual categories, and semantic memory networks. *European Journal of Behavior Analysis*, *17*(1), 2–18.
doi:10.1080/15021149.2015.1084713
- Fields, L., Arntzen, E., & Moksness, M. (2014). Stimulus sorting: A quick and sensitive index of equivalence class formation. *The Psychological Record*, *64*(3), 487–498.
doi:10.1007/s40732-014-0034-y
- Fields, L., Arntzen, E., Nartey, R. K., & Eilifsen, C. (2012). Effects of a meaningful, a discriminative, and a meaningless stimulus on equivalence class formation. *Journal of experimental analysis of behavior*, *97*(2), 163–181.
doi:10.1901/jeab.2012.97-163

- Fields, L., Hobbie-Reeve, S. A., Adams, B. J., & Reeve, K. F. (1999). Effects of training directionality and class size on equivalence class formation by adults. *The Psychological Record*, 49(4), 703–724. doi:10.1007/BF03395336
- Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of experimental analysis of behavior*, 48(2), 317–332. doi:10.1901/jeab.1987.48-317
- Fields, L., Verhave, T., & Fath, S. (1984). Stimulus equivalence and transitive associations: A methodological analysis. *Journal of the experimental analysis of behavior*, 42(1), 143–157. doi:10.1901/jeab.1984.42-143
- Fienup, D. M., & Dixon, M. R. (2006). Acquisition and maintenance in visual–visual and visual–olfactory equivalence classes. *European Journal of Behavior Analysis*, 7(1), 87–98. doi:10.1080/15021149.2006.11434266
- Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404–411. doi:10.1037/h0059831
- Green, G. (1990). Differences in development of visual and auditory-visual equivalence relations. *American Journal on Mental Retardation*, 95(3), 260–270.
- Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 229–262). New York, NY: Plenum Press.
- Hayes, S. C., Kohlenberg, B. S., & Hayes, L. J. (1991). The transfer of specific and general consequential functions through simple and conditional equivalence relations. *Journal of the experimental analysis of behavior*, 56(1), 119–137. doi:10.1901/jeab.1991.56-119

- Holth, P., & Arntzen, E. (1998). Stimulus familiarity and the delayed emergence of stimulus equivalence or consistent nonequivalence. *The Psychological Record, 48*(1), 81–110. doi:10.1007/BF03395260
- Holth, P., & Arntzen, E. (2000). Reaction times and the emergence of class consistent responding: A case for precurrent responding? *The Psychological Record, 50*(2), 305–337. doi:10.1007/bf03395358
- Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the experimental analysis of behavior, 65*(1), 185–241. doi:10.1901/jeab.1996.65-185
- Hove, O. (2003). Differential probability of equivalence class formation following a one-to-many versus a many-to-one training structure. *The Psychological Record, 53*(4), 617–634. doi:10.1007/bf03395456
- Huziwara, E., De Souza, D., & Tomanari, G. (2016). Patterns of eye movement in matching-to-sample tasks. *Psicologia: Reflexao e critica, 29*(2), 2–10. doi:10.1186/s41155-016-0010-3
- Imam, A. A. (2006). Experimental control of nodality via equal presentations of conditional discriminations in different equivalence protocols under speed and no-speed conditions. *Journal of the experimental analysis of behavior, 85*(1), 107–124. doi:10.1901/jeab.2006.58-04
- Johnson, C., & Sidman, M. (1993). Conditional discrimination and equivalence relations: Control by negative stimuli. *Journal of the experimental analysis of behavior, 59*(2), 333–347. doi:10.1901/jeab.1993.59-333
- Lian, T., & Arntzen, E. (2013). Delayed matching-to-sample and linear series training structures. *The Psychological Record, 63*(3), 545–561. doi:10.11133/j.tpr.2013.63.3.010

