

# Hvor robust er samsvaret mellom respondering på *matching-to-sample* tester og sorteringstester?

Kristiane Rustad Bevolden og Erik Arntzen  
OsloMet — storbyuniversitetet

Formålet med dette eksperimentet var å undersøke om samsvaret mellom resultatene på en *matching-to-sample* (MTS) test og sorteringstester også er gjeldende ved bruk en *Many-to-One* (MTO) treningsstruktur. Vi ønsket å gjøre en systematisk replikasjon av Arntzen, Granmo og Fields (2017) som hadde vist et slik samsvar ved bruk av en *Linear Series* (LS) treningsstruktur. Tjuen voksne forsøkspersoner ble trent 12 betingede diskriminasjoner og testet for om tre ekvivalensklasser med fem medlemmer framkom. Deltakerne i Gruppe 1 ble eksponert for en fase med trening av baseline relasjoner, to sorteringstester, to blokker med MTS tester og to sorteringstester. Deltakerne i Gruppe 2 ble eksponert for en fase med trening av baseline relasjoner, to blokker med MTS tester, to sorteringstester, og to blokker med MTS tester. I Gruppe 1 viste syv deltakere umiddelbar emergens av tre stimulusklasser på begge sorteringstestene, og fire deltakere viste andre responsmønstre. I Gruppe 2 viste ni deltakere umiddelbar emergens av tre klasser i de to blokkene med MTS tester, mens en deltaker viste et annet responsmønster. Tjue av deltakerne viste inndeling i de tre eksperimentatordefinerte klassene i den siste testfasen, uavhengig av type testformat. Det var 100% samsvar i respondering mellom de to sorteringstestene i hver sorteringstestfase på 28 av totalt 32 presenterte sorteringstester. 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, samt anvendelsen av sorteringstester. Oppsummert så utvider dette eksperimentet kunnskapen om robustheten ved samsvaret mellom respondering på *matching-to-sample* tests and sorteringstester.

*Nøkkelord:* Stimulusekvivalens, sortering, stimulusklasser, umiddelbar emergens, utsatt emergens

Stimulusekvivalens studeres som regel ved bruk av *matching-to-sample* (MTS) prosedyrer hvor man trener systematisk på noen betingede relasjoner mellom stimuli, og tester for om nye, emergente relasjoner mellom stimuliene har oppstått (Sidman, 1994, 2009; Sidman & Tailby, 1982). Dersom den betingede relasjonstreningen har ført til etablering av betinget diskriminasjon og partisjonering eller inndeling 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 meningsfulle for forsøkspersonen, og være av felles eller forskjellige modaliteter (Green & Saunders, 1998; Sidman, 1994). Dersom det er seks diskriminasjoner ( $A \rightarrow B \rightarrow C$ ) som skal trenes 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 tilba-

Takk til to anonyme fagfeller for verdifulle kommentarer til en tidligere versjon av manuskriptet. Korrespondanse vedrørende manuskriptet sendes til erik.arntzen@equivalence.net

kemelding 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, B1/C1C2C3, B2/C1C2C3 og B3/C1C2C3<sup>1</sup> (Fields & Verhave, 1987; Sidman, 2009; Sidman & Tailby, 1982).

Under ekstinksjonsbetingelser testes det for emergens av utrente relasjoner, og om de emergente relasjonene innehar egenskapene som definerer stimulusekvivalens: refleksivitet, symmetri og transitivitet (e.g., Sidman & Tailby, 1982). Fra eksemplet ovenfor, så kan refleksivitet testes ved å undersøke om hver stimulus er relatert til seg selv, for eksempel hvis A2 er utvalgsstimulus så velges A2 som sammenligningsstimulus. 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 velge 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 (for eksempel A1B1 og B1C1). Transitivitetstesten går ut på å undersøke om forsøkspersonen matcher sammenligningsstimulus fra den første betingede relasjonen til utvalgsstimulusen i den første betingede relasjonen (A1C1). Egenskapen kombinert symmetri og transitivitet —global ekvivalens—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, innebærer det at forsøkspersonen responderer i henhold til stimulusekvivalens (Sidman,

1986, 1994; Sidman & Tailby, 1982).

På 90-tallet ble sorteringstester tatt i bruk i stimulusekvivalensforskning (e.g., Cowley, Green, & Braunling-McMorrow, 1992; Dymond & Rehfeldt, 2001), samt benyttet som et alternativ til tradisjonelle MTS tester for å måle etablering og opprettholdelse av ekvivalensklasser (e.g., Arntzen, 2004; Arntzen et al., 2017; Arntzen, Nartey, & Fields, 2014; Arntzen, Nartey, & Fields, 2015; Arntzen, Norbom, & Fields, 2015; Arntzen & Steingrimsdottir, 2017; Dickins, 2015; Eilifsen & Arntzen, 2009; Fienup & Dixon, 2006; Leader, Barnes-Holmes, & Smeets, 2000; Pilgrim & Galizio, 1996).

Sorteringstester har vist stor nytteverdi både av generaliserings- og generalitets-hensyn (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, Norbom, et al., 2015; Fields, Arntzen, & Moksness, 2014). Opprinnelig har sorteringstester blitt presentert etter MTS tester for å undersøke om ekvivalensklassene kan generaliseres til et annet testformat. Det har flere ganger blitt dokumentert samsvar mellom emergens av klasser i MTS tester og sorteringstester, og korrekt respondering på MTS tester har dermed blitt forstått som en indikator for korrekt respondering på sorteringstester (Arntzen et al., 2017; Arntzen, Norbom, et al., 2015; Fields et al., 2014). Arntzen, Norbom, et al. (2015) presenterte sorteringstester etter trening av baselinerelasjoner for å undersøke om det var mulig å dokumentere umiddelbar emergens av ekvivalensklasser i andre formater enn MTS format. Tre deltakere viste umiddelbar emergens av klassene, men Arntzen, Norbom, et al. (2015) påpekte faren for interaksjonseffekter fordi disse deltakerne hadde erfaring med emergens av andre ekvivalensklasser tidligere i den samme eksperimentet.

