



UNIVERSITETET I AGDER

Betydningen av fysisk aktivitet for kognitiv funksjon og akademisk prestasjon

Et systematisk review

ERIK BORGERSEN

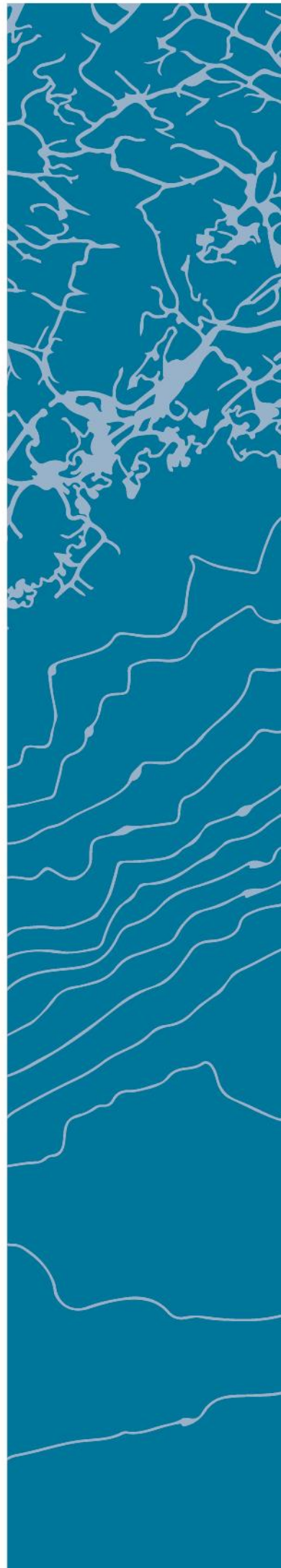
VEILEDER

Yngvar Ommundsen

Universitetet i Agder, 2018

Fakultet for Helse- og Idrettsvitenskap

Institutt for folkehelse, idrett og ernæring



Forord

Denne rapporten er del av masterstudiet mitt. Min utdanningsbakgrunn er elektronikingeniør fra Ingeniørhøgskolen i Telemark, cand. scient. grad i fysikk fra Universitetet i Oslo, og bachelorgrad i idrett fra Fakultet for helse- og idrettsfag ved Universitetet i Agder.

Jeg bestemte meg tidlig i masterprosessen for å gjøre en selvstendig litteraturstudie. Tema kom relativt seint, og tilhører ikke de tematiske valg som ble presentert for vårt kull. Jeg søkte veiledere, undersøkte hva den enkelte foreleser hadde publisert, og kom over en interessant artikkel av Yngvar Ommundsen. Han sa seg villig til å veilede meg mot en oppgave som omhandlet sammenhengen mellom fysisk aktivitet og kognisjon eller akademisk prestasjon.

Dette er et forskningsarbeid i kurset ME-517 Masteroppgave i idrettsvitenskap.

Masteroppgaven er skrevet som en monografi. Rapporten kan publiseres som en artikkel ved små endringer og ved å utelate teorikapittelet. Jeg vil takke min veileder for råd og samtaler, støtte og oppmuntring under arbeidet med oppgaven. Jeg vil takke familie og venner for iverrike diskusjoner underveis.

Sammendrag

Hensikt: Dette reviewet har til hensikt å presentere en oversikt over studier som bidrar til kunnskapsutvikling om betydningen av fysisk aktivitet for kognitiv funksjon eller akademisk prestasjon hos barn og unge i normal utvikling og alder til og med 18 år.

Teori: Status for forskningsfeltet blir introdusert i teorikapitlet.

Metode: Studien er et systematisk review. Kilde til relevante studier var databasene SPORTDiscus, CINAHL Plus with full text, MEDLINE, Academic Search Complete, Health and Psychosocial Instruments og ERIC. I dette systematiske reviewet ble 23 artikler inkludert. Metodologisk kvalitet ble vurdert. Koder for kategorier av intervensjonsulikhet ble utviklet. Gradert evidens ble utslagsgivende for troverdigheten i funn.

Resultater: En syntese ble utviklet for å presentere alle funn samlet. Syntesen viser funn av kognitiv funksjon (75.7%) og akademisk prestasjon (24.3%) med tilhørende metodologisk kvalitet. Tilstrekkelig metodologisk kvalitet ble funnet i 76.5% av studiene. En empiri-tro modell ble utviklet på grunnlag av syntesen. Modellen sammen med den nevrobiologiske eller psykososiale hypotesen støtter en kausal sammenheng. Progresjon i kognitiv funksjon etter akutt fysisk aktivitetsintervensjoner ble funnet i 63.6% av undersøkelsene. Blant kronisk fysisk aktivitetsintervensjoner ble progresjon i hhv. kognitiv funksjon og akademisk prestasjon påvist i 38.5% og 11.5 % av undersøkelsene.

Diskusjon: Alle studiene har merknader med betydning for ekstern validitet og generalisering. Alle studiene har merknader med betydning for internvaliditet og kausalitet. Kun 35 % av studiene er gjennomført med tilstrekkelig utvalgsstørrelse. Sensitivitetsanalysen som vektla mangler i søkeprosess og utvelgelse av studier, samt primærstudienes resultater og mangler relatert til publikasjonsavvik, viste at inkluderte studier var «sunne».

Konklusjon: Fysisk aktivitet har betydning for kognitiv funksjon eller akademisk prestasjon. Det systematiske reviewet oppnår en robusthet i evidens som gjør det mulig å anbefale intervensjoner hvor fysisk aktivitet er kognitivt beriket.

Abstract

Objectives: This review aims to give an overview of studies providing evidence for a relationship between physical activity and cognitive function or academic achievement in children of normal development ≤ 18 years of age.

Theory: The relevant research are introduced.

Methods: This study is a systematic review. SPORTDiscus, CINAHL Plus with full text, MEDLINE, Academic Search Complete, Health and Psychosocial Instruments and ERIC were searched for relevant articles. A total of 23 articles were included. Methodological quality was assessed. Codes on dissimilar categories of intervention differences were developed. The notion of credibility in outcomes came forth through four levels of evidence that considered methodological quality and the outcome of the studies.

Results: A synthesis were developed to present an assembly of all the findings. Quality score included, the synthesis provides findings of cognitive function (75.7%) and academic achievement (24.3%). Sufficient quality score was found in 76.5% of the studies. A model was developed on the bases of the synthesis. Both the model and the neurobiological or psycho social hypothesis support a causal relationship. Progression in cognitive function was found after acute physical activity interventions in 63.6% of the outcomes. Among the chronic physical activity interventions, progression in cognitive function and academic achievement were found in respectively 38.5% and 11.5% of the outcomes.

Discussion: All studies have remarks on external validity and generalization. All studies have remarks on internal validity and causality. Only 35% of the studies had adequate sample size. A sensitivity analysis examining the search method, eligibility of studies, study outcomes and questions in relation to publication bias, was found to be “sound”.

Conclusions: Physical activity has an effect on cognitive function and academic achievement. This systematic review provides a robustness in evidence that give implications for physical activity that contain cognitive enrichment.

Innholdsfortegnelse

Forord.....	ii
Sammendrag.....	iii
Abstract.....	iv
Innholdsfortegnelse.....	v
1.0 Innledning.....	1
1.1 Bakgrunn for tema.....	1
1.2 Formulering og drøfting av problemstilling.....	3
1.3 Disposisjon for oppgaven.....	4
2.0 Teoridel.....	6
2.1 Antatte forklaringsmekanismer mellom fysisk aktivitet og utvikling av mentale prosesser....	7
2.1.1 Den nevrobiologiske hypotesen.....	7
2.1.2 Den psykososiale hypotesen.....	8
2.1.3 Atferdshypotesen.....	10
2.2 Teoretiske modeller for betydningen av fysisk aktivitet på kognitive funksjoner og akademisk prestasjon.....	10
2.2.1 Metakognisjon og eksekutive funksjoner som mediator mellom fysisk aktivitet og akademisk prestasjon.....	11
2.2.2 Kognitiv læringsteori kan introdusere en aldersmoderator.....	11
2.2.3 Læreridentitet som mediator mellom fysisk aktivitet og kognisjon.....	12
2.3 Kognitive funksjoner - begrepsavklaring.....	13
2.3.1 Eksekutive funksjoner.....	13
2.3.2 Eksekutive funksjoner som dynamisk begrep.....	16
2.3.3 Eksekutive funksjoner i et utviklingsperspektiv.....	17
2.3.4 Operasjonalisering av eksekutive, ikke-eksekutive og metakognitive funksjoner.....	17
2.3.5 Måter å forbedre eksekutive funksjoner.....	19
2.4 Fysisk aktivitetsintervensjoner – tidligere forskning.....	19
2.4.1 Grunnleggende bevegelsesferdighet som intervensjon.....	19
2.4.2 Kvantitative og kvalitative, akutte og kroniske intervensjoner.....	20

2.4.3 Kroppslig kognisjon	20
2.4.4 Leke aktiviteter	20
2.4.5 Ordinær kroppøving som kontrollaktivitet	21
2.5 Betydningen av ulike former fysisk aktivitet for kognitive funksjoner og akademisk prestasjon	21
2.5.2 Kvantitative akutte intervensjoner	22
2.5.3 Kvantitative kroniske intervensjoner	22
2.5.4 Kvalitative akutte intervensjoner	23
2.5.5 Kvalitative kroniske intervensjoner	23
2.6 Måleresultater fra den senere tids forskning	25
2.6.1 Ikke-eksekutive og kjerne-eksekutive funksjoner	25
2.6.2 Metakognisjon	25
2.6.3 Delanalyser som undersøker effekten av ulike intervensjoner	26
3.0 Metode	26
3.1 Inklusjonskriterier	26
3.1.1 Valg av populasjon og utvalg	27
3.1.2 Valg av intervensjonsaktivitet	27
3.1.3 Kontroll eller sammenligningsaktivitet	27
3.1.4 Valg av utfall	27
3.1.5 Valg av studiedesign	28
3.2 Søkeprosessen	28
3.2.1 Valg av søkeord	28
3.2.2 Tverrfaglig databasesøk	29
3.2.3 Systematiske søk i enkeltdatabaser	30
3.3 Utvelgelse av litteraturreferanser ved PRISMA's flytdiagram	30
3.4 Beskrivelse av inkluderte studier	33
3.4.1 Deskriptiv informasjon	33
3.4.2 Metode for koding av inkluderte studier	34
3.5 Metodologisk kvalitet	36