- Lowe, C. F., Horne, P. J., Harris, F. D. A., & Randle, V. R. L. (2002). Naming and categorization in young children: Vocal tact training. *Journal of the experimental analysis of behavior*, 78(3), 527–549. doi:10.1901/jeab.2002.78-527
- Ludvigson, H. W., & Caul, W. F. (1964). Relative effect of overlearning on reversal and nonreversal shifts with two and four sorting categories. *Journal of Experimental Psychology*, 68(3), 301–306. doi:10.1037/h0042254
- Mackay, H. A., Wilkinson, K. M., Farrell, C., & Serna, R. W. (2011). Evaluating merger and intersection of equivalence classes with one member in common. *Journal of the experimental analysis of behavior*, 96(1), 87–105. doi:10.1901/jeab.2011.96-87
- Mensah, J., & Arntzen, E. (2017). Effects of meaningful stimuli contained in different numbers of classes on equivalence class formation. *The Psychological Record*, 67(3), 325–336. doi:10.1007/s40732-016-0215-y
- Nartey, R. K., Fields, L., & Arntzen, E. (2015). Enhancement of equivalence class formation by pretraining discriminative functions. *Learning and Behavior*, 43(1), 20–31. doi:10.3758/s13420-014-0158-6
- Nedelcu, R. I., Fields, L., & Arntzen, E. (2015). Arbitrary conditional discriminative functions of meaningful stimuli and enhanced equivalence class formation. *Journal of the experimental analysis of behavior*, 103(2), 349–360. doi:10.1002/jeab.141.
- Palmer, D. C. (2010). Behavior under the microscope: Increasing the resolution of our experimental procedures. *The Behavior Analyst*, 33(1), 37–45. doi:10.1007/BF03392202
- Pilgrim, C. (2016). Considering definitions of stimulus equivalence. *European Journal of Behavior Analysis*, 17(1), 105–114. doi:10.1080/15021149.2016.1156312
- Pilgrim, C., & Galizio, M. (1996). Stimulus equivalence: A class of correlations, or a correlation of classes? In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class*

- formulaiton in humans and animals* (pp. 173–195). Amsterdam: The Netherlands: Elsevier Science B. V.
- Saunders, K. J., Saunders, R. R., Williams, D. C., & Spradlin, J. E. (1993). An interaction of instructions and training design on stimulus class formation: Extending the analysis of equivalence *The Psychological Record*, *43*(4), 725–744. doi:10.1007/BF03395909
- Saunders, R. R., Chaney, L., & Marquis, J. G. (2005). Equivalence class establishment with two-, three-, and four-choice matching to sample by senior citizens. *The Psychological Record*, *55*(4), 539–559. doi:10.1007/BF03395526
- Saunders, R. R., Drake, K. M., & Spradlin, J. E. (1999). Equivalence class establishment, expansion, and modification in preschool children. *Journal of the experimental analysis of behavior*, *71*(2), 195–214. doi:10.1901/jeab.1999.71-195
- Saunders, R. R., & McEntee, J. E. (2004). Increasing the probability of stimulus equivalence with adults with mild mental retardation *The Psychological Record*, *54*(3), 423–435. doi:10.1007/BF03395483
- Saunders, R. R., Wachter, J., & Spradlin, J. E. (1988). Establishing auditory stimulus control over an eight-member equivalence class via conditional discrimination procedures. *Journal of the experimental analysis of behavior*, *49*(1), 95–115. doi:10.1901/jeab.1988.49-95
- Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and Integration of Behavioral Units* (pp. 213–245). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sidman, M. (1987). Two choices are not enough. *Behavior Analysis*, *22*(1), 11–18. Hentet fra http://www.equivalence.net/pdf/Sidman_1987.pdf
- Sidman, M. (1992). Equivalence relations: Some basic considerations. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 15–27). Reno, NV: Context Press.

- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative, Inc.
- Sidman, M. (2009). Equivalence relations and behavior: An introductory tutorial. *The Analysis of Verbal Behavior*, 25(1), 5–17. doi:10.1007/BF03393066
- Sidman, M., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. *Journal of the experimental analysis of behavior*, 43(1), 21–42. doi:10.1901/jeab.1985.43-21
- Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the experimental analysis of behavior*, 37(1), 5–22. doi:10.1901/jeab.1982.37-5
- Sigurdardottir, Z. G., Mackay, H. A., & Green, G. (2012). Stimulus equivalence, generalization, and contextual stimulus control in verbal classes. *The Analysis of Verbal Behavior*, 28(1), 3–29. doi:10.1007/BF03393105
- Smeets, P. M., Dymond, S., & Barnes-Holmes, D. (2000). Instructions, stimulus equivalence, and stimulus sorting: Effects of sequential testing arrangements and a default option. *The Psychological Record*, 50(2), 339–354. doi:10.1007/bf03395359
- Spencer, T. J., & Chase, P. N. (1996). Speed analyses of stimulus equivalence *Journal of experimental analysis of behavior*, 65(3), 643–659. doi:10.1901/jeab.1996.65-643
- Steingrimsdottir, H. S., & Arntzen, E. (2014). Performance by older adults on identity and arbitrary matching-to-sample tasks. *The Psychological Record*, 64(4), 827–839. doi:10.1007/s40732-014-0053-8
- Steingrimsdottir, H. S., & Arntzen, E. (2016). Eye movements during conditional discrimination training. *The Psychological Record*, 66, 201–212. doi:10.1007/s40732-015-0156-x