Arntzen et al. (2017) kontrollerte for interaksjonseffekter som nevnt ovenfor ved å benytte en mellomgruppe design, og hvor deltakerne i den ene gruppen ble eksponert

<sup>1</sup>Bokstav og tallkombinasjonen foran skråstreken illustrerer utvalgsstimulus, mens den understrekte bokstavtallkombinasjonen illustrerer den eksperimentatorkorrekte sammenligningsstimulus.

for sorteringstester rett etter trening av baseline relasjoner, uten å ha erfaring med MTS tester. Ti forsøkspersoner i Gruppe 1 ble eksponert for sorteringstester, to blokker med MTS tester og til slutt sorteringstester. I Gruppe 2 ble 10 forsøkspersoner eksponert for to blokker med MTS tester, sorteringstester, og til slutt to blokker med MTS tester. Fem av deltakerne i Gruppe 1 viste umiddelbar emergens 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 emergens av de tre eksperimentatordefinerte klassene da sortering var andre testformat. To deltakere i hver gruppe viste emergens av de tre eksperimentatordefinerte klassene på sorteringstesten, men viste ikke opprettholdelse av klassene i påfølgende MTS test (Arntzen et al., 2017). Data har vist at noen deltakere som ikke nådde mestringskriteriet for stimulusekvivalens likevel viste emergens av en, to eller tre klasser i påfølgende sorteringstester (Arntzen et al., 2017; Arntzen & Steingrimsdottir, 2017; Fields et al., 2014; Mensah & Arntzen, 2017).

Det har blant annet blitt diskutert om forskjeller i henhold til eksperimentatordefinerte klasser i 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 mindre strikt stimuluskontroll. Sistnevnte har vært brukt som argument for at sorteringstestene dermed er mer sensitive for å avdekke klasseslithørighet (Arntzen et al., 2017; Arntzen, Norbom, et al., 2015; Fields et al., 2014). 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 klasseslithørigheten 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.

Effektene av ulike treningsstrukturer har vært omtalt innen forskningen på stimulusekvivalens. Opprinnelig så var det en antakelse om at de tre forskjellige strukturene, *Linear Series* (LS), *Many-to-One* (MTO) og *One-to-Many* (OTM) ikke skulle ha betydning for respondering i henhold til stimulusekvivalens, mens det senere er vist i en rekke forskningsrapporter at det er en forskjell. Med bakgrunn i eksemplet tidligere i denne artikkelen vil LS innebære at man trener  $A \rightarrow B$  og deretter  $B \rightarrow C$ . MTO innebærer at det blir trent  $A \rightarrow C$  og  $B \rightarrow C$ , mens OTM innebærer trening av  $A \rightarrow B$  og  $A \rightarrow C$ . LS er den treningsstrukturen som har vist seg å være den minst effektive strukturen med hensyn til respondering i henhold til stimulusekvivalens (se Arntzen, 2012, for en oversikt og diskusjon).

I flere studier hvor sorteringstester har vært inkludert i forbindelse med undersøkelser av emergens og opprettholdelse av ekvivalensklasser, har LS blitt benyttet (Arntzen et al., 2017; Arntzen, Norbom, et al., 2015; Fields et al., 2014). Ettersom det i disse nevnte studiene med LS har vært få personer som har respondert i henhold til stimulusekvivalens og samtidig sortert korrekt i forhold de eksperimentatordefinerte klassene, vil det være nyttig å anvende en treningsstruktur som gir en mye høyere andel personer som responderer i henhold til stimulusekvivalens for å få bekreftet antakelsen om dette samsvaret mellom de to testformatene (MTS og sortering).

Det presenterte eksperimentet er en systematisk replikasjon av Arntzen et al. (2017). Hovedformålet med eksperimentet var å undersøke robustheten av samsvaret mellom respondering på de to testformatene MTS og sortering ved å bruke MTO. Videre om det var mulig å replikere resultatene med partisjonering av stimulusklasser i henhold til

de eksperimentatordefinerte klasser ved bruk av sorteringstester når disse kom direkte etter treningen hvor de betingede diskriminasjonene ble etablert. I tillegg ville vi undersøke om stimulusklassene i sorteringstesten var løsere definert enn ekvivalensklassene i MTS testene, og det ble derfor administrert to blokker med sorteringstester i hver testfase for sortering.

## Metode

### Deltakere

Tjuen voksne personer deltok i eksperimentet, hvor alle deltakere ble rekruttert blant personlige kontakter. Deltakerne utgjorde syv kvinner og 14 menn fra 24 til 60 år, med en gjennomsnittsalder på 36 år. 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 stimuluskvivalens fra tidligere.

Ved ankomst til eksperimentlokalet ble deltakerne bedt om å lese og skrive under på en informert samtykkeklæring. 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 at eksperimentet var forventet å vare i maksimum to timer. Etter at eksperimentet var avsluttet fikk deltakerne en gjennomgang av egne resultater, en innføring i stimuluskvivalens, 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 artikkel om stimuluskvivalens skrevet på norsk av andreforfatter.

### Setting

Det ble benyttet to eksperimentrom. Det ene var forskningsgruppens laborato-

rium ved universitetet, og det andre var i en privat bolig. Forskningslaboratoriet ved universitetet 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. I den private boligen var rommet 380 x 240 cm og møblert med et skrivebord på 155 (b) x 85 (d) cm og en stol, og inneholdt i tillegg en bokhylle og en safe. Mens eksperimentene pågikk var det ingen kjente potensielt forstyrrende stimuli i eller i nærheten av eksperimentlokalene. Deltakerne brukte fra en og en halv til to timer på å gjennomføre eksperimentøkten inklusive nødvendige pauser. De som ikke var ferdige med trening av baseline relasjoner innen det hadde gått to og en halv timer fikk tilbud om å avbryte. Det ble sørget for at alle deltakerne hadde drikkevann tilgjengelig i umiddelbar nærhet under hele forsøket. Eksperimentator satt alltid utenfor døren til eksperimentlokalet mens eksperimentet pågikk.

### Utstyr

Det ble benyttet en 15 tommer Hewlett Packard bærbar datamaskin med operativsystem Windows 8 (64-bit) og mus for å klikke på stimuli på skjermen. To dataprogram ble benyttet i eksperimentet. 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. Når 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 feilresponser per trial og testblokk, samt forekomst av programmerte konsekvenser under trening. 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

Tabell 1. Rekkefølge på betingelser for deltakere i Gruppe 1 og 2.