3.5.1 Valg av sjekkliste for vurdering av metodologisk kvalitet	36
3.5.2 Anvendelsen av metodologisk kvalitet.....	37
3.5.3 Kategorisert metodologisk kvalitet og gradert evidens.....	37
3.6 Typen review	38
3.6.1 Presentasjon av syntese.....	38
3.6.2 Diskusjon av begrensning i syntesens robusthet.....	40
4.0 Resultater	41
4.1 Beskrivelse av inkluderte studier	41
4.1.1 Akutte intervensjoner	41
4.1.2 Kroniske intervensjoner med varighet 1 uke – 3 mnd.....	43
4.1.3 Kroniske intervensjoner med varighet 3 – 6 mnd.	44
4.1.4 Kroniske intervensjoner med varighet 6 – 12 mnd.	46
4.1.5 Kroniske intervensjoner med varighet 12 – 24 mnd.	47
4.2 Koding av inkluderte studier.....	48
4.3 Metodologisk kvalitet.....	51
4.3.1 Metodologisk kvalitet for inkluderte RCT.....	51
4.3.2 Metodologisk kvalitet for inkluderte CRT.....	53
4.4 Syntese av inkluderte studier – et narrativ	56
4.4.1 Akutte fysiske aktivitetsintervensjoner	56
4.4.2 Kroniske fysiske aktivitetsintervensjoner	57
4.4.3 Utvikling av en teoretisk modell om hvordan intervensjonen fungerer	59
4.4.4 Identifisering av mønstre i syntesen	60
4.4.5 Heterogenitet i synteseplanet.....	61
4.4.6 Syntesens robusthet.....	62
4.4.7 En «sann» effektstørrelse.....	65
4.4.8 Undersøkelse av publikasjonsavvik	66
5.0 Diskusjon.....	67
5.1 Vurdering av validitet	68

5.1.1 Eksternvaliditet	68
5.1.2 Beregning av utvalgsstørrelse.....	70
5.1.3 Internvaliditet	72
5.2 Progresjon i kognitiv funksjon og akademisk prestasjon.....	76
5.2.1 Akutte kvantitative og kvalitative fysisk aktivitetsintervensjoner.....	77
5.2.2 Kroniske kvantitative og kvalitative fysisk aktivitetsintervensjoner.....	77
5.3 Anbefalinger for kroppsøving	79
5.4 Begrensinger i syntesens robusthet.....	80
5.5 Sensitivitetsanalyse.....	82
5.5.1 Hva om studier ikke fredsstiller eligeringskriterier?.....	82
5.5.2 Hva om studier med lav metodologisk kvalitet ekskluderes?.....	83
5.5.3 Hva om studier er heterogene?	83
5.5.4 Hva om ett resultat er veldig forskjellig fra resten?.....	84
5.5.5 Hva om resultater mangler?.....	84
5.5.6 Hva om ett studie med spesiell innflytelse på resultatene ekskluderes?.....	84
5.6 Identifisering av mangler i søk og utvelgelse av studier	85
5.6.1 Har alle berettigede studier blitt identifisert?	85
5.6.2 Hvordan blir forskningsspørsmålet adressert i primærstudiene og blir spørsmålet besvart i konklusjonen?	85
5.6.3 Ble primærstudier feilaktig akseptert eller ekskludert?.....	85
5.7 Identifisering av mangler i primærstudienes resultater	86
5.7.1 Har primærstudier unnlatt å rapportere sentrale resultater?	86
5.7.2 Mangler reviewet måleresultater?	86
5.7.3 I hvilken grad rapporterer primærstudier mediert og endelig utfall?.....	86
5.7.4 Hva er status for antallet tillagte funn rapportert i primærstudier?	87
5.7.5 Hvilke utfall brukes til å beregne utvalgsstørrelse i primærresultater?	87
5.7.6 Er det samsvar i primærstudienes konferansebidrag og siste publiserte resultater?	88
5.7.7 Er syntesen analysert?	89
5.7.8 Er funn og konklusjon presentert med en klar sammenheng?	89

5.8 Identifisering av mangler som kan relateres til publikasjonsavvik	89
5.8.1 Er alle rapporterte resultater signifikante?	89
5.8.2 Kan Harvestplot indikere publikasjonsavvik?	90
5.8.3 Har publikasjoner blitt tilført ved kontakt med forfattere?	90
5.8.4 Andre bidrag til publikasjonsavvik	90
5.9 Styrker og begrensninger.....	91
6.0 Konklusjon.....	92
6.1 Hovedhypotese 1.....	93
6.2 Hovedhypotese 2.....	93
6.3 Hovedhypotese 3.....	94
6.4 Implikasjoner for fremtidige arbeider	94
7.0 Referanser	96
8.0 Vedlegg	106
8.1 Vedlegg 1: Glossarium – begreper, forkortelser og administrering av testinstrumenter	106
8.1.1 Begreper	106
8.1.2 Forkortelser	109
8.1.3 Administrering av testinstrumenter	111
8.2 Vedlegg 2: Tverrfaglig databasesøk.....	121
8.3 Vedlegg 3: Ekskluderte artikler med begrunnelse	122
8.4 Vedlegg 4: Prosjektbeskrivelse for Masteroppgave i Idrett	126

1.0 Innledning

I læreplanen for kroppsøving fastsatt som forskrift av Kunnskapsdepartementet 11.05.2015 (Utdanningsdirektoratet, 2015, s. 2), heter det blant annet at «Kroppsøvfaget skal medvirke til at mennesker sanser, opplever, lærer og skaper med kroppen.» Dette er eneste sitat i formålet for kroppsøving som antyder at fysisk aktivitet kan ha betydning for kognisjon og akademisk prestasjon. Utvikling av nye undervisningsmetoder for å stimulere barn og unges akademiske prestasjon, er sentralt innen skoleforskning. Denne forskningen vil kunne lede til at også kroppsøvfaget sees med nye øyne når det gjelder elevenes kognitive funksjon og faglig læringsutbytte. Forskning på fysisk aktivitet dreier seg ofte om å undersøke fysiologiske parametere som leder til anbefalinger for en formålstjeneslig livsstil eller effektive treningsrutiner. Forskning på sammenhengen mellom fysisk aktivitet og kognisjon eller akademisk prestasjon går et steg videre og undersøker muligheten for å stimulere elever kognitivt og akademisk gjennom fysisk aktivitet. Denne rapporten retter søkelyset på det vitenskapelige grunnlaget for betydningen av fysisk aktivitet for elevenes kognitive funksjon og akademiske prestasjon i skolen. Tilnærmingen til problemstillingen er å gjøre en systematisk gjennomgang av empiriske studier fra de 10 siste år og studere betydningen av fysisk aktivitet for kognisjon eller akademisk prestasjon.

1.1 Bakgrunn for tema

Flere forskergrupper har publisert ulike litteraturstudier (reviews) om emnet. De utgjør konteksten for denne rapporten. Nyere forskning viser at økt fysisk aktivitet kan være viktig for elevers kognitive funksjon og læring i skolen. Man kan nærme seg forskningsfeltet på en rekke måter. Fysisk-motorisk ferdighet har betydning for kognitiv funksjon og skoleprestasjoner (Ommundsen, 2013). Nyere perspektiver innen nevro-psykologisk forskning (Hillman, Kamijo & Scudder, 2011) gir anatomiske fremstillinger av fysisk aktivitet. Empiriske studier (Tomprowski, Davis, Miller & Naglieri, 2008) indikerer at systematiske treningsprogrammer fremmer mentale prosesser som har betydning for akademisk prestasjon.

En litteraturstudie av Best (2010) undersøker hvorfor akutte (enkelt) økter og kroniske (repeterte) økter har positiv effekt på eksekutiv funksjon. Forskningsdesignet er konstruert ved enten umiddelbart å måle kognitiv effekt av en aerob økt eller måle differansen av kognitiv effekt før og etter noen ukers aerob trening. Resultatene pekte i retning av en effekt

på eksekutiv funksjonen etter enkeltøkter. Ved aerob trening over tid finner Best (2010) en varig forbedring av eksekutiv funksjonen.

En litteraturstudie av Tomporowski et al. (2008) undersøker hvordan fysisk aktivitet har en effekt på barns intelligens, kognisjon og akademisk prestasjon. Teorigrunnlaget i studien er kognitiv læringsteori som retter seg mot fysisk aktivitet. Problemstillingen belyser effekten av trening på kognitiv funksjon hos voksne. Studier gjennomført på dyr gir resultater med sammenheng til nevrologisk bildedannelse. Tomporowski et al. (2008) tolker resultatene som evidens på en kausal kobling mellom fitnessnivå og vitalitet. Hypotesen om påvirkning av kognitiv funksjon har blitt testet på eldre utøvere ved å undersøke effekten av aerob trening både på eksekutiv- og ikke-eksekutiv kognitiv funksjon. En rekke kognitive tester avdekker en differanse for utøvere før og etter treningsintervensjon (Tomporowski et al., 2008). Singh et al. (2012) er spesielt opptatt av longitudinelle studier og effekten på akademisk prestasjon. Fysisk aktive barn har bedret kognitiv atferdskontroll og øker sin konsentrasjonsevne og oppmerksomhet i innlærings situasjoner (Singh, Uijtdewilligen, Twisk, Van Mechelen & Chinapaw, 2012).

Vazou, Pesce, Lakes & Smiley-Oyen (2016) bruker randomiserte kontrollerte eksperimenter som undersøker sammenhenger mellom fysisk aktivitet og kognisjon med et ønske om å kunne forklare årsak og virkning. Vazou et al. (2016) er i sin meta-analyse opptatt av å tallfeste samlet effektstørrelse for kroniske intervensjoner sammenlignet med ulike kontrolltilstander. Vazou et al. (2016) predikerer en sammenheng og mangler en kausal forklaring til måleresultatene. Keeley & Fox (2009) etterlyser design med kvalitet og dybde i evidens ut over korrelasjonsstudier med svake sammenhenger. Keeley & Fox (2009) etterlyser også reviews som på en systematisk måte tilfører forskningsfeltet robust og objektiv kunnskap. Donnely et al. (2016) fremmer avansert teknologi (Magnetic resonance imaging MRI og Elektroencefalogram EEG) som danner anatomiske og biologiske modeller for å etablere en biologisk basis for observert effekt på kognisjon og akademisk prestasjon. Fedewa og Ahn (2011) etterlyser grundige undersøkelser slik at anbefalinger om ikke å redusere mengden kroppsøving i skolen kan formidles beslutningstagere. Fedewa & Ahn (2011) konkluderer med gevinst av fysisk aktivitet for barn og unges kognitive funksjoner og skoleprestasjoner (særlig lesing og matematikk). Singh et al. (2012) fremhever studienes metodologiske kvalitet som avgjørende for rapportering av fakta omkring problemstillingen.

Rasjonale for valg av typen review kom som et resultat av masterprosessen. I utgangspunktet var det lagt til rette for å skrive et metodologisk review for å belyse fagfeltets metodologiske problemstillinger. Dernest ble resultatene av datainnsamlingen avgjørende. Inkluderte primærstudier er kvantitative, men inneholder ikke tilstrekkelig eller entydig mengde kvantitativ informasjon som kreves i en meta-analyse. Valget falt på et systematisk review med en datakategorisering, en dataekstrahering, en kvalitetsvurdering og en narrativ syntese med tabellarisk presentasjon (Booth, Sutton & Papaioannou, 2016).