Varelas, A., & Fields, L. (2017). Equivalence based instruction by group based clicker training and sorting tests. *The Psychological Record, 67*(1), 71–80.

doi:10.1007/s40732-016-0208-x

Vie, A., & Arntzen, E. (2017). Talk-aloud protocols during conditional discrimination training and equivalence class formation. *The Analysis of Verbal Behavior, 33*(1), 80–

97. doi:10.1007/s40616-017-0081-y

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record, 58*(3), 475–482. doi:10.1007/BF03395630

Tabell 1

Rekkefølge på betingelser for deltakere i gruppe 1 og 2

Fase	1	2	3	4
Gruppe 1	TBR	SRT 1.1 og 1.2	MTS 1a og 1b	SRT 2.1 og 2.2
Gruppe 2	TBR	MTS 1a og 1b	SRT 1.1 og 2.1	MTS 2a og 2b

Merknad. TBR = trening av baseline relasjoner; SRT 1.1 og 1.2 = første og andre testblokk i den første sorteringstestfasen; MTS 1a og 1b = første og andre testblokk i den første matching-to-sample testfasen; SRT 2.1 og 2.2 første og andre testblokk i den andre sorterings- testfasen; MTS 2a og 2b = første og andre testblokk i den andre matching-to-sample testfasen.

Tabell 2

Trenings- og testfase i eksperimentets betingede diskriminasjonsprosedyre

Faser	Type relasjoner	Programmerte konsekvenser (%)	Minimum trials	Mestringskriterium(%)
TBR (etablering)	BR A1E1, A2E2, A3E3 B1E1, B2E2, B3E3 C1E1, C2E2, C3E3 D1E1, D2E2, D3E3	100	60	100
TBR (opprett- holdelse)	BR A1E1, A2E2, A3E3 B1E1, B2E2, B3E3 C1E1, C2E2, C3E3 D1E1, D2E2, D3E3	50 0	60 60	100 100
MTS- test	BR A1E1, A2E2, A3E3 B1E1, B2E2, B3E3 C1E1, C2E2, C3E3 D1E1, D2E2, D3E3 SYM E1A1, E2A2, A3E3 E1B1, E2B2, E3B3 E1C1, E2C2, E3C3 E1D1, E2D2, E3D3 EQ A1B1, A2B2, A3B3 A1C1, A2C2, A3C3 A1D1, A2D2, A3D3 B1A1, B2A2, B3A3 B1C1, B2C2, B3C3 B1D1, B2D2, B3D3 C1A1, C2A2, C3A3 C1B1, C2B2, C3B3 C1D1, C2D2, C3D3 D1A1, D2A2, D3A3 D1B1, D2B2, D3B3 D1C1, D2C2, D3C3	0	180	95 ^a

Merknad. Programmerte konsekvenser indikerer prosentvis antall trials med tilbakemelding i hver blokk. Minimum trials indikerer minimum antall trials tilstrekkelig for å avansere til neste betingelse. Mestringskriterium indikerer prosentvis antall korrekte responser i en blokk som var nødvendig for å avansere til neste betingelse. TBR = trening av baseline relasjoner; BR = baseline relasjoner; SYM = symmetrirelasjoner; EQ = ekvivalensrelasjoner; MTS-test = *matching-to-sample* test. Alle relasjonstypene ble presentert blandet og i tilfeldig rekkefølge i hver blokk.

^a Det ble kontrollert for at deltakerne som responderte i henhold til ekvivalens med total testskåre på 95 %, også hadde minimum 103 (95 %) korrekte ekvivalenstrials.

Tabell 3

Oppsummerte resultater på tvers av faser for deltakerne i Gruppe 1

D#	Antall treningstrials	Sortering 1.1	Sortering 1.2	Prosent korrekte trials på MTS 1a	Prosent korrekte trials på MTS 1b	Sortering 2.1	Sortering 2.2
16030	720	500 050 005	500 050 005	99	100	500 050 005	500 050 005
16039	780	500 050 005	500 050 005	100	99	500 050 005	500 050 005
16018	720	500 050 005	500 050 005	99	99	500 050 005	500 050 005
16019	660	500 050 005	500 050 005	99	99	500 050 005	500 050 005
16029	660	500 050 005	500 050 005	97	99	500 050 005	500 050 005
16031	360	500 050 005	500 050 005	98	97	500 050 005	500 050 005
16025	960	500 050 005	500 050 005	96	99	400 150 005	500 050 005
16033	1320	400 140 015	500 040 015	97	99	500 050 005	500 050 005
16037	660	555 000 000	555 000 000	100	100	500 050 005	500 050 005
16027	1380	I/A ^a	300 050 005 200	83	88	510 040 005	500 050 005
16023	660	102 120 111 222	102 120 111 222	31	33	102 120 111 222	102 111 222 102