Fase	1	2	3	4
Gruppe 1	TBR	SRT 1a og 1b	MTS 1a og 1b	SRT 2a og 2b
Gruppe 2	TBR	MTS 1a og 1b	SRT 1a og 2a	MTS 2a og 2b

Merknad. TBR = trening av baseline relasjoner; SRT 1a og 1b = første og andre sorteringstest i den første runden; MTS 1a og 1b = første og andre testblokk i den første matching-to-sample testfasen; SRT 2a og 2b første og andre sorteringstest i den andre andre runden; MTS 2a og 2b = første og andre testblokk i den andre matching-to-sample testfasen.

benyttet var en spesiallaget programvare for sorteringstester (SRT 1a, 1b, 2a og 2b). Sorteringsprogrammet styrte presentasjon av stimuliene i en stablet bunke midt på skjermen, og lagret data i form av skjerm-bilder 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 en mellom-gruppe design med randomisert fordeling av deltakere til en av to grupper (se Tabell 1). I Gruppe 1 var det 11 deltakere, mens det var 10 deltakere i Gruppe 2. Alle deltakerne startet med samme fase med trening av baseline relasjoner (TBR). Deltakerne i Gruppe 1 gikk videre igjennom en fase med to tester i sorteringsformat (SRT 1a og SRT 1b), 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 2a og SRT 2b). 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 1a og SRT 1b) og avslutningsvis en fase med to tester i *matching-to-sample*

format (MTS 2a og MTS 2b).

### Prosedyre

**MTS trening.** 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.

	1	2	3
A			
B			
C			
D			
E			

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

Tabell 2. Trenings- og testfase i eksperimentets betingede diskriminasjonsprosedyre.

Fase	Type relasjoner	Programmerte konsekvenser (%)	Minimum trials	Mestringskriterium (%)
TBR (etablert)	BR A1E1, A2E2, A3E3 B1E1, B2E2, B3E3 C1E1, C2E2, C3E3 D1E1, D2E2, D3E3	100	60	100
TBR (oppnått- holdning)	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	100	95 <sup>a</sup>

Merknad. Programmerte konsekvenser indikerer prosentvis antall trials med programmerte konsekvenser 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 mikset i hver blokk.

<sup>a</sup> Det ble kontrollert for at deltakerne som responderte i henhold til stimulusekvivalens med total testskåre på 95 %, også hadde minimum 103 (95 %) korrekte ekvivalens-trials.

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.

De betingede relasjonene mellom stimuliene ble etablert ved bruk av en MTO

treningsstruktur og en simultan trenings- og testprotokoll. Dette innebar at baselinere-lasjonene AE, BE, CE, og DE ble presentert randomisert i blokker før testene ble introdusert.

Treningsfasen (TBR) bestod av blokker med 60 trials hvor hver av de 12 ulike typene av trials ble presentert fem ganger i randomisert rekkefølge (se øverste del i Tabell 2). Mestringskriteriet i TBR var 60 av

60 korrekte trials i en blokk. En feilrespons 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 for hver trial i en blokk, ble antall programmerte konsekvenser redusert til 50 % per blokk. Etter oppnådd mestringskriterium med kun 50 % programmerte konsekvenser per blokk, bestod den siste delen av treningen av 0 % programmerte konsekvenser per blokk. Programmerte konsekvenser ble vist på skjermen i 500 ms, etterfulgt av et inter-trial intervall på 1000 ms.

**MTS test.** Ingen programmerte konsekvenser ble gitt underveis i testen. Hver testblokk inneholdt 180 trials hvor av 36 var baseline relasjoner (AE, BE, CE og DE), 36 var symmetri relasjoner (EA, EB, EC og ED) og 108 relasjoner hvor det ble testet for stimulusekvivalens (AB, AC, AD, BA, BC, BD, CA, CB, CD, DA, DB og DC) (se nederste del i Tabell 2). Hver type av trials 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 ekvivalens trials.

**Sortering.** 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*. Når deltakerne 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 skjerm bilde 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 inndeling 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 1b ble sjekket fortløpende, mellom presentasjonen av de ulike 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 eksperimentet. 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.

## Resultater

### Etablering av baselinerelasjoner

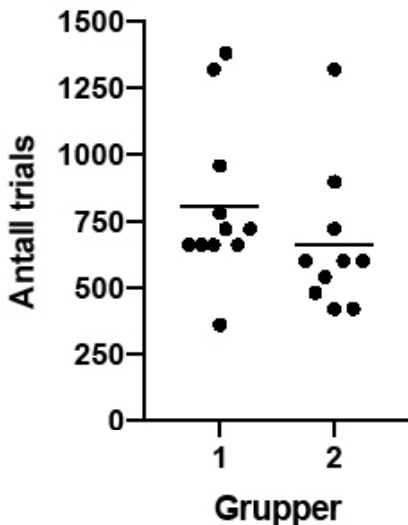
Analyse av data basert på gruppetilhørighet viste et gjennomsnitt på 807 trenings-trials i Gruppe 1, med en variasjonsbredde fra 360 til 1380 trials (se Figur 2). I Gruppe 2 var gjennomsnittet på 660 trials, med en

Tabell 3. Oppsummerte resultater på tvers av faser for deltakerne i Gruppe 1.

D#	Sortering		MTS		Sortering	
	1a	1b	1a	1b	2a	2b
16020	500 050 005	500 050 005	99	100	500 050 005	500 050 005
16029	500 050 005	500 050 005	100	99	500 050 005	500 050 005
16018	500 050 005	500 050 005	99	99	500 050 005	500 050 005
16019	500 050 005	500 050 005	99	99	500 050 005	500 050 005
16029	500 050 005	500 050 005	97	99	500 050 005	500 050 005
16031	500 050 005	500 050 005	98	97	500 050 005	500 050 005
16025	500 050 005	500 050 005	96	99	400 150 005	500 050 005
16023	400 140 015	500 040 015	97	99	500 050 005	500 050 005
16027	555 000 000	555 000 000	100	100	500 050 005	500 050 005
16027	IA <sup>a</sup>	300 050 005 200	83	88	510 040 005	500 050 005
16023	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. Hvert tall i MTS kolonnene er prosent korrekt respondering og uthevede tall betyr at deltakerne responderte i henhold til stimulusekvivalens. Tall i sorteringstester viser deltakerdefinerte klasser hos hver deltaker. Uthevet skrift betyr at deltakerdefinerte klasser ble sortert i samsvar med de eksperimentatordefinerte klassene.  
<sup>a</sup> Ikke anvendbare data.

variasjonsbredde fra 420 til 1320 trials. En uavhengig t-test viste at det ikke var en statis-



Figur 2. Figuren viser antall trials til mestringskriterium. Hver prikk er antallet for hver av forsøkspersonene. Streken indikerer gjennomsnittet for gruppene.

tisk 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.