1.2 Formulering og drøfting av problemstilling

Med et metodologisk bakteppe som reflekterte et fagfelt dels med tvetydighet og liten grad av substans i måleresultater, ble følgende problemstilling formulert:

Kan fysisk, aerob, motorisk aktivitet, motorisk ferdighet og tradisjonell kroppsøving være årsak til kognitiv psykologisk respons som eksekutiv funksjon og akademisk prestasjon?

Fysiske aktivitetsintervensjoner har søkt svar på en rekke spørsmål ved å variere typen av stimuli ved å undersøke ulike utfall. Søken etter sterk evidens har gjort at inkluderte forskningsdesign er randomiserte eksperimenter (Randomized controlled trials, RCT design) eller kluster randomiserte eksperimenter (Cluster randomized controlled trials, CRT). Men, fysiske aktivitetsintervensjoner skiller seg fra kliniske eksperimenter på en rekke områder. For eksempel vil ikke fysiske aktivitetsintervensjoner kunne teste medisin A, B og C for sammenligning med placebo. De metodologiske utfordringene avdekkes ikke minst når det gjelder blinding og smitte. Eksperimentene har foregått på barn og unge som naturlig har et fellesskap i en klasse og som ikke fullt ut er uavhengige individer. Inkluderte studier er gjennomført i skolen. Systematiske søk danner grunnlag for valg av inkluderte studier. De inkluderte utvalg har deltagere i normal utvikling mellom 7 og 17 år. Studiene griper inn i en skolehverdag med en erkjennelse av at elever stimuleres, akademisk og fysisk hele tiden. Populasjonen er dannet av barn og ungdom hovedsakelig fra vestlige samfunn som Europa, USA, Canada og Australia. Noen unntak er: en studie (Chen, Yan, Yin, Pan & Chang, 2014) fra Kina og to studier (Telles, Singh, Bhardwaj, Kumar & Balkrishna, 2013; og Subramanian, Sharma, Arunachalam, Radhakrishnan & Ramamurthy, 2015) fra India.

Problemstillingen har sitt rasjonale som i arbeidene til Keeley & Fox (2009), Fedewa & Ahn (2011) og Singh et al. (2012). Problemstillingen blir rendyrket gjennom tre hovedhypoteser:

1. Fysiske aktivitetsintervensjoner vil styrke kognitiv funksjon (eks. eksekutiv funksjon) eller akademisk prestasjon sammenlignet med ordinær undervisning eller kroppsøving.
2. Primærstudiene har tilstrekkelig metodologisk kvalitet. Tilstrekkelig metodologisk kvalitet krever at det er 60% sannsynlighet for at primærstudiene er fri for metodologiske avvik.
3. Det systematiske reviewet oppnår en robusthet i evidens som gir implikasjoner for anbefaling av bestemte fysiske aktivitetsintervensjoner.

Rapporten besvarer problemstillingen ved å besvare disse hovedhypoteser.

1.3 Disposisjon for oppgaven

Studien legitimeres spesielt gjennom arbeidsoppgavene som leder til masteroppgaven.

Arbeidsoppgavene er:

1. Å vurdere effekten av fysisk aktivitetsintervensjoner på eksekutiv funksjon, kognisjon eller akademisk prestasjon, sammenlignet med akademisk undervisning eller ordinær kroppsøving.
2. Å vurdere den metodologiske kvaliteten som fremgår i hver primærstudie.
3. Å vurdere robusthet ved evidens for klinisk relevante funn ved å sammenligne fysiske aktivitetsintervensjoner og kontrollaktivitet, metodologisk kvalitet samt å vurdere om flere studier peker i en kausal retning.

Disposisjonen for oppgaven inkluderer de arbeidsoppgaver (1 - 3 over) som den systematiske gjennomgangen av den aktuelle forskningslitteraturen byr på. I forkant av dette presenteres teori som en ramme for arbeidsoppgavene. Teorikapittelet vektlegger å kunne synliggjøre målinger av kognisjon og akademisk prestasjon som noe mer enn «black box» instrumentering. Hypoteser introduseres for å forklare betydningen av fysisk aktivitet for utvikling av mentale prosesser. Antatte forklaringsmekanismer kan gi oss et innblikk i hvorfor fysisk aktivitet eventuelt påvirker kognitiv funksjon og akademisk prestasjon. Eksekutive funksjoner beskrives som forankret i et hierarki av kognitive funksjoner. Eksempler på psykomotoriske tester brukes for å konkretisere metoder for operasjonalisering av eksekutive funksjoner, kognitive funksjoner og metakognisjon. På bakgrunn av litteraturen kategoriseres fysiske aktivitetsintervensjoner i aktivitetsprogrammer som inneholder henholdsvis kvantitative økter og kvalitative økter, samt økter med sikte på avklaring av akutte eller kroniske effekter av fysisk aktivitet. Kroppsøvfaget i skolen utgjør en viktig potensiell arena for styrking av kognitiv funksjon og akademisk prestasjon, og flere fysisk aktivitetsintervensjoner har vært gjennomført der kroppsøvfaget har vært en del av aktivitets-

opplegget. Teorikapittelet avsluttes med en oppsummering av forskningsresultater for mulig forklaring på betydningen av ulike former for fysisk aktivitet på mentale prosesser.

Metodekapittelet vektlegger å gjøre studien transparent og reproduserbar ved at prosess og progresjon synliggjøres. Utgangspunktet for metoden var å identifisere relevante studier. Primærstudier ble generert ved systematiske databasesøk. Primærstudier fra de siste 10 år gir en oppdatering av kunnskap om effekten av fysisk aktivitet på kognisjon og akademisk prestasjon. Bakgrunnen for valg av typen review, et systematisk review, blir forklart. Variablene som brukes for fysisk aktivitet og kognisjon og akademisk prestasjon, blir konkretisert og beskrevet. Flytdiagram gir en oversikt med identifikasjon, screening og valg av inkluderte og ekskluderte artikler. En del av metodekapittelet består i å beskrive hvordan studier blir kategorisert og kodet, og hvordan metodologisk kvalitet vurderes og anvendes. Metodologiske kvalitet og gradert evidens blir brukt som mål på studiepreferanse eller hvilke måleresultater som har størst troverdighet. Resultatene presenteres som et narrativ sidestilt med en tabellarisk fremstilling i Harvestplott. Metodekapittelet inneholder en beskrivelse av hvordan Harvestplottet brukes i denne rapporten, og hvordan «best evidens» avgjør konkurrerende hypoteser.

Resultatkapittelet vektlegger å få klarhet i innledende hovedhypoteser som operasjonaliserer forskningsspørsmålet. Resultatene fungerer i to plan: et grunnplan bestående av primærstudier og et synteseplan som er en kompilering av innsamlede data. Inkluderte studier blir beskrevet og presentert. Studiene blir kodet og vurdert for metodologisk kvalitet. Data transformeres til synteseplanen og beskrives narrativt. Ut av syntesen utvikles en empiri-tro modell. Syntesen undersøkes for mønstre og ulike karakteristika. Metodologisk kvalitet brukes i vurderingen av syntesens robusthet. Store studier presenteres som indikasjon på generelle resultater og små studier presenteres for å sannsynliggjøre publikasjonsavvik.

Diskusjonskapittelet utforsker funn og mangler ved undersøkelsene som er presentert i resultatdel. Diskusjonskapittelet følger malen til Booth et al. (2016) som formell tilnærming til det systematiske litteraturreviewet. Validitetsvurderingen tar for seg ekstern validitet, utvalgsstørrelse og internvaliditet for inkluderte studier. Funn som har gitt progresjon i kognitiv funksjon og akademisk prestasjon blir presentert. Anbefalinger om bestemte typer fysisk aktivitet fremmes på bakgrunn av robusthet i evidens. Analysen har som mål å vurdere styrken i evidens for å kunne trekke slutninger om resultatene. Hensikten er videre å kunne

generalisere resultatene til en målpopulasjon eller til en valgt kontekst. Holdbarhet i funn for konkurrerende hypoteser avgjøres ved prinsippet om «best evidens». Prinsippet om «best evidens» kombinerer teknikker som tar hensyn til robusthet ved syntesen og grunnlaget for analysen (Slavin, 1995). En sensitivitetsanalyse utforsker konsekvensen av manglende, uklare eller åpenbart uriktige, metoder eller data i inkluderte studier (Booth et al., 2016).

Sensitivitetsanalysen kan danne et utgangspunkt for hvordan enkelte primærstudier gir implikasjoner for dette reviewet. Gjennom en rekke spørsmål identifiseres manglende rapportering og avvik i søk og utvelgelse av studier, mangler ved måleresultater i primærstudier og publikasjonsavvik for reviewet. Kapittelet avsluttes med styrker og begrensninger.

Konklusjonskapittelet presenterer essensen i masteroppgaven. Innledende hovedhypoteser blir besvart. Begrensinger i inkluderte primærstudier og dette reviewet gir implikasjoner for fremtidige arbeider.

Masteroppgaven avsluttes med referanser i APA stil (American Psychological Association, 2010). Vedlegg 1 inneholder glossarium for begreper, forkortelser og administrering av testinstrumenter aktuelle for masteroppgaven. Vedlegg 2 og 3 beskriver et fullstendig databasesøk og ekskluderte fulltekstvurderte artikler. Vedlegg 4 inneholder prosjektbeskrivelsen for masteroppgaven. Inkludert er de etiske overveielser som ligger til grunn for oppgaven.

2.0 Teoridel

Teoridelen setter fokus på antatte mekanismer som kan forklare betydningen av fysisk aktivitet for utvikling av mentale prosesser. Slike mentale prosesser kan gi oss et innblikk i hvorfor fysisk aktivitet eventuelt påvirker kognitiv funksjon og akademisk prestasjon. I den forbindelse introduseres metakognisjon som en aktuell medierende variabel for å forklare betydningen av fysisk aktivitet for kognitive funksjoner, samt betydningen av kognitive funksjoner for akademisk prestasjon. I forbindelse med læringsprosessen berøres også elevenes kognitive utvikling og læreridentitet. Dette omtales innen rammen av kognitiv læringsteori. Med basis i kognitiv psykologi blir eksekutive funksjoner beskrevet som forankret i et hierarki av kognitive funksjoner, og eksekutive funksjoner presenteres som et dynamisk begrep som utvikles med alderen. Deretter gjøres det rede for et utvalg tester for operasjonalisering av eksekutive funksjoner, kognitive funksjoner og metakognisjon. Flere av

disse er anvendt i intervensjoner med bruk av fysisk aktivitet med mål om å påvirke kognitiv funksjon og/eller akademisk prestasjon. Med basis i litteraturen kategoriseres slike fysiske aktivitetsintervensjoner i tiltak med aktivitetsopplegg som inneholder henholdsvis kvantitative, kvalitative økter samt, økter med sikte på avklaring av akutte og kroniske og kroniske effekter av fysisk aktivitet. Kroppsøvningsfaget i skolen utgjør en viktig potensiell arena for styrking av kognitiv funksjon og akademisk prestasjon, og flere fysisk aktivitetsintervensjoner har vært gjennomført der kroppsøvningsfaget har vært en del av aktivitetsopplegget. Derfor ser jeg nærmere på de muligheter faget byr på med sikte på stimulering av kognitiv funksjon og akademisk prestasjon i lys av fagets plass og funksjon i skolen. Kapitlet avsluttes med en oppsummering av forskningsresultater basert på tre ulike hypoteser fremlagt i litteraturen til mulig forklaring på betydningen av ulike former fysisk aktivitet på mentale prosesser som underlag for å forstå forholdet mellom fysisk aktivitet, kognisjon og akademisk prestasjon. I gjennomgangen av mentale prosesser står, en mulig effekt av fysisk aktivitet på ikke-eksekutive, kjerne-eksekutive og metakognitive funksjoner sentralt.