Merknad. Resultater fra alle faser i eksperimentet i kronologisk rekkefølge fra venstre, for alle deltakere i gruppe 1. D# = deltaker nummer; MTS = *matching-to-sample* tester. Tall i vertikale kolonner under sorteringstester viser deltakerdefinerte klasser hos hver deltaker. Uthevet skrift betyr at deltakerdefinerte klasser ble sortert i samsvar med de eksperimentatordefinerte klassene. Uthevet skrift i vertikale kolonner under prosent korrekte trials på MTS tester betyr at deltakerne responderte i henhold til ekvivalens.

^a Ikke anvendbare data.

Tabell 4

Oppsummerte resultater på tvers av faser for deltakerne i Gruppe 2

D#	Antall treningstrials	Prosent korrekte trials på MTS 1a	Prosent korrekte trials på MTS 1b	Sortering 1.1	Sortering 1.2	Prosent korrekte trials på MTS 2a	Prosent korrekte trials på MTS 2b
16028	420	100	100	500 050 005	500 050 005	100	100
16021	720	100	100	500 050 005	500 050 005	100	99
16022	900	100	100	500 050 005	500 050 005	99	99
16038	600	100	100	500 050 005	500 050 005	100	99
16032	540	100	97	500 050 005	500 050 005	99	99
16036	1320	99	100	500 050 005	500 050 005	99	100
16034	420	97	100	500 050 005	500 050 005	99	100
16026	480	99	99	500 050 005	500 050 005	100	100
16035	600	99	99	500 050 005	500 050 005	99	100
16024	600	52	89	500 050 005	500 050 005	100	99

Merknad. Resultater fra alle faser i eksperimentet i kronologisk rekkefølge fra venstre, for alle deltakere i gruppe 2. D# = deltaker nummer; MTS = *matching-to-sample* tester. Tall i vertikale kolonner under sorteringstestene viser deltakerdefinerte klasser hos hver deltaker. Uthevet skrift betyr at deltakerdefinerte klasser ble sortert i samsvar med de eksperimentatordefinerte klassene. Uthevet skrift i vertikale kolonner under prosent korrekte trials på MTS tester betyr at deltakerne responderte i henhold til ekvivalens.

Tabell 5

Indeks over korrekte responser per klasse i testtrials for ekvivalens

D#	Test-blokk	MTS 1			MTS 2		
		Kl. 1	Kl. 2	Kl. 3	Kl. 1	Kl.2	Kl.3
16027	a	0.66	0.63	0.94			
	b	0.83	0.77	1.00			
16023	a	0.25	0.33	0.27			
	b	0.16	0.36	0.47			
16024	a	0.30	0.38	0.36	1.00	1.00	1.00
	b	0.97	0.88	0.69	1.00	0.97	1.00
















Merknad. Tabellen viser en indeks over korrekte responser per klasse i testtrials for ekvivalens, hos deltakere som ikke responderte i henhold til ekvivalens på en eller flere MTS tester. Indeksen tilsvarer antall korrekte responser per klasse dividert på antall mulige korrekte responser per klasse. 1.00 tilsvarer en komplett etablert klasse. D# = deltakernummer; MTS 1 og MTS 2 = første og andre fase med *matching-to-sample* tester; Kl. = klasse. Grå markerte felter viser mulig utsatt emergens av klasser og er beskrevet i teksten.