#### MTS test data

Atten av 21 (85%) deltakere responderte i henhold til stimulusekvivalens på MTS test 1a og 1b (se Tabell 3 og 4). To av 21 deltakere (D16023 og D16027) responderte ikke i henhold til stimulusekvivalens på noen av MTS testene. En deltaker (D16024) responderte i henhold til stimulusekvivalens kun i siste testfase etter erfaring med sorteringstester. Det ble benyttet en *Fisher's exact test* (FET) som viste at det ikke var statistisk signifikant forskjell mellom gruppene med hensyn til antall deltakere som responderte i henhold til stimulusekvivalens på MTS test 1,  $p > .05$ .

#### Sorteringsdata



Tabell 4. Oppsummerte resultater på tvers av faser for deltakerne i Gruppe 2

D#	MTS		Sortering		MTS	
	1a	1b	1a	1b	2a	2b
16028	100	100	500 050 005	500 050 005	100	100
16021	100	100	500 050 005	500 050 005	100	99
16022	100	100	500 050 005	500 050 005	99	99
16038	100	100	500 050 005	500 050 005	100	99
16032	100	99	500 050 005	500 050 005	99	99
16036	99	100	500 050 005	500 050 005	99	100
16034	99	100	500 050 005	500 050 005	99	100
16026	99	99	500 050 005	500 050 005	100	100
16035	99	99	500 050 005	500 050 005	99	100
16024	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. Hvert tall i MTS kolonnene er prosent korrekt respondering og uthevede tall betyr at deltakerne responderte i henhold til stimulusekvivalens. Tall i sorteringstester viser deltakerdefinerte klasser hos hver deltaker. Uthevet skrift betyr at deltakerdefinerte klasser ble sortert i samsvar med de eksperimentatordefinerte klassene. En deltaker trakk seg tidlig under trening av baselinereelasjonene. Data fra denne deltakeren er derfor utelatt.

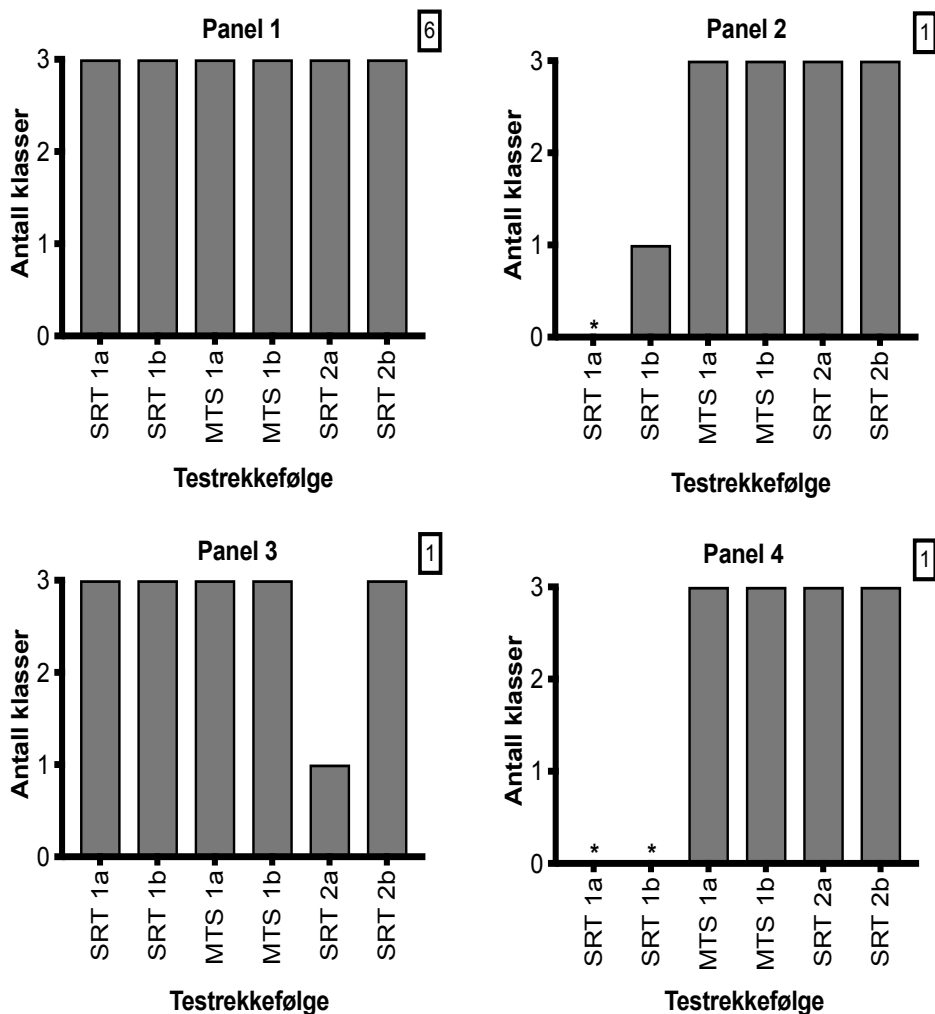
For å beskrive sorteringsresultatene kvantitativt ble det benyttet et trestrengssystem bestående av klynger med tall, der hver klynge indikerte gruppene i sorteringen (se Arntzen et al., 2017, for detaljer). Tallene i hver klynge viser hvor mange stimuli fra hver klasse som var plassert i gruppen. Sorteringsdata som reflekterte de eksperimentatordefinerte klassene ble notert som 500 050 005, det vil si at alle fem stimuliene fra Klasse 1 ble gruppert sammen, alle 5 stimuliene fra Klasse 2 ble gruppert sammen og alle stimuliene fra Klasse 3 ble gruppert sammen. Eksperimentator (førsteforfatter) og en medobservatør evaluerte uavhengig av hverandre skjerm-billedata fra alle sorteringstestene på hvert sitt noteringsark. Disse registreringene viste 100% enighet mellom observatørene på sorteringsdata for begge gruppene.