2.1 Antatte forklaringsmekanismer mellom fysisk aktivitet og utvikling av mentale prosesser

Forskningen på utvikling av kognisjon og mentale prosesser som følge av fysisk aktivitet har resultert i en rekke av hypoteser. Disse hypotesene har dannet utgangspunkt for planlegging og gjennomføring av ulike forskningsdesign, og er aktuelle som forklaringsunderlag i studier som har funnet evidens for betydningen av fysisk aktivitet for kognitiv funksjon og akademisk prestasjon. Lubans et al. (2016) framsetter tre forklaringsmekanismer basert på tre hypoteser: En nevrobiologisk hypotese, en psykososial hypotese og en atferdshypotese.

2.1.1 Den nevrobiologiske hypotesen

Ifølge den nevrobiologiske hypotesen (Lubans et al., 2016) styrker deltagelse i fysisk aktivitet kognisjon og mental helse gjennom strukturelle endringer i hjernen. Hypotesene underbygges av arbeidene til Voss, Vivar, Kramer & Van Praag (2013), og peker på tre kategorier nevrobiologiske forklaringsmekanismer som årsak til utvikling av kognitiv funksjon, som innebærer endring i nervesystemet:

1) CNS: Celler, molekyler og nerveledning – nevrogenese, relevante reseptorer (BDNF) er målt i dyreforsøk.

2) CNS: Biomarkører - volumet av grå substans og cerebralt blodvolum, aktivitetsmønstre (fMRI) og elektrofysiologi (EEG/eRP) er målt hos mennesker ved kognitive oppgaver.

3) Perifere biomarkører - vekstfaktorer (BDNF, IGF-1, VEGF), betennelsesmarkører, metabolsk funksjon, arteriell «compliance» er målt perifert hos mennesker.

Nevrologisk scanning eller bildedannelse (MRI, funksjonell MRI, Hendelses-relatert neurologisk potensial, «Event-related brain potential» (eRP)) har blitt anvendt for å identifisere strukturelle og funksjonelle mekanismer som kan forklare sammenhengen mellom fysisk aktivitet, kardiorespiratorisk kapasitet og kognisjon. Bildedannelse gir ikke en direkte måling av endring i mekanisme (Van Praag, 2008), men dyreforsøk har identifisert en rekke mekanismer blant annet endring i nevrologisk faktor BDNF, nervecelleproliferasjon (deling), -overlevelse og -differensiering i hippocampus. Nevrokjemikalier (BDNF, IGF-1, VEGF) viser økt produksjon ved fysisk aktivitet og kan påvirke hjernens struktur, funksjon og kognisjon. Konsentrasjonen av disse nevrokemikalierne er stor i hippocampus og andre deler av hjernen. Fysisk aktivitet antas å stimulere vekst av nye kapillærer (angiogenese) som i sin tur er avgjørende for transport av næringsstoffer til nevroner. Angiogenese er ikke nødvendigvis prediktor for økt nevrogenese (dannelsen av nye nerveceller) eller visa versa (Van Praag, 2008). Det har for øvrig blitt påvist en sammenheng mellom kardiorespiratorisk kapasitet og hjernebarkens struktur (frontallappen, anterior cingulum) og subkortikalt (hippocampus, basal gangliene) (Chaddock et al., 2010).

Økt psykologisk velvære som følge av fysisk aktivitet, som i sin tur kan styrke kognitiv funksjon og akademisk prestasjon ved å gi økt psykologisk overskudd og økt konsentrasjon (jmf. 2.1.2 Den psykososiale hypotesen (Lubans et al., 2016)) kan også forklares nevrobiologisk ved at økt fysisk aktivitet utløser endogene opioider (eks. endorfiner) og dermed også andre nevrotransmittere de er forbundet med. Endorfiner (neuropeptider) som kan dempe smerte og skape eufori, frigjøres trolig ved fysisk aktivitet. Men empiriske forsøk har ikke påvist at velvære etter fysisk aktivitet skyldes endorfiner eller at den kortsiktige effekten av fysisk aktivitet bedrer mental helse over tid. Godfølelsen etter trening kan også skyldes økte konsentrasjon av monoaminer i hjernen (transmittermolekyler - dopamin, noradrenalin, serotonin) (Lubans et al., 2016).

2.1.2 Den psykososiale hypotesen

Den psykososiale hypotesen legger til grunn at deltagelse i fysisk aktivitet styrker velvære gjennom en rekke psykososiale mekanismer. Ryff & Keyes (1995) gir en strukturert

beskrivelse av komponentene i begrepet psykologisk velvære. Ved å tilfredsstille grunnleggende psykologiske behov som sosiale relasjoner, autonomi, selvksept, mestring av omgivelsen, personlig vekst og mening med livet styrkes psykologisk velvære.

Fysisk aktivitet gir mulighet for sosial interaksjon (positive relasjoner), mestring av det fysiske domenet (mestringstro og persepsjon av kompetanse), styrker fysiske aspekter ved selvbilde i form av endret fysisk selvpersepsjon, kroppsbilde, og opplevd selvstendighet og tiltro til egen handlekraft (autonomi). I tillegg kan fysisk aktivitet gi bedret sosial relasjon til omgivelsen og potensielt bedre humør. Humør kan påvirke en mer generell affektiv tilstand og andre indikatorer på velvære. Den psykososiale hypotesen innebærer at fysisk aktivitet kan lede til oppgavespesifikk mestringstro (eks. «exercise efficacy»). Slik mestringstro kan gi overføringer til generell fysisk selvverd («physical selfworth») og global selvoppfatning («self-esteem») (Sonstroem, Harlow & Josephs, 1994). Under visse forutsetninger kan fysisk aktivitet blant barn og ungdom også ha en negativ effekt på indikatorer for mental helse. Eksempel: konsekvensen av dårlig instruksjon og svakt sammensatte treningsøkter kan være redusert behovstilfredsstillelse og lede til redusert persepsjon av kompetanse og global selvoppfatning. I slike tilfeller kan man forvente eventuelle negative eller manglende effekter av fysisk aktivitet på kognitive prosesser og akademisk prestasjon.

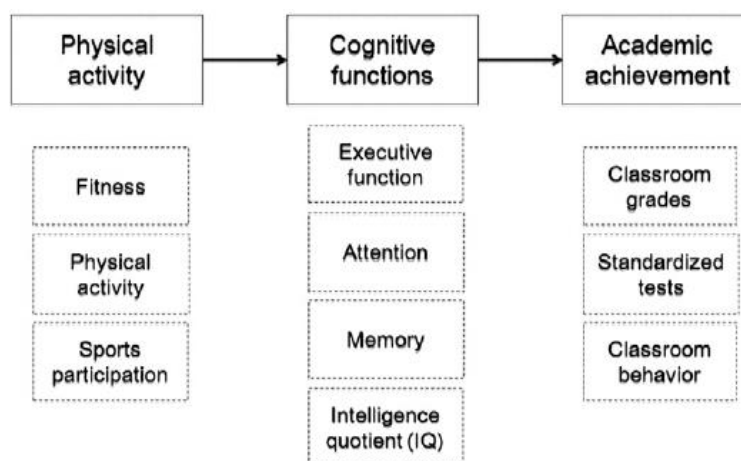
Innen domenet fysisk selvpersepsjon («physical self-perception») kan deltagelse i fysisk aktivitet forsterke persepsjonen av fysisk kompetanse, og styrket fysisk fremtoning (Kipp & Weiss i Lubans et al., 2016). Eksperimentelle studier har vist at fysisk aktivitet har en positiv effekt på selvrapportert velvære umiddelbart etter trening ute i naturlige omgivelser, men ikke etter innendørstrening (Coon et al., 2011). Forklaringsmekanismen er basert på at mennesket har en biologisk predisposisjon for å være tiltrukket til naturen. Mennesket har tilbrakt mesteparten av den evolusjonære tiden i naturen. Studier av voksne har vist at tilknytning til naturen har en positiv sammenheng med indikatorer for mental helse. I tillegg har naturen en helbredende effekt som også kan forklare at fysisk aktivitet i naturen har en positiv effekt på mental helse (Coon et al., 2011). Dersom hypotesen har støtte, vil det kunne ha betydning for hvilken fysisk kontekst man tilrettelegger for fysisk aktivitet med bedret kognitiv funksjon og akademisk prestasjon. Det er et tiltagende forskningsfokus på betydningen av fysisk aktivitet i naturlige utemiljøer for barn som tenkes å kunne styrke kognitive læringsprosesser via bedret motorikk, initiativ, kreativitet og konsentrasjon.

2.1.3 Atferdshypotesen

Atferdshypotesen peker på at fysisk aktivitet kan ha en effekt på indikatorer for mental helse mediert av relevant og assosiert atferd (Lubans et al., 2016). Fysisk aktivitet kan påvirke *ferdighetsmestring og selvregulering* som har implikasjoner for mental helse. Deltagelse i fysisk aktivitet kan bedre søvntid, søvnkvalitet, tidsbruk før innsovning og redusere trøtthet på dagtid (Astill, Van der Heijden, Van Ijzendoorn & Van Someren, 2012). Studier har vist at søvnmangel er negativt assosiert med høyere ordens komplekse kognitive funksjoner og økt atferdsproblematikk hos barn (Astill et al., 2012). Studier av Yoga (Khalsa, Hickey-Schultz, Cohen, Steiner & Cope, 2012) har vist en positiv effekt på ferdighetsmestring og kan også anvendes for å behandle engstelse og angstrelaterte forstyrrelser hos unge og voksne. Utvikling av selvregulering og mestring kan forklare at yoga og Martial arts har en positiv effekt på mental helse (Khalsa et al., 2012), og at slik aktivitet dermed har potensiale til å styrke kognitiv funksjon og akademisk prestasjon.