Tabell 6

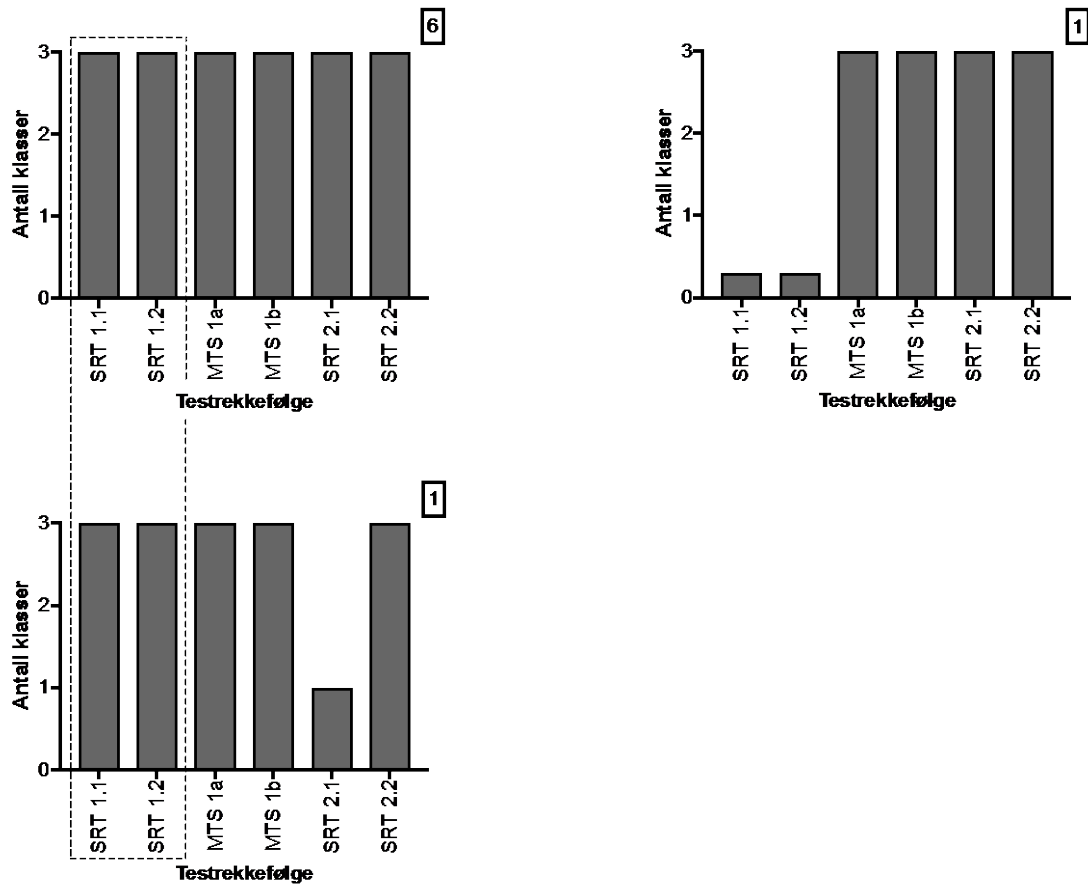
Prosentvis korrekt seleksjon av hver relasjonstype i Matching-to-sample test 1a

D#	Gruppe	Korrekte baseline relasjoner (%)	Korrekte symmetri- relasjoner (%)	Korrekte ekvivalens- relasjoner (%)
16027	1	100	92	75
16023	1	42	25	29
16024	2	97	58	35

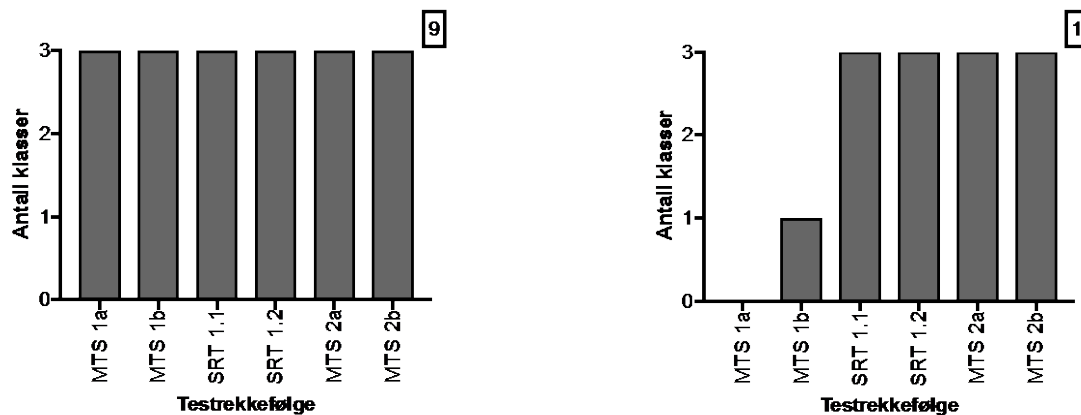
Merknad. D# = Deltaker nummer. Tall med uthevet skrift indikerer mer enn 95% korrekte trials per relasjon, og er i samsvar med eksperimentatordefinert mestringskriterium.