16 av 21 (76%) deltakere viste emergens av de tre eksperimentatordefinerte klassene på alle sorteringstestene, uavhengig om sortering var første eller andre testfase (se Tabell

3 og 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 1a hos D16027 ikke var anvendbar som sammenligningsgrunnlag, ble sorteringsdata fra sorteringstest 1b i Gruppe 1 vurdert opp mot sorteringsdata fra sorteringstest 1b i Gruppe 2. Det ble benyttet en 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 for deltakere i Gruppe 1

Syv av 11 deltakere viste umiddelbar emergens av de tre eksperimentatordefinerte klassene på sorteringstest 1a og viste opprettholdelse av klassene på sorteringstest 1b (se de to panelene til venstre i Figur 3). Samtlige av de syv deltakerne viste emergens



Figur 3. Umiddelbar emergens av tre klasser når 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. \* betyr at ingen klasser er framvist.

av tre ekvivalensklasser på de to påfølgende blokkene med MTS tester. Seks av syv deltakere viste opprettholdelse av de tre eksperimentatordefinerte klassene på sorteringstest 2a. På den siste sorteringstesten (2b) viste alle 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 emergens av de eksperimentatordefinerte klassene da det første testformatet var sortering. Data fra to av disse deltakerne vises i de to panelene på høyre side i Figur 3. De to deltakerne 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. De to andre av de fire deltakerne viste andre responsmønstre (se

omtale av disse senere).

### Umiddelbar emergens av klasser for deltakere i Gruppe 2

Ni av 10 deltakere viste umiddelbar emergens av de tre ekvivalensklassene når MTS var første testformat (se venstre panel i Figur 4). Samtlige av disse deltakerne viste opprettholdelse av de tre eksperimentator-definerte klassene under påfølgende faser med sortering og MTS tester. En deltaker i Gruppe 2 viste ikke emergens av ekvivalensklassene i første testfase med MTS tester, men viste emergens av tre eksperimentator-definerte klasser i den påfølgende fasen på sorteringstest 1a og 1b (se høyre panel i Figur 4). Denne deltakeren responderte i henhold til ekvivalens på siste fase med MTS tester.

### Samsvar mellom sorteringstestene i hver testfase for deltakere som responderte i henhold til ekvivalens

I Gruppe 1 var det 100% samsvar i respondering mellom sorteringstestblokkene 1a og 1b hos åtte av ni deltakere (se panel 1, 3 og 4 i Figur 3). En deltaker viste endring fra ingen korrekt sorte klasser i sortering 1a til at en klasse var sortert korrekt i sortering 1b (se panel 2 i Figur 3). Som vist i de fire panelene

i Figur 3 var det 100% samsvar i prestasjoner mellom sortering 2a og 2b hos åtte av ni deltakere. En deltaker viste forbedring i sortering 2a og 2b, fra to korrekte klasser tre korrekte klasser (se panel 3 i Figur 3).

Hos samtlige deltakere i Gruppe 2 var det et 100% samsvar i prestasjoner på sorteringstestene 1a og 1b (se Figur 4).

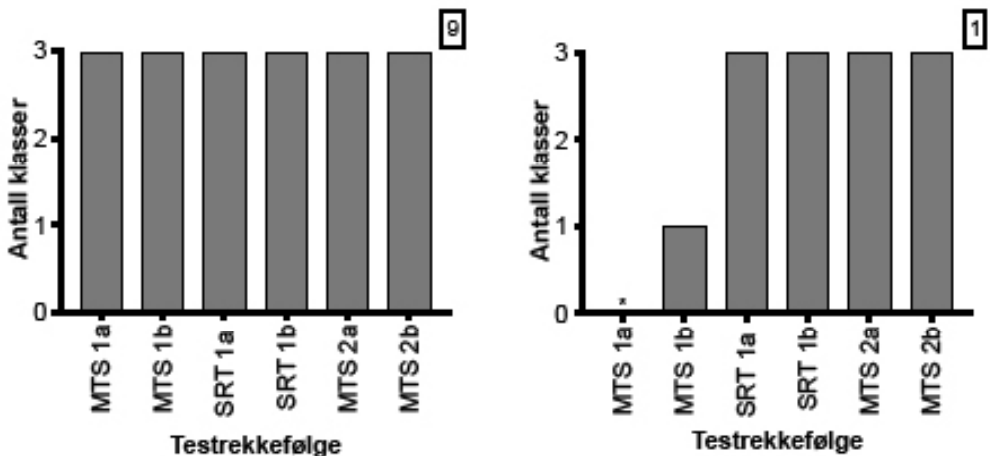
### Analyser av respondering hos deltakere som ikke responderte i henhold til ekvivalens

Vi gjorde utvidete analyser av respondering hos to deltakere, D16027 og D16023 i Gruppe 1, som ikke responderte i henhold til stimulusekvivalens. Det ble undersøkt om en eller to av klassene likevel var framkommet i MTS testen hos disse deltakerne, og om disse klassene samsvarte med inndeling av klassene i sorteringstesten.

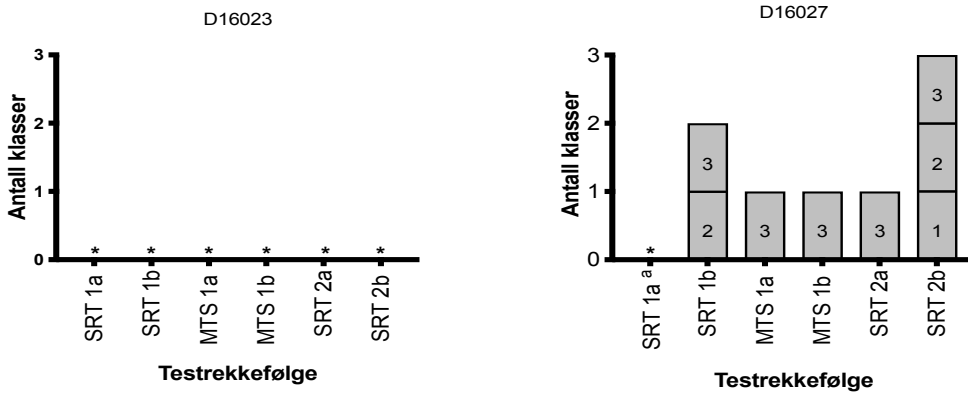
**Emergens av noen av klassene.** D16023 viste ikke emergens av noen av ekvivalensklasser på MTS testene. D16027 viste kun emergens av Klasse 3 på MTS testene (se Figur 5).