2.2 Teoretiske modeller for betydningen av fysisk aktivitet på kognitive funksjoner og akademisk prestasjon

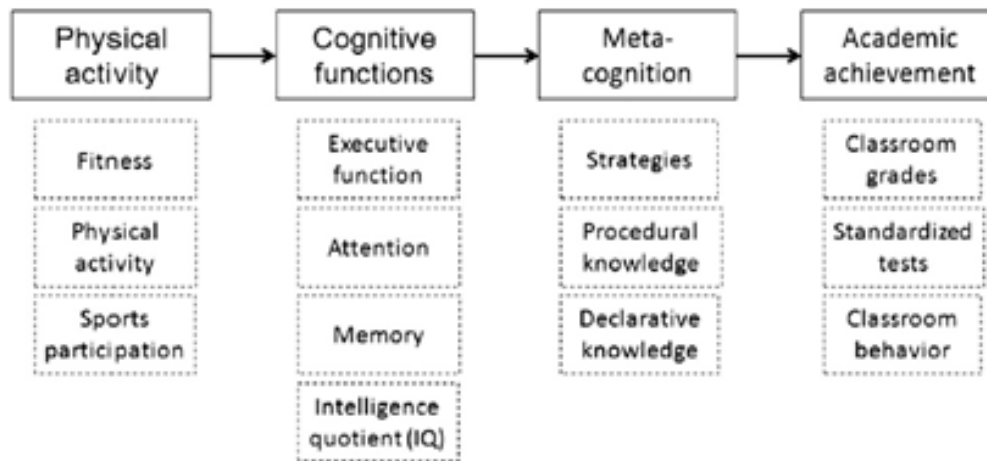
Howie & Pate (2012) foreslo en modell basert på hypotesen om at fysisk aktivitet endrer barns kognitive funksjoner – eksekutive funksjoner, oppmerksomhet, hukommelse og intelligens. En styrket kognitiv funksjon vil fungere som grunnlag for progresjon ved akademisk prestasjon – karakterer, testresultater fra standardiserte prøver og atferd i klasserommet (Figur 1).



Figur 1: Modell av sammenhengen mellom fysisk aktivitet og akademisk prestasjon (Howie and Pate, 2012).

2.2.1 Metakognisjon og eksekutive funksjoner som mediator mellom fysisk aktivitet og akademisk prestasjon

Figur 2 viser, en tilpasset modell (Tomporowski et al., 2015), der også metakognisjon ble inkludert som medierende faktor for akademisk prestasjon. Metakognisjon innebærer individets forståelse av hvordan kunnskap er organisert og av læringsstrategier, faktorer av betydning for å regulere egen atferd i innlærings situasjoner.



Figur 2: Modell av sammenhengen mellom fysisk aktivitet, kognitive funksjoner, metakognisjon og akademisk prestasjon (Tomporowski et al., 2015).

Eksekutive funksjoner og metakognisjon overlapper, men er likevel forskjellige. Vesentligst er kanskje prosessvarigheten. Testing av eksekutive funksjoner inkluderer testinstruksjon, at deltagerne responderer så raskt og presist som mulig. Responstiden omkring 2 sekunder, indikerer en «on-line» prosessering, og kapasitet til raske endringer etter situasjonsbestemte oppgaver. Metakognisjon måles ved at deltagerne må løse flertrinnsoppgaver. Deltagerne bruker strategier som krever grundig refleksjon og vurdering underveis mot et spesifisert mål. Responstiden er fra sekunder til flere minutter. Meta-kognisjon med fokus på mental og motorisk kreativitet måles ved problemløsningsstrategier uten tidskrav (Tomporowski et al., 2015). Modellen som både inneholder kognitive funksjoner og metakognitive funksjoner er en måte å illustrere hvordan ulike fysisk aktivitetsintervensjoner kan påvirke barns akademiske prestasjon (Tomporowski et al., 2015).

2.2.2 Kognitiv læringsteori kan introdusere en aldersmoderator

Effekten av fysisk aktivitet på kognisjon og akademisk prestasjon vil rimeligvis kunne være moderert av alder ved at elevene befinner seg på ulike stadier i sin kognitive utvikling. Piagets forskning (Birkeland, Venheim & Breiteig, 2011) har ledet frem til erkjennelsen av at

kognitiv utvikling skjer i trinn, der hvert trinn har sine karakteristiske trekk: I «den sansemotoriske perioden» (0.5 – 2.5 år) er sansinger og handlinger viktige for barnets læring. I «den før operasjonelle eller preoperasjonelle fasen» (4 – 9 år) er språk, imitasjon og tegning viktig for barnets evne til å representere tanker og ideer. Barnets konkrete opplevelse og intuitive tolkning gjør at barnet lærer begreper som størrelse, form og relasjoner. Barnet mangler generalitet og konservering av antall, lengde, avstand og volum. I «den konkret-operasjonelle fasen» (9.5 – 17.5 år) kan ungdommer på bakgrunn av erfaring med konkret materiell eller spesielle situasjoner lære og utføre abstrakte logiske resonnement. I «det formelt operasjonelle stadiet» (9.5 – voksen alder) kan ungdommer ved hypoteser lære resonnement. Hvis - så resonnement kan gjennomføres, og logiske operasjoner kan utføres også uten bakgrunn i erfaringer med konkret materiell (Birkeland et al., 2011).

2.2.3 Læreridentitet som mediator mellom fysisk aktivitet og kognisjon

Fysiske aktivitetsintervensjoner baseres på at elever blir veiledet av en instruktør enten ved akutte eller korniske treningsøkter. Instruktøren blir avgjørende for hvordan øktene gjennomføres og oppfattes av elevene. Læreridentitet som følger en bestemt undervisnings-tradisjon kan påvirke i hvilken grad de ulike kognitive systemene blir stimulert hos elevene (Birkeland et al., 2011). Skjemaet (Tabell 1) som er en spissformulering av en konstruktivistisk og en tradisjonell behavioristisk undervisning (Birkeland et al., 2011), viser ytterpunktene for kommunikasjonen i klasserommet, de kognitive strukturer som utvikles hos elevene, de oppgavene som arbeides med, og konsekvensen som følger av feil og misoppfatninger underveis.

Tabell 1: Sammenhengen mellom undervisningstradisjon og kognitive strukturer (Birkeland et al., 2011).

	<i>Konstruktivistisk syn på kunnskap</i>	<i>Tradisjonelt behavioristisk syn på kunnskap</i>
Kommunikasjonen	Spørrende undersøkende	Ordre Instruksjon
Kognitive strukturer	Refleksjon forståelse	Imitasjon Memorering
Oppgaver	Prosesspreget oppdagende	Produktet teller Resultatet teller
Status for feil og misoppfatninger	Stadier på veien mot å konstruere en kunnskap	Tolkes som mangler Nederlag Negativt

Ambisjonen i matematikk er eksempelvis å overføre evnen til å resonnerer og føre bevis for klar og logisk tenkning som senere kan benyttes i naturvitenskap eller jus. Men

undervisningstradisjoner kan også hemme en slik utvikling og produsere elever med en forsinket kognitiv utvikling, som baserer sine strategier på imitasjon og memorering heller enn refleksjon og forståelse (Birkeland et al., 2011). Forskningsdesign er følsomme for kontekst. Intervensjoner bør derfor utøves i skolehverdagen med daglige undervisnings- og læringsrutiner og innenfor et relevant pensum som gir en økologisk validitet (Fleischer & From, 2017, s. 117).

2.3 Kognitive funksjoner - begrepsavklaring

Ifølge Fleischer et al. (2017, s. 15) kan kognitive funksjoner beskrives som verktøy vi bruker for å oppfatte og forstå oss selv og andre. «Kognitiv» (Imsen, 2005) står for det som har med de intellektuelle funksjonene å gjøre (læring, hukommelse, tenking og problemløsning). Et hierarki av kognitive funksjoner inneholder også eksekutive funksjoner (Fleischer et al., 2017, s. 15):

- Eksekutive funksjoner
- Språk
- Læreevne og hukommelse
- Oppmerksomhet
- Persepsjon
- Årvåkenhet

Fra Gestaltpsykologien oppfatter mennesket helheter (Imsen, 2005). Helheten kommer i stand ved at en organiserer nerveimpulser i mønstre eller gestalter (etter gestaltlover). Alle delene av et hierarki henger sammen og påvirker hverandre (Fleischer et al., 2017, s. 16). De nederste funksjonene er en forutsetning for at de øverste skal fungere fullt ut. Eksekutiv funksjon er individets evne til å bruke verktøyene til problemløsning. Individet kan også ha enkelte kognitive funksjoner godt utviklet uten å beherske selvstendig problemløsning særlig godt. En må kanskje i større grad basere seg på rutinepreget arbeid. Eksempel: et barn kan aldersadekvat lese og stave, men likevel være ute av stand til å skrive en historie.

2.3.1 Eksekutive funksjoner

Ifølge Wasserman & Wasserman (2013) og Fleischer & From (2017, s. 19) kan eksekutive funksjoner defineres gjennom 4 – fire trinn som bærer preg av problemløsning, kompleksitet og til en viss grad at situasjonen er ny (ikke rutine eller automatiserte handlinger):

- Få en ide`.
- Planlegge utførelsen av ideen.

- Utføre planen gjennom målrettede handlinger hvor en tar hensyn til både egne behov og til situasjonens krav.
- Vurdere utførelsen underveis og til slutt foreta eventuelle justeringer.

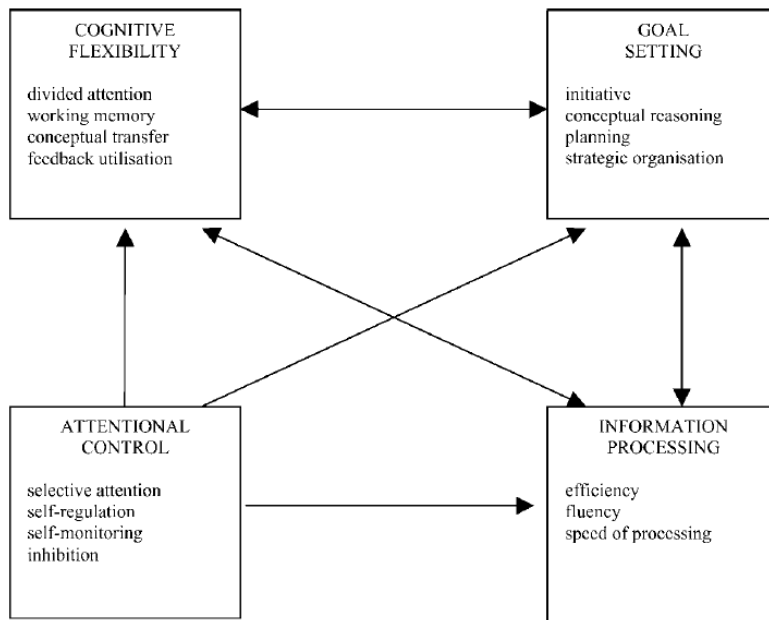
En ide` forutsetter at en kan forestille seg noe. Ideene bygger på tidligere erfaringer (hukommelse) og dels på den aktuelle situasjonen. Å planlegge innebærer at en kan fastholde og dele opp ideen sin, samt utføre de enkelte delene, i den rekkefølge en har valgt (krever arbeidsminne, sekvens-syntese-tenkning og sekvensforståelse). Utføring av planen handler om å utføre de forskjellige delhandlingene i riktig rekkefølge frem mot målet. Uvesentlige eller forstyrrende elementer skal fjernes, samtidig som en må være oppmerksom på andres krav og forventninger, slik at en unngår å overskride grenser eller etiske regler. Utførelsen krever avansert oppmerksomhetskontroll, et godt informasjonsforarbeid, selvbeherskelse og målrettethet. Vurdering og justering gjennom hele utførelsen krever evne til å vurdere opp mot noe annet, samt fleksibilitet og innlevelse.