	1	2	3
A			
B			
C			
D			
E			

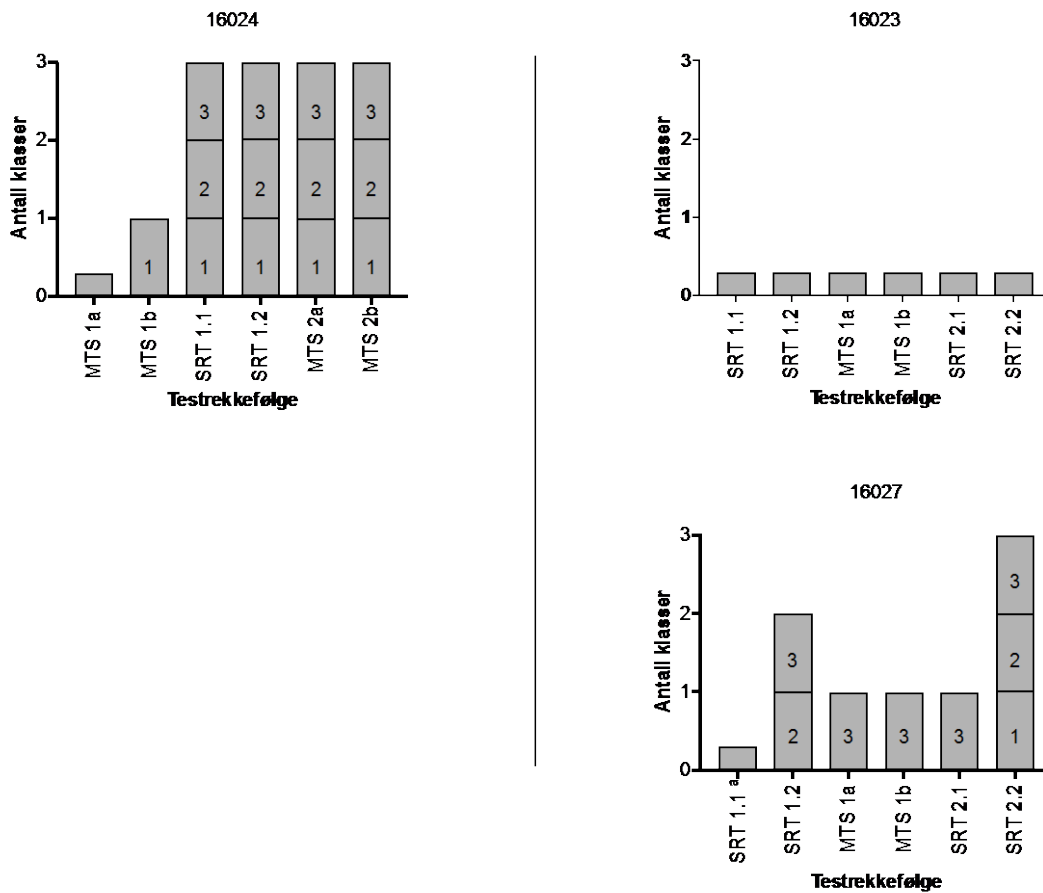
Figur 1. Oversikt over stimuli benyttet i eksperimentet. Nummer 1-3 øverst indikerer de tre eksperimentatordefinerte klassene, og bokstavene i kolonnen til venstre indikerer de fem medlemmene.



Figur 2. Umiddelbar formasjon av tre klasser da sortering var første testformat. Hvert panel viser en type responsmønster på tvers av testbetingelsene. Øverst til høyre for hvert panel indikeres antall deltakere som responderte i hver type mønster. SRT = sorteringstester; MTS = *matching-to-sample* tester.