### Samsvar mellom MTS og sorteringstestene.

D16023 viste kun samsvar av to rela-



Figur 4. 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. \* betyr at ingen klasser er framvist.



Figur 5. Responsmønstre for deltakere som ikke responderte i henhold til stimuluskvalens, to deltakere i Gruppe 1. SRT = sorteringstester; MTS = matching-to-sample tester. <sup>a</sup> Data for SRT 1a hos deltaker 16027 var ikke anvendbare. \* betyr at ingen klasser er framvist.

sjoner (C2D2 og A3B3) mellom respondering på MTS testene og inndeling i de deltakerdefinerte klassene i sorteringstesten, men altså ikke tilstrekkelig korrekt respondering på noen av testene. D16027 viste kun samsvar mellom respondering på de to testformatene med hensyn til Klasse 3.

## Diskusjon

I dette eksperimentet ble 12 betingede diskriminasjoner trent i et MTS format hos 21 forsøkspersoner. Deretter ble det testet for emergens av tre ekvivalensklasser med fem medlemmer. Umiddelbar partisjonering av de eksperimentatordefinerte klassene ble dokumentert med bruk av sorteringstester hos over halvparten av deltakerne i Gruppe 1. Ni av 10 deltakere i Gruppe 2 viste umiddelbar emergens av stimuluskvalens når det første testformatet var MTS. Data viste samsvar i prestasjoner mellom de to testblokkene i hver sorteringstest på totalt 28 av 32 presenterte sorteringstester, som er en replikasjon av Arntzen et al. (2017).

### Samsvar i respondering på MTS tester og sorteringstester

Resultatene fra eksperimentet replikerer funnene fra tidligere eksperimenter om samsvaret mellom korrekt respondering

på MTS tester og korrekt respondering på sorteringstester (Arntzen et al., 2017; Arntzen et al., 2014; Arntzen, Norbom, et al., 2015; Arntzen & Steingrimsdottir, 2017; Fields et al., 2014; Nartey, Arntzen, & Fields, 2015).

Videre viste resultatene et samsvar mellom korrekt respondering ved sorteringstester og korrekt respondering 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).

Resultatene i dette eksperimentet er spesielt viktige ettersom de underbygger at deltakere som responderer i henhold til stimuluskvalens vil også sortere i forhold til de eksperimentatordefinerte klassene når det brukes MTO, en treningsstruktur som produserer en høy andel personer som responderer i henhold til stimuluskvalens. Det kan altså ikke være slik at den høye korrespondansen mellom de to testformatene er vist i tidligere eksperimenter med LS (Arntzen et al., 2017; Arntzen, Norbom, et al., 2015; Fields et al., 2014) skyldes at relativt få deltakere responderte i henhold til ekvivalens på MTS testene,

men at denne korrespondansen er et relativt robust funn.

### **Samsvar 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. Samtidig som de klassene som dokumenteres med sorteringstester er løserer definert enn klassene som dokumenteres i MTS testene (Arntzen et al., 2017; Arntzen, Norbom, et al., 2015; Fields et al., 2014). I dette eksperimentet var det stor grad av samsvar i prestasjoner mellom hver testblokk i de korrekt suksessivt presenterte sorteringstestblokkene, og det var dermed ikke grunnlag for å si at klassene som ble påvist med sorteringstester var løserer 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, bortsett fra D16025 som viste bortfall av en klasse på sorteringstest 2a da en stimulus var sortert i feil klasse (se Tabell 3). Siden D16025 hadde sortert korrekt på de to forutgående blokkene med sorteringstester, responderte korrekt i to blokker med MTS tester, samt sorterte korrekt i den siste sorteringstestblokken (2b), er det sannsynlig at fravær av korrekt sortering av Klasse 1 på sorteringstest 2a skyldes at eksperimentsituasjonen ble utmattende eller lignende.

### **Umiddelbar emergens av klasser**

Dersom resultatene fra sorteringstesten viste at deltakerne i Gruppe 1 sorterte stimuliene i de eksperimentatordefinerte klassene og samtidig responderte i henhold til stimulusekvivalens på MTS testen, så er det rimelig å si at dette er eksempler på umiddelbar emergens av ekvivalensklasser vist ved sortering. Antall personer som viste dette responsmønsteret var mer enn dobbelt så mange som i Arntzen et al. (2017). En tilsvarende forskjell gjaldt for deltakerne i Gruppe 2 som hadde MTS test etter den betingede diskriminasjonstreningen. Årsaken til disse forskjellene kan tilskrives at det ble

brukt en annen treningsstruktur, LS. Denne påstanden er begrunnet i at bruken av LS har vist i en rekke eksperimenter å gi et lavere antall forsøkspersoner som responderer i henhold til stimulusekvivalens (Arntzen, Grøndahl, & Eilifsen, 2010; Arntzen & Hansen, 2011; Fields, Hobbie-Reeve, Adams, & Reeve, 1999). Samtidig vil en ikke kunne forvente andre resultater ved bruk OTM ettersom forskjellen mellom MTO og OTM har vist seg å være minimale (e.g., Ayres Pereira & Arntzen, i trykk).

### **Utsatt emergens**

At deltakere ikke responderer i henhold til stimulusekvivalens på MTS tester, men likevel viser emergens av de eksperimentatordefinerte klassene i sorteringstestene er et vanlig funn (Arntzen et al., 2017; Arntzen, Norbom, 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 (Dube & McIlvane, 1996).

D16024 sitt responsmønster kan fortolkes om et eksempel på utsatt emergens av ekvivalensklasser. Slike data er interessante fordi de viser at det skjer en endring 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 and McIlvane (1996) foreslo at feilresponser 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 klasse-medlemskap 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).

Fields og medarbeidere (e.g., Fields & Moss, 2007) argumenterer for at utsatt emergens kan skyldes at diskriminasjonen mellom de ulike stimulussettene ikke er strikt nok i den betingede diskriminasjonstreningen på tross av at mestringskriterium før testing av emergente relasjoner er oppnådd. Forandringen i løpet av testen til det som kalles utsatt emergens skyldes en økt diskriminasjon mellom stimulussettene.