Ifølge Anderson (2002) kan ikke eksekutive funksjoner deles uten at mening går tapt. Eksekutive funksjoner kan heller ikke betraktes som et overordnet kontrollsystem. Forskere er enige om at eksekutiv funksjonen består av 3 til 5 enheter som sammen utgjør konstruksjonen eksekutive funksjoner. Modellen i Figur 3 (Anderson, 2002) baseres på 4 overordnede eksekutive domener som sammen utgjør et eksekutivt kontrollsystem – de eksekutive funksjoner. Modellen kan betraktes som et teoretisk instrument for å få et overblikk. Områdene påvirker hverandre innbyrdes. Et domene vil aldri kunne arbeide helt alene eller uavhengig av andre kognitive ferdigheter. Modellen stemmer godt overens med nyere forskning. Det er ikke empirisk belegg for at det finnes en mengde forskjellige separate eksekutive funksjoner.

Modellens utgangspunkt er styrt oppmerksomhet – «attentional control» (eksekutiv og dirigerende). Styrt oppmerksomhet har sammenheng med oppmerksomhetsbegrepet. Oppmerksomhet kan forenklet teoretiseres med tre grunnleggende systemer: et opprettholdende system (sikrer årvåkenhet), et orienteringssystem og et eksekutivt system (Petersen & Posner, 2012). Orienteringssystemet sysler med å oppdage, velge og flytte oppmerksomhet, mens det fremre systemet er mer eksekutivt og dirigerende og holder i tillegg vedvarende fokus. Styrt oppmerksomhet vil si å holde fast ved valgt fokus, men krever også en evne til å hemme impulser, innfall og forstyrrelser utenifra. Selvregulering med evnen til å se seg selv utenfra er også viktig. Til sist må en kunne selv-monitorere. Det vil si å holde

øye med eller overvåke det en holder på med: at planen utføres i riktig rekkefølge, at feil korrigeres og at målene nås (Petersen et al., 2012).

ASSESSMENT AND DEVELOPMENT OF EF



Figur 3: Modell av de eksekutive ferdighetene (Anderson, 2002).

Styrt oppmerksomhet påvirker alle øvrige domener. Kognitiv fleksibilitet – «cognitive flexibility», som inneholder delt oppmerksomhet, arbeidsminne, begrepsoverføring og bruk av tilbakemelding. Forenklet handler kognitiv fleksibilitet om evnen til å kunne skifte mellom forskjellige reaksjonsmåter eller være oppmerksom på flere aspekter ved en oppgave eller situasjon på samme tid. Det handler også om å kunne lære av feilene sine og om å benytte seg av alternative strategier (skifte innstilling). Arbeidsminnet handler om evnene til å holde på eller bearbeide informasjon i et kortere tidsrom. Arbeidsminnet gjør en form for mental sjonglering som er avhengig av en fleksibel tanke og som kan mislykkes dersom arbeidsminnet er overbelastet. Arbeidsminne og fleksibilitet er forutsetning for å ta i bruk det en har lært for overføring til en annen sammenheng. Arbeidsminne og fleksibilitet er også forutsetning for å kunne benytte tilbakemelding fra omgivelsen og i tillegg kunne holde på informasjonen tilstrekkelig lenge til å kunne agere på den. I modellen (Anderson, 2002) er kognitiv fleksibilitet synonymt med ulike måter å vinkle informasjon på. Kognitiv fleksibilitet påvirker målrettethet og informasjonsbearbeiding (Anderson, 2002). I begrepet målrettethet – «goal setting» ligger evnen til å igangsette, planlegge, strukturere og legge strategier. Planlegging krever at en kan forutse hva som kommer til å skje og at en må kunne formulere mål ut ifra de forskjellige trinn en skal gjennom (vanskelig å skille fra strukturering og

organisering). Organisering kan forstås som evnen til å sette sammen forskjellige trinn i prosessen, til effektive handlingssekvenser. Planlegging krever også abstrakt tenkning og begrepsdannelse. Dersom en ikke har overblikk gjennom målrettet planlegging hele tiden, kan kognitiv fleksibilitet påvirkes ved at det blir vanskeligere å tenke alternativt eller å utnytte arbeidsminnet. Domenet målrettethet kan påvirke effektiviteten i informasjonsbearbeidelsen. Effektivitet er nøkkelbegrep i domenet informasjonsbearbeiding – «information processing». Effektivitet handler om hvor raskt en kan oppfatte. Ulike syn gjør at en del forskere peker på at effektivitet og hurtighet ikke er grunnleggende funksjoner, men egne komponenter i det eksekutive systemet. Effektivitetsdomenene har med kvantitet (hvor mye innen en viss tid) og kvalitet (hva) i forhold til det en oppfatter. Vanskeligheter med informasjonsbearbeiding, vil påvirke arbeidsminnet, og det tar lengre tid å bearbeide alle delene. Evnen til målrettethet påvirkes.

Det eksekutive systemet med 4 adskilte domener fungerer optimalt ved at systemene arbeider sammen på kryss og tvers. Modellen gir en forståelsesramme som kan bidra til å belyse atferd. Forståelsen av disse mønstrene har stor betydning for tilrettelegging for en passende innsats. Modellen kan også hjelpe til å tolke empiri og guide oss slik at vi stiller de rette spørsmålene (Anderson, 2002).

2.3.2 Eksekutive funksjoner som dynamisk begrep

En oppgave kan oppfattes både som problemløsning og rutinemessig handling. Dersom individet først forsøker noe nytt og komplekst, kan en gitt atferd dreie seg om problemløsning og derav utfordre de eksekutive funksjoner. Eksekutive funksjoner er et dynamisk begrep i den forstand at de blir påvirket av en persons følelsesmessige tilstand. Barn som føler seg redde eller presset kan ha problemer med å komme med nye ideer og problemløsninger.

Eksekutive funksjoner er avhengig av erfaring og læring, men også av hjernens modning. Spesielt i barne- og ungdomsårene er eksekutive funksjoner i sterk utvikling. Toåringen med liten indre struktur, er avhengig av en høy grad av ytre struktur (hvor, når, hvordan barnet skal spise, leke, sove osv.). Barn i 13 års alder har en indre struktur som krever mindre ytre struktur. Trettenåringen kan selv organisere samvær og aktiviteter, kun med den ytre struktur at han/hun må være hjemme senest kl. 21:00 (Fleischer et al., 2017, s. 18).

2.3.3 Eksekutive funksjoner i et utviklingsperspektiv

De eksekutive funksjonene er ennå ikke ferdig utviklet hos tenåringene. Fram til tenårene (8-12 år) snakker en om en utviklingsmessig «spurt» (Fleischer et al., 2017, s. 151). De målorienterte evnene foredles. Evnen til fleksibilitet bedres, og evnen til impulshemming, fokusert oppmerksomhet og arbeidsminnet tar store modningsskritt. De sosiale ferdighetene tar også store skritt i denne perioden. Billedspråk, overført betydning og små hvite løgner er blitt en del av det mellomstore barnets repertoarer. Informasjonsbearbeiding ser ut til også å ta store skritt (6 – 12 år). I tenårene skjer det fortsatt store fysiske forandringer i pannelappen og nettverkene omkring (Fleischer et al., 2017, s. 151). Forbindelsen mellom hjernecellene effektiviseres og minskes dermed i antall. Signalene mellom cellene blir bedre i takt med myeliniseringen av nervecellene. Oppmerksomhetskontrollen blir bedre sammenlignet med yngre barn. Prosesseringshastigheten stiger og arbeidsminnet samt evnen til å planlegge økes fortsatt. Følelsesmessige og sosiale ferdigheter foredles ytterligere. Informasjonsbearbeiding ser ut til å stabiliseres i 15 års alderen. Gjennom tyveårene skjer det positive endringer i hjernens fremste del (et eksekutivt høydepunkt), med marginale endring i informasjonsbearbeiding. Hjernen arbeider best og mest effektivt på alle plan. I tyveårene faller det eksekutive systemet på plass, men allerede i 30 års alderen begynner hjernens vekt å synke som et resultat av hjernecellenes stille forfall. Det sies at 90 åringen har en hjernestørrelse som er redusert med ca 10% (Fleischer et al., 2017, s. 152).

2.3.4 Operasjonalisering av eksekutive, ikke-eksekutive og metakognitive funksjoner

Meta – analysen til Álvarez-Bueno et al. (2017) kartlegger effekten av ulike fysisk aktivitetsintervensjoner på barns kognisjon og metakognisjon. Artikkelen skiller mellom «kjerne - eksekutive funksjoner» som arbeidsminne, kognitiv fleksibilitet, og inhiberingskontroll som inneholder interferenskontroll av kognitiv inhibering og selektiv oppmerksomhet (Diamond, 2013) og «ikke - eksekutive kognitive funksjoner».

Ikke - eksekutive kognitive funksjoner (Álvarez-Bueno et al., 2017) måles ved hjelp av intelligenstester (IGF-M = Spanish overall and Fractional Intelligence test), kognitiv kartlegging (CAS = Cognitive assessment system), test av risikoatferd (Bart = Balloon Analog Risk task) og testing av beslutningstaking (Decision making).

Kjerne - eksekutive tester (Álvarez-Bueno et al., 2017) måles innen tre områder: arbeidsminne, selektiv oppmerksomhet/inhibering og kognitiv fleksibilitet/skifte av setting.

Testene som er brukt for å måle arbeidsminne er:

Digit/visual Span backward, CANTAB Memory Span/Error, Sternberg Task, WISC Digit Span = Wechsler intelligence scale for children, Visual Span/Spatial Memory, Free/Cued recall, RNG = random number generation, Non spatial n-back task, Digit/Visual Span backward (Álvarez-Bueno et al., 2017).

Testene som er brukt for å måle selektiv oppmerksomhet/inhibering er:

RNG = random number generation, CAS attention, Golden Stroop test, D2-R test, Continue Performance test, Flanker task, Heart and Flowers og Cognitrone (Álvarez-Bueno et al., 2017).

Testene som er brukt for å måle kognitiv fleksibilitet/skifte av setting er:

M-WCST = modifisert versjon av Wisconsin Card Sorting Test, Color-shape switch task, Flanker task (Additional block) og Ruff Figural Fluency test (Álvarez-Bueno et al., 2017).