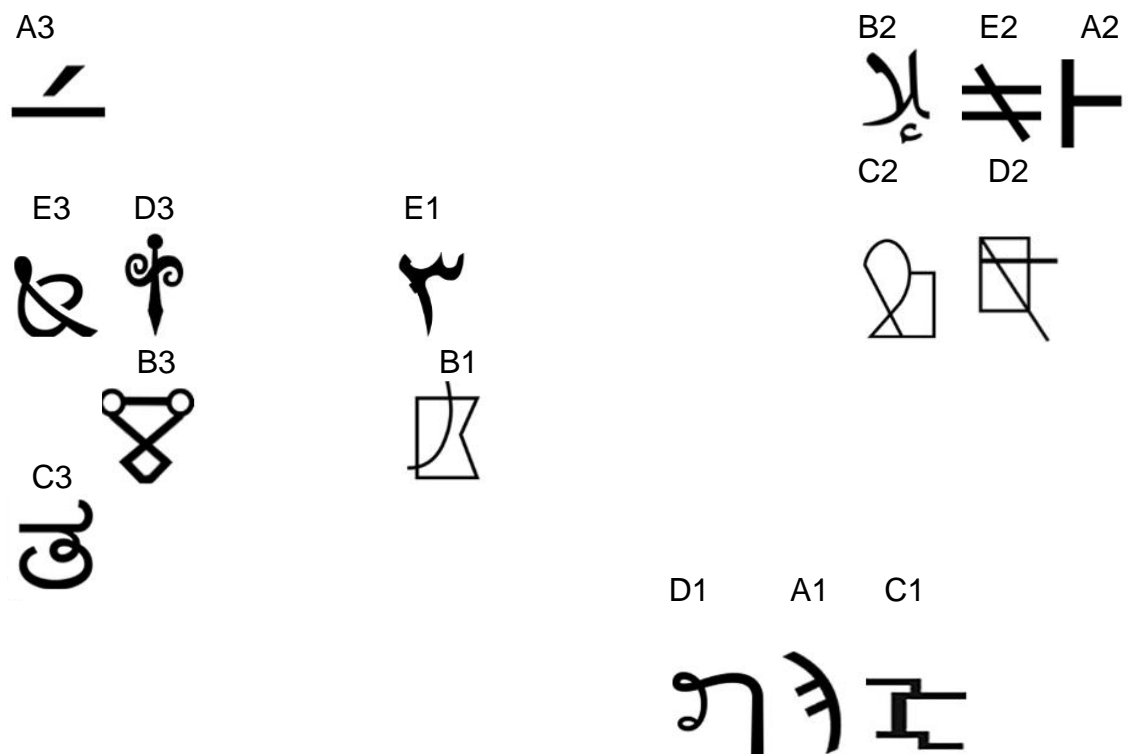


Figur 3. Umiddelbar emergens av tre klasser da MTS var første testformat. Hvert panel viser en type responsmønster på tvers av testbetingelsene. Øverst til høyre for hvert panel indikeres antall deltakere som responderte i hver type mønster. SRT = sorteringstester; MTS = *matching-to-sample* tester.

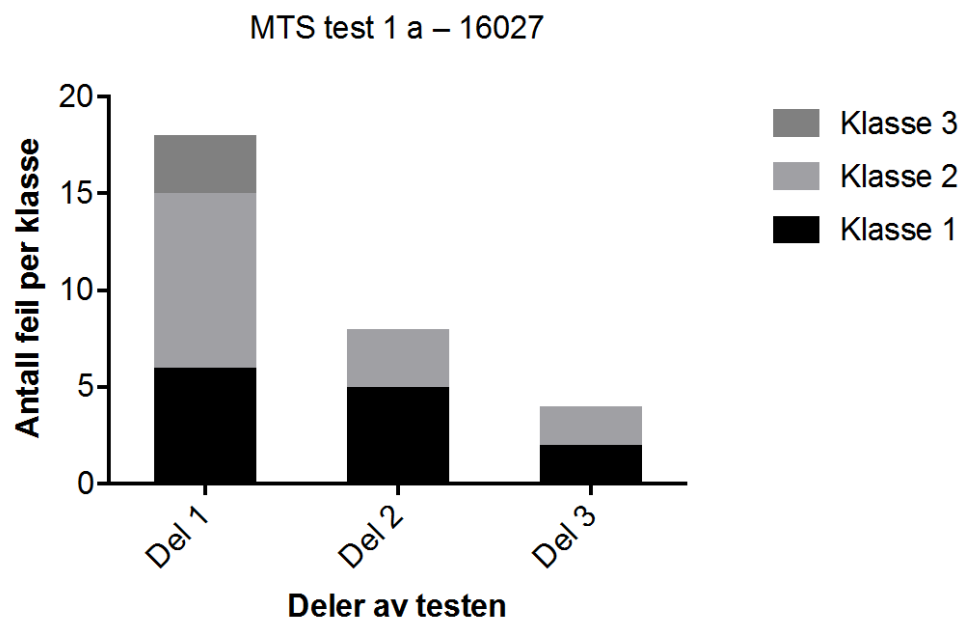


Figur 4. Suboptimal formasjon av klasser hos en deltaker i gruppe 2 (venstre), og to deltakere i gruppe 1 (høyre). SRT = sorteringstester; MTS = *matching-to-sample* tester.