Mye av forskningen vedrørende utsatt emergens har stort sett vært deskriptiv, for eksempel det er bare rapport om antall forekomster av utsatt emergens (e.g., Arntzen & Nartey, 2018; Varelas & Fields, 2015). Derfor er det påpekt at det er vesentlig med forskning som fokuserer på de mekanismene som ligger til grunn for utsatt emergens. For eksempel er sannsynlige mekanismer for at utsatt emergens forekommer: (1) hvordan og mengden av baseline trials presenteres i testblokkene og (2) diskriminasjon av stimuli på tvers av klassene (Arntzen & Mensah, akseptert avhengig av revidering).

### **Forskjeller mellom testformatene**

Ulike stimuluskontrolltopografier kan ses i sammenheng med forskjeller i trenings- og testbetingelser, og potensielle forskjeller mellom de to typene testformatene, MTS og sortering. Det har 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 stimuluskvalensklasser, slik de defineres ved

demonstrasjon av egenskapene refleksivitet, symmetri og transitivitet. 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 trial 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, vil resultatet kun si noe om hvordan forsøkspersonene kategoriserer stimuliene. Dette avgjøres på grunnlag av om stimuliene 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 korrekt respondering på de to testformatene. En gitt prosentandel korrekte testtrials som indikator på respondering i henhold til stimuluskvalens, gir rom for noen få feilresponser til hver relasjonstype på MTS testene. På sorteringstestene derimot, kan en stimulus plassert feil ødelegge for emergens av flere klasser.

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. Framtidig forskning vil kunne teste ut forskjellene mellom testformatene ved å organisere sorteringstesten slik at stimuli plasseres i bunker i stedet for å plassere dem ved siden av hverandre.

Oppsummert fra forskningen hvor en sammenligner respondering på de to testformatene viser resultatene for det første 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 respondering på de to testformatene, og hvor emergent atferd påvises i størst grad med sorteringstester. Pilgrim påpeker

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 ved testing for emergente relasjoner. Dette er interessant i forbindelse med sorteringstester, spesielt i tilfellene hvor man ser partisjonering av alle de eksperimentatordefinerte klassene i sorteringstestene, men uten at forsøkspersonen responderer i henhold til stimulusekvivalens på MTS tester.

### Anvendelse

Det er en rekke forhold som gjør at bruken av sorteringstester er viktig når det gjelder inndeling i stimulusklasser. For eksempel kan sorteringstester være enklere å administrere hos små barn (Barron, Leslie, & Smyth, 2018) eller når man skal dokumentere emergent atferd i settinger med mange deltakere (Varelas & Fields, 2017). Videre er det ikke alltid slik at man i en del anvendte sammenhenger trenger å dokumentere at personer responderer i henhold til den formelle definisjonen av stimulusekvivalens. Derimot kan det være tilstrekkelig at personene klarer å gruppere stimuli i henhold til klasser etter at den betingede diskriminasjonstreningen er mestret. Et siste argument for å bruke sorteringstester som indikasjon på klasseinndeling er at en sorteringstest tar omtrent 1/10 av tiden som en MTS test (Arntzen et al., 2017).

### Oppsummering

I dette eksperimentet ble det bekreftet at sorteringstester kan benyttes til å avdekke umiddelbar emergens av stimulusklasser etter trening av baseline relasjoner i MTS format. Resultatene viste en 100% sammenheng mellom respondering på MTS testen og sorteringstesten når det ble brukt en treningsstruktur som er effektiv med hensyn til antall personer som responderer i henhold til stimulusekvivalens. Dette tyder på at samsvaret mellom respondering på de to

testformatene er robust. Videre var det stor grad av samsvar mellom respondering som var korrekt i to suksessivt presenterte sorteringstester i hver sorteringstestfase, samt samsvar mellom respondering på sorteringstestene og MTS testene.

Partisjonering eller inndeling i stimulusklasser som påvises gjennom sorteringstester etter betinget diskriminasjonstrening, men som teknisk sett ikke er testet for egenskapene refleksivitet, symmetri og transitivitet, kan være viktig i bruk opplæringsammenhenger ettersom MTS tester kan ta relativt lang tid å gjennomføre. Sorteringstester derimot er raske å administrere.

### Overholdelse av etiske retningslinjer

**Finansiering:** Forskingen ble støttet av OsloMet — storbyuniversitetet.

**Interessekonflikter:** Forfatterne erklærer at det er ingen interessekonflikter.

**Forskningsetiske forhold:** Alle prosedyrer er utført i henhold til etiske retningslinjer og er i overenstemmelse med 1964 Helsinki Deklarasjonen og dens senere tilføyelser.