Historisk er metakognisjon (Ylvisaker & Feeney, 2008, s. 411) blitt brukt som begrep for eksekutiv selvregulering og kontroll over kognitive prosesser i sammenheng med læring og akademisk prestasjon. Metakognisjon har en komponent - kunnskap om læring og en komponent selv-regulering og ble delt i to hovedkategorier: «eksekutive funksjoner (høyere orden)» inkluderte planlegging, fluid intelligens- abstrakt resonnement og problemløsning og «bevissthet om læring» som inkluderte strategi i forhold til målsetning, problemløsning, positivitet (Álvarez-Bueno et al., 2017).

Testene som ble brukt for å måle eksekutive funksjoner (høyere orden) er:

Intelligens tester (IGF-M), kognitiv kartlegging (CAS planning), fluid intelligens (Standard Progressive Matrices), kreativitet (RFFT designs/rotations), og planlegging (Tower of London) (Álvarez-Bueno et al., 2017).

Testene som ble brukt for å måle bevissthet om læring er:

Positivitet og målrettethet (Self- beliefs test: positive thinking, goal setting), selvregulering (Response to change scale og self-efficacy questionnaire), målrettethet (Goal questionnaire) og mestringstro (Self-efficacy questionnaire) (Álvarez-Bueno et al., 2017).

2.3.5 Måter å forbedre eksekutive funksjoner

Ifølge Diamond og Ling (2016) har en rekke studier vist at ulike aktiviteter egner seg til trening av eksekutive funksjoner (trening med pc spill, aerobic, styrketrening, Martial arts, yoga, mindfulness (meditasjon), teater og en del skolepensum. På bakgrunn av disse studiene (Diamond et al., 2016) er følgende slutninger formulert i den hensikt å bedre eksekutive funksjoner:

1. Trening av eksekutive funksjoner (EF) gir en overføringsverdi, men overføringen virker å være snever. Eksempel: Pc spill som stimulerer arbeidsminne, styrker kun arbeidsminne ikke selvkontroll, kreativitet eller kognitiv fleksibilitet.
2. Progresjon i EF er avhengig av treningstiden. Tilstrekkelig mengde trening er generelt avgjørende for å bli god.
3. Progresjon i EF er avhengig av måten aktiviteten blir presentert og gjennomført på.
4. Progresjon krever at de eksekutive funksjonene utfordres kontinuerlig. En regelmessig stimulans av eksekutive funksjoner er ikke tilstrekkelig for progresjon.
5. Individuer med svake eksekutive funksjoner har best utbytte av EF trening.
6. Progresjon krever kontinuitet i treningen. Dersom treningen opphører forsvinner også forbedring i EF.
7. Ofte er differansen mellom behandling (fysisk aktivitet) og kontroll målbar bare hvis eksekutive ferdigheter utfordres maksimalt.
8. Aerob trening eller styrketrening uten en kognitiv komponent gir lite eller ingen bedring av EF.

2.4 Fysisk aktivitetsintervensjoner – tidligere forskning

Et utdrag av fysisk aktivitetsintervensjoner som allerede har blitt undersøkt vil i det følgende bli presentert. Både ordinær kroppsøving og ordinær undervisning er ofte brukt som kontrollaktivitet. Innholdet i ordinær kroppsøving blir utdypet.

2.4.1 Grunnleggende bevegelsesferdighet som intervensjon

En review av Lubans et al. (2010) beskriver grunnleggende bevegelsesferdigheter (FMS) som bidrag til barns utvikling fysisk, kognitivt og sosialt. Typisk utvikles grunnleggende bevegelse som foredles i en situasjonsavhengig- og idrettsspesifikk ferdighet. Her nevnes bevegelse på stedet (løping og hopping), manipulende objektkontroll (kast eller mottak) og

stabilitetsferdighet (balanse og vridning). Modning av FMS kompetanse kan lettere oppnås ved korrekt praksis, oppmuntring, feedback og instruksjon (Lubans et al., 2010).

2.4.2 Kvantitative og kvalitative, akutte og kroniske intervensjoner

I Tomporowskis narrative review (Tomporowski, McCullick, Pendleton & Pesce, 2015) ble studiene karakterisert etter enten kvantitative- eller kvalitative intervensjoner. De kvantitative intervensjonene hadde kjennetegn ved at de krevde minimal ferdighet. Løp på tredemølle, ergometer sykkel, eller rytmikk, inneholder repetert bevegelser med liten «topp-down kontroll». Kartlegging av intensitet i disse intervensjonene var hovedsakelig basert på kardiorespiratorisk kapasitet (hjerterefrekvens, oksygenopptak, akselerometri). De kvalitative intervensjonene ble karakterisert som trening med høy kognitiv innsats eller som innebar innlæring av ferdighet. Exergames (videospill med bevegelse), koordineringsspill, strategi/innlæringspill er aktiviteter hvor forsterkningen skjer langs aksen - mental mobilisering. Innsats ble målt ved observasjon eller egenrapportering. Mental mobilisering – «mental engagement», ble definert som refleksjon over egen atferd, inklusive anstrengelsen som kreves for å forstå ny informasjon og å mestre nye ferdigheter (Tomporowski et al., 2015).

Videre ble studiene delt etter akutt- eller kroniske intervensjoner. Kvantitative akutte intervensjoner var enkeltøkter basert på kvantitative mål, mens kroniske kvantitative intervensjoner var repeterte økter basert på kvantitative mål. Tilsvarende inndeling ble gjort for kvalitative akutte- og kroniske intervensjoner (Tomporowski et al., 2015).

2.4.3 Kroppslig kognisjon

Forskning på kroppslig læring («embodied learning») har sitt rasjonale ved at bevegelse er avgjørende for ens eksistens. Fra et evolusjonært perspektiv argumenteres det for at hjernen har utviklet seg for å kontrollere og organisere bevegelse og forklarer dermed en kognitiv utvikling (Lians i Tomporowski, McCullick & Horvat, 2010). Sentral for kroppslig kognisjon («embodied cognition») er at kognitive prosesser er rotfestet i kroppslig interaksjon med resten av verden (Wilson, 2002).

2.4.4 Leke aktiviteter

Forskning på kognitiv utvikling har vist at de fundamentale eksekutive prosessene har ulike utviklingskurver (Best, Miller & Jones, 2009). Forskning på fysisk aktivitet i form av leker og

ulike spill har vist at lek har en viktig rolle i normal modning og utvikling av barns kognitive prosesser (Tomporowski et al., 2010). I restriktive omgivelser hvor tilgangen til fri lek har vært begrenset, har en hindret utviklingen av et godt sosialt atferdsmønster (Tomporowski et al., 2010).

2.4.5 Ordinær kroppsøving som kontrollaktivitet

I mange studier fungerer ordinær kroppsøving som kontroll eller sammenligningsaktivitet. Kroppsøvingfaget er innholdsmessig i utvikling (Tomporowski et al., 2010). Hensikten med kroppsøving er at elevene får en fysisk utdanning og forblir fysisk aktive hele livet. Dagens kroppsøving tar sikte på at elevene skal tilegne seg en motorisk ferdighet, kunne anvende koordinative egenskaper (balanse, rytme, reaksjon, romorientering, tilpasset kraft og øye-hånd, øye-fot koordinasjon). Elevene skal anvende strategier, være fysisk aktive, bli sunne og friske, bli sosiale og personlig ansvarlige, og verdsette fysisk aktivitet av en rekke grunner (Tomporowski et al., 2010). I den senere tid har også overvekt- og fedmeproblematikk lagt grunnlag for innholdet i kroppsøvingstimene. Pensum har fokusert på moderat til høy intensitetstrening (MVPA), uten å vektlegge tilegnelse av spillferdigheter som krever mental involvering. En del kroppsøvingsprogrammer har dreid seg om enkle aerobe aktiviteter. Det har vært mindre fokus på fritid med samtale og lek som fremmer mental involvering. Den mest effektive måten å redusere prevalens for overvekt og fedme kan være MVPA i kombinasjon med mental mobilisering (Tomporowski et al., 2010). Glede og trivsel er hovedmotivator for å delta i fysisk aktivitet blant lavinntekts-, og kulturelt divergente ungdommer og voksne (Bragg, Tucker, Kaye & Desmond, 2009). Kvalifiserte instruktører kan stimulere til en egnet mental mobilisering i fysisk aktive spill med moderat til høy intensitet (Tomporowski et al., 2010).

2.5 Betydningen av ulike former fysisk aktivitet for kognitive funksjoner og akademisk prestasjon

Måleresultater med tolkning basert på den nevrobiologiske hypotesen blir beskrevet inndelt i kvantitative akutt- og kroniske økter og kvalitative akutt- og kroniske økter. Status for effekten av fysisk aktivitet på ikke-eksekutive, kjerne-eksekutive funksjoner og metakognisjon blir presentert.

2.5.2 Kvantitative akutte intervensjoner

Hoveddelen av studier som har undersøkt effekten av fysisk aktivitet på barns kognisjon har basert seg på hypotesen om at nevropsykologiske mekanismer gir en underliggende tolkning av empiri (Tomporowski et al., 2015). I kvantitative studier hvor treningsdosering undersøkes, kan en måle en direkte kapasitet til å endre kognitive funksjoner hos barn. Akutte kvantitative intervensjoner har i samsvar med modeller fra teori om informasjons-prosessering, vist seg å endre barns *oppmerksomhet, prosesseringshastighet og eksekutive kontroll* (Tomporowski et al., 2015). En enkel informasjonsprosesseringsmodell, er basert på en hypotese om at informasjon fra omgivelsen tilflyter sentralnervesystemet. Måten stimuli reguleres ved mental innsats er mediert av eksekutive funksjoner. Stimuli behandles i en persepsjonsenhet, en enhet for beslutningsprosesser og valg av respons og en enhet for responsprogrammering som resulterer i atferd (Tomporowski et al., 2010). Best (2010) hevder at akutte kvantitative økter bedrer *reaksjonstid*. Matlab-tester med multiple choice oppgaver har vist en bedret reaksjonstid målt etter akutte økter (Elleberg & St-Louis-Deschênes, 2010). Resultatet ble tolket som en effekt på eksekutive funksjoner – *kognitiv fleksibilitet og styrt oppmerksomhet, inhibering*. Studien styrker antagelsen om at akutte kvantitative økter påvirker deler av eksekutiv funksjonen (Best, 2010). Elleberg et al. (2010) hevder samtidig at enkle oppgaver som måler reaksjonstid involverer enkle sansemotoriske funksjoner i syns- og motorisk bark.