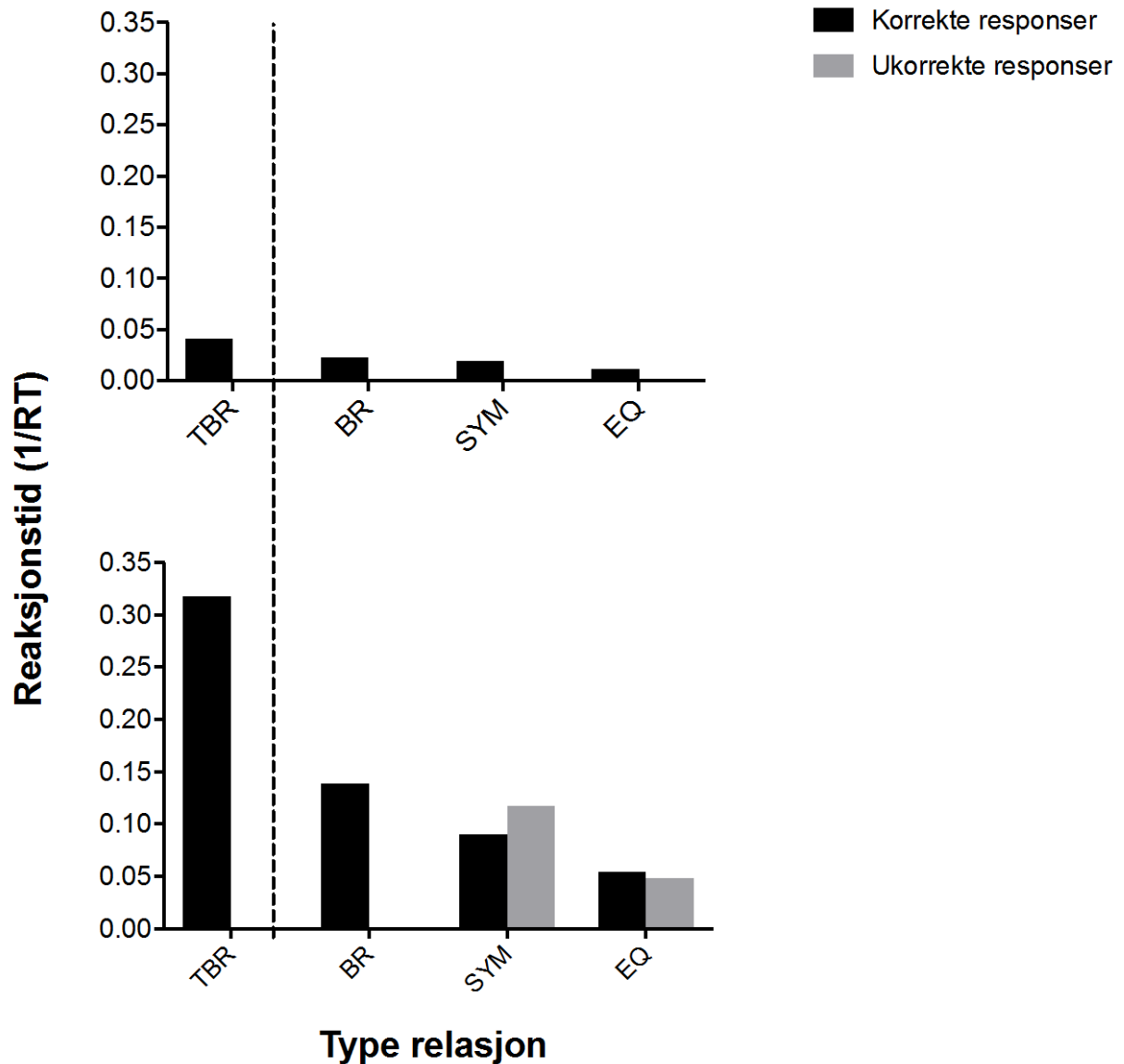
^a Data for SRT 1.1 hos deltaker 16027 var ikke anvendbare. Beskrives i tekst.



Figur 5. Resultater fra sorteringstest 1.2. hos deltaker 16027. Bokstaver over stimuli indikerer medlem, og tall over stimuli indikerer klassesilhørighet.



Figur 6. Antall ikke korrekte responser per testdel når antall trials i løpet av matching-to-sample test 1a ble delt og analysert i tre like store deler. Y-aksen viser antall ikke korrekte responser per klasse, og x-aksen viser testdel 1, 2 og 3.



Figur 7. Inverse reaksjonstid (InvRT) hos deltakere som responderte i henhold til ekvivalens vises i øverste panel (n = 18), og hos deltakere som ikke responderte i henhold til ekvivalens i nederste panel (n = 3). TBR = trening av baseline relasjoner; BR = testtrials for baseline relasjoner; SYM = testtrials for symmetri relasjoner; EQ = testtrials for ekvivalensrelasjoner. InvRT data er beregnet fra de fem siste treningstrials i treningsfasen, de fem første korrekte og ikke korrekte testtrials for baseline relasjoner, symmetri relasjoner og ekvivalens relasjoner hos hver deltaker.