**Informert samtykke:** Informert samtykke ble innhentet fra alle deltakerne i det presenterte eksperimentet.

### Referanser

- Arntzen, E. (2004). Probability of equivalence formation: Familiar stimuli and training sequence. *The Psychological Record, 54*, 275–291. Retrieved from <https://link.springer.com/article/10.1007/BF03395474>.
- Arntzen, E. (2012). Training and testing parameters in formation of stimulus equivalence: Methodological issues. *European Journal of Behavior Analysis, 13*, 123–135. doi: 10.1080/15021149.2012.11434412
- Arntzen, E., Granmo, S., & Fields, L. (2017). The relation between sorting tests and matching-to-sample tests in the formation of equivalence classes. *The Psychological Record, 67*, 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 the subsequent performance on the tests for stimulus equivalence. *The Psychological Record*, *60*, 437–462. Retrieved from <http://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1000&context=tp>.
- Arntzen, E., & Hansen, S. (2011). Training structures and the formation of equivalence classes. *European Journal of Behavior Analysis*, *12*, 483–503. doi: 10.1080/15021149.2011.11434397
- Arntzen, E., & Mensah, J. (akseptert avhengig av revideringer). On the effectiveness of including meaningful pictures in the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior*.
- Arntzen, E., & Narthey, R. K. (2018). Equivalence class formation as a function of preliminary training with pictorial stimuli. *J Exp Anal Behav*, *110*, 275–291. doi: 10.1002/jeab.466 Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/30182379>.
- Arntzen, E., Narthey, 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 Retrieved from <http://dx.doi.org/10.1007/s40732-014-0066-3>.
- Arntzen, E., Narthey, R. K., & Fields, L. (2015). Enhanced equivalence class formation by the delay and relational functions of meaningful stimuli. *Journal of the Experimental Analysis of Behavior*, *103*, 524–541. doi: 10.1002/jeab.152
- Arntzen, E., Norbom, A., & Fields, L. (2015). Sorting: An alternative measure of class formation? *The Psychological Record*, *65*, 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*. doi: 10.3389/fnhum.2017.00058 Retrieved from <http://journal.frontiersin.org/article/10.3389/fnhum.2017.00058>.
- Ayres Pereira, V., & Arntzen, E. (i trykk). Effect of presenting baseline probes during or after emergent relations tests on equivalence class formation. *The Psychological Record*.
- Barron, R., Leslie, J. C., & Smyth, S. (2018). Teaching Real-World Categories Using Touchscreen Equivalence-Based Instruction. *The Psychological Record*. doi: 10.1007/s40732-018-0277-0 Retrieved from <https://doi.org/10.1007/s40732-018-0277-0>.
- 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*, 461–475. doi: 10.1901/jaba.1992.25-461
- 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*, 1–11. doi: 10.1007/s40732-015-0134-3 Retrieved from <http://dx.doi.org/10.1007/s40732-015-0134-3>.
- Dube, W. V., & McIlvane, W. J. (1996). Some implications of a stimulus control topography analysis. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus Class Formation in Humans and Animals* (pp. 197–218). Amsterdam: Elsevier.
- Dymond, S., & Rehfeldt, R. A. (2001). Supplemental measures and derived stimulus relations. *Experimental Analysis of Human Behavior Bulletin*, *19*, 8–12. Retrieved from <http://www.eahb.org/NewSitePages/BulletinHomepage.htm>.
- Eilifsen, C., & Arntzen, E. (2009). On the



- role of trial types in tests for stimulus equivalence. *European Journal of Behavior Analysis*, *10*, 187–202. Retrieved from <http://www.ejoba.org/>.
- Fields, L., Arntzen, E., & Moksness, M. (2014). Stimulus Sorting: A quick and sensitive index of equivalence class formation. *The Psychological Record*, *64*, 487–498. doi: 10.1007/s40732-014-0034-y
- 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*, 703–724. Retrieved from <http://thepsychological-record.siuc.edu/>.
- Fields, L., & Moss, P. (2007). Stimulus relatedness in equivalence classes: Interaction of nodality and contingency. *European Journal of Behavior Analysis*, *8*, 141–159. doi: 10.1080/15021149.2007.11434279
- Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of the Experimental Analysis of Behavior*, *48*, 317–332. doi: 10.1901/jeab.1987.48-317
- Fienup, D. M., & Dixon, M. R. (2006). Acquisition and maintenance of visual-visual and visual-olfactory equivalence classes. *European Journal of Behavior Analysis*, *6*, 87–98.
- 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.
- Leader, G., Barnes-Holmes, D., & Smeets, P. M. (2000). Establishing equivalence relations using a respondent-type training procedure III. *The Psychological Record*, *50*, 63–78.
- Mensah, J., & Arntzen, E. (2017). Effects of meaningful stimuli contained in different numbers of classes on equivalence class formation. *The Psychological Record*, *67*, 325–336. doi: 10.1007/s40732-016-0215-y
- Nartey, R. K., Arntzen, E., & Fields, L. (2015). Enhancement of equivalence class formation by pretraining discriminative functions. *Learning & Behavior*, *43*, 20–31. doi: 10.3758/s13420-014-0158-6 Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/25425295>.
- Pilgrim, C. (2016). Considering definitions of stimulus equivalence. *European Journal of Behavior Analysis*, *17*, 105–114. doi: 10.1080/15021149.2016.1156312 Retrieved from <http://dx.doi.org/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 Formation in Humans and Animals* (pp. 173–195). Amsterdam: Elsevier.
- 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, New Jersey: Lawrence Erlbaum.
- Sidman, M. (1992). Equivalence relations: Some basic considerations. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 15–27). Reno, Nevada: Context Press.
- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston, MA: Authors Cooperative.
- Sidman, M. (2009). Equivalence relations and behavior: An introductory tutorial. *The Analysis of Verbal Behavior*, *25*, 5–18. Retrieved from <http://www.springer.com/psychology/psychology+general/journal/40616>.
- 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*, 5–22. doi: 10.1901/jeab.1982.37-5
- Varelas, A., & Fields, L. (2015). Induction of a Generalized Transitivity Repertoire Via Multiple-Exemplar Training and Staged Testing. *The Psychological Record*,

65, 595-614. doi: 10.1007/s40732-015-0129-0

*Psychological Record*, 67, 71-80. doi: 10.1007/s40732-016-0208-x

Varelas, A., & Fields, L. (2017). Equivalence Based Instruction by Group Based Clicker Training and Sorting Tests. *The*

---

## **How robust is the correspondence between responding on matching-to-sample tests and sorting tests?**

Kristiane Rustad Bevolden og Erik Arntzen  
OsloMet – Oslo Metropolitan University

The present experiment aimed to study the correspondence between the results on the matching-to-sample test and the sorting test when employing a Many-to-One (MTO) training structure. The present experiment is a systematic replication of earlier findings from Arntzen, Granmo, and Fields (2017) which showed a high correspondence when using the Linear Series (LS) training structure. Twenty-one participants were taught 12 conditional discriminations and tested for the emergence of three 5-member classes. In Group 1, baseline training was followed by two sorting test, two MTS test blocks and finally two sorting test. 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, the initial sorting test showed immediate partitioning of three experimenter-defined classes for 7 participants. Four participants showed other patterns of responding. In Group 2, the initial MTS test blocks showed immediate emergence of three equivalence classes for nine participants. One participant showed another pattern of responding. Twenty of the 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% correspondence in performance between test blocks on 28 of 32 presented sorting tests. The degree of correspondence 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 stimulus classes in sorting tests was discussed and the application of sorting tests. In conclusion, the present experiment extends the knowledge about the robustness of the correspondence between responding on matching-to-sample tests and sorting tests.

*Keywords:* Stimulus equivalence, sorting, stimulus classes, immediate emergence, delayed emergence