2.5.3 Kvantitative kroniske intervensjoner

Regelmessig trening kan gi treningstilpasning eller fysiologisk tilpasning på en rekke områder (Tomporowski et al., 2015). Resultater som viste at kronisk fysisk aktivitet kan endre kardiorespiratorisk kapasitet ledet til spekulasjonen om at effekten på kognitive funksjoner skyldes økt blodsirkulasjon og oksygentilgjengelighet til hjernen (Tomporowski et al., 2015). Sent på 1990 tallet kom forskere frem til at aerob trening påvirker hjernens strukturer og nettverk. I dag eksisterer sterk evidens for at aerob trening endrer hjernens strukturer og funksjoner (Hillman, 2008; Chaddock, Pontifex, Hillman & Kramer, 2011). Disse funnene ledet til flere hypoteser som har utspring i nevrologi. En har forsøkt å forklare hvordan kronisk trening endrer kognitive funksjoner og hvordan hjernen endres nevrofysiologisk. Relasjonen mellom kronisk aerob trening og kognisjon er undersøkt ved empiriske studier som har variert i intensitet, varighet og treningsfrekvens. Randomiserte eksperimenter har vist en fordel kronisk aerob fysisk aktivitet hos barn med overvekt (Davis et al., 2011). MR-målinger avdekket en tentativ evidens for økt aktivitet i bilateral prefrontal bark og redusert aktivitet i bilateral posterior parietal bark i testsituasjon bestående av eksekutive oppgaver

(Davis et al., 2011). Testene som kartla kognitive funksjoner og akademisk prestasjon etter intervensjon (dose-respons forsøk, 20 min/dag og 40 min/dag, ca 3 mnd), viste bedret *eksekutiv funksjon*. Sammenligningsgruppen besto av overvektige inaktive barn (Davis et al., 2011). Den fysiske aktivitetsintervensjonen ble introdusert med vektlegging av intensitet, glede, sikkerhet og uten konkurranse eller bedring av ferdighet. Imidlertid besto øktene av løpsleker, hoppetau, modifisert basketball og modifisert fotball (Davis et al., 2011).

Ifølge Tomporowski et al. (2015) har også effekten av kroniske kvantitative intervensjoner blitt undersøkt med utgangspunkt i hypotesen om en nevrofysiologisk effekt. Dishman et al. (2006) hevder at fysisk aktivitet og trening kan influere fordelaktig på *synaptisk plastisitet* og *vekst i aksoner* og på den måten fasilitere prosesser som inngår i dannelsen, tilpasning og beskyttelse av nerveceller. Noen av disse prosessene er påvirket av *neurotrophin faktor* (BDNF) (Binder & Scharfman, 2004). Foreløpig antagelser (Dishman et al., 2006) er også at kronisk fysisk aktivitet kan *svekke stressrespons* i deler av hjernen som regulerer perifer sympatisk aktivitet, og foreslår en redusert sympatisk respons til stress som videre reduserer faren for kliniske forstyrrelser (som høyt arterielt blodtrykk (hypertensjon), hjertesvikt, oksidativt stress og undertrykking av immunrespons).

2.5.4 Kvalitative akutte intervensjoner

Nylig har forskere utviklet intervensjoner som fokuserer spesielt på oppgaver som er komplekse og som gir en kognitiv effekt av mental mobilisering – «mental engagement» (Tomporowski et al., 2015). Felles for disse intervensjonene er bruken av aktiviteter som inneholder komplekse bevegelsessekvenser og som krever rask beslutningstaking. Bedret *oppmerksomhet* ble målt etter en akutt kvalitativ økt (Budde, Voelcker-Rehage, Pietraßyk-Kendziorra, Ribeiro & Tidow, 2008) sammenlignet med ordinær kroppsøving. Bedret *arbeidsminne*, ved lagring og rekapitulering - «memory encoding», ble målt etter en akutt kvalitativ økt (Pesce, Crova, Cereatti, Casella & Bellucci, 2009) sammenlignet med en økt uten fysisk aktivitet eller mental mobilisering. Bedret *inhibering* (Flanker test) ble målt for akutt fysisk aktivitet med og uten kognitiv mobilisering (Best & Eccles, 2012).

2.5.5 Kvalitative kroniske intervensjoner

Nylig har teoretisering og forskning ført til design av kroniske treningsintervensjoner som med hensikt inkluderer spill og fysisk aktivitet som har som mål å utfordre både fysisk og

kognitivt og som drar nytte av sosiale relasjoner mellom barn og lærere (Tomporowski et al., 2015). Bedret *matematikkprestasjon, kognitiv- og affektiv selvregulering* og klasseromsatferd ble målt etter et kronisk kvalitativt treningsprogram (3 mnd.) sammenlignet med ordinær kroppsøving. Treningen vektla instruksjon som inneholdt selv-monitorering og et treningsmiljø som var basert på respekt, disiplin, og selv-kontroll. Treningslærer var spesialist innen Martial arts (kampsport) (Lakes & Hoyt, 2004). Bedret *kognitiv funksjon* og økt nevrologisk aktivitet ble målt etter et kronisk kvalitativt treningsprogram (9 mnd.), men uten sammenligningsgruppe (Castelli, Hillman, Hirsch, Hirsch & Drollette, 2011). Bedret *inhibering* ble målt etter et kronisk kvalitativt treningsprogram (2 timers økter, 6 mnd.) sammenlignet med 1 times ordinær kroppsøving (Crova et al., 2013). Bedret *responstid og presisjon* ble målt etter et kronisk kvalitativt treningsprogram (8 uker), men uten sammenligningsgruppe (Chang, Tsai, Chen & Hung, 2013). Bedret *akademisk prestasjon* og klasseromsatferd ble målt etter et kronisk kvalitativt treningsprogram (Take 10!) (Kibbe et al., 2011). Bedret *oppmerksomhet* ble målt etter kronisk kvalitativt klasseromsbasert treningsprogram – «10 min energizers» (3 mnd.) sammenlignet med ordinær undervisning (Mahar et al., 2006). Bedret *akademisk prestasjon* ble målt etter kronisk kvalitativt klasserombasert treningsprogram – «short PA breaks» (5 mnd.) sammenlignet med ordinær undervisning (Erwin, Fedewa & Ahn, 2012). Pesce et al. (2013) målte bedret *eksekutiv oppmerksomhet* etter kroniske kvalitative intervensjoner med kognitive krav som utfordret eksekutive funksjoner.

Det er foreslått at kvalitative intervensjoner gir en fordelaktig kognitiv effekt som ikke oppnås gjennom kvantitative intervensjoner (Davis et al. (2011) i Pesce et al., 2012). Mental mobilisering gir en tilleggseffekt. Rasjonale bak tilleggseffekten er hypotesen om en kontekstuell påvirkning, som har anvendelse for motorisk ferdighet og forskning på rehabilitering (Tomporowski et al., 2010). Flere studier har vist at læring påvirkes av kontekst. Tomporowski et al. (2010) bruker begrepet «kontekst» om instruksjon under trening. Instruksjonen kan gi et stabilt miljø med repeterende praksis versus uforutsigbare forhold med tilfeldig praksis (Tomporowski et al., 2010). Studier har vist en fasiliterende effekt ved variert ferdighetstrening. Bedret *innlæring* ble forklart med mengden mental involvering som krevdes. Under variert instruksjon må individet inhibere tidligere bevegelsesplan og erstatte med en ny bevegelsesplan. Endring av bevegelsesplaner fra en oppgave til en annen innebærer flere mentale operasjoner. Individet må gjenkjenne omgivelsen som definerer oppgaven. Innhente et prosedyreprogram fra langtidshukommelsen, og iverksette

bevegelse i form av atferd. Under en uforutsigbar kontekst kreves mer mental innsats enn ved repetisjon av samme bevegelse i hver økt. En nyttig konsekvens av varierte treningsprogram er individets evne til å tilpasse atferd ved overføring til en generell kontekstuell variasjon (Tomporowski et al., 2010).

2.6 Måleresultater fra den senere tids forskning

Álvarez-Bueno et al. (2017) presenterer kroniske skolebaserte studier (barn og ungdom 4 -18 år) hvorav de fleste ble gjennomført som randomiserte eksperimenter. Metaanalysen baseres på en modell av eksekutive funksjoner (Diamond, 2013) som avdekker samlet effektstørrelse (pooled ES estimation) og heterogenitet (jmf. 8.1 Vedlegg 1: 8.1.1 Begreper) basert på individuelle karakteristika, intervensjonskarakteristika og kontekst (hvor i skolens planer).

2.6.1 Ikke-eksekutive og kjerne-eksekutive funksjoner

For måling av *ikke-eksekutive funksjoner*, ble ES beregnet til 0.23 (95% KI = 0.09 – 0.37). Liten heterogenitet ble funnet mellom studier ($I^2 = 21.9\%$; $p = 0.199$) (Álvarez-Bueno et al., 2017).

For målinger av *kjerne-eksekutive funksjoner*, ble samlet ES beregnet til 0.20 (95% KI = 0.10 – 0.30). Stor heterogenitet ble funnet mellom studier ($I^2 = 70.0\%$; $p = 0.001$) (Álvarez-Bueno et al., 2017).

En delanalyse undersøkte spesifikke eksekutive funksjoner. Delanalysen ga en beregning av ES på 0.14 (95% KI = 0.00 – 0.27) for *arbeidsminne*. Medium heterogenitet ble funnet mellom studier ($I^2 = 48.0\%$; $p = 0.027$). Delanalysen ga en beregning av samlet ES på 0.26 (95% KI = 0.10 – 0.41) for *selektiv oppmerksomhet – inhibering*, i tester som krevde kontroll av forstyrrelser. Stor heterogenitet ble målt mellom studier ($I^2 = 76.0\%$; $p = 0.001$). Delanalysen ble gjentatt separat for *inhibering* (ES = 0.38) og *selektiv oppmerksomhet* (ES = 0.13). Stor heterogenitet ble påvist i begge delanalyser (hhv. $I^2 = 86.2\%$ og 66.8%) (Álvarez-Bueno et al., 2017).

2.6.2 Metakognisjon

For målinger av metakognisjon, ble samlet ES beregnet til 0.23 (95% KI = 0.13 – 0.32). Liten heterogenitet ble funnet mellom studier ($I^2 = 4.7\%$; $p = 0.398$). Delanalysen av *eksekutive funksjoner (høyere orden)* ble beregnet til 0.19 (95% KI = 0.06 – 0.31). Liten heterogenitet

ble funnet mellom studier ($I^2 = 12.9\%$; $p = 0.315$). For delanalysen av «*bevissthet om læring*» ble samlet ES beregnet til 0.30 (95% KI = 0.15 – 0.45). Ingen heterogenitet ble funnet mellom studier ($I^2 = 0.0\%$; $p = 0.549$) (Álvarez-Bueno et al., 2017).

2.6.3 Delanalyser som undersøker effekten av ulike intervensjoner

Delanalyser som undersøkte effekten av ulike typer kroniske intervensjoner (kvantitative, kvalitative eller begge), viste at forsterket fysisk aktivitet (kvantitative økter) styrker ikke-eksekutive funksjoner. Forsterket fysisk aktivitet styrker arbeidsminne. Beriket fysisk aktivitet styrker selektiv oppmerksomhet – inhibering. Kognitiv fleksibilitet styrkes ikke av noen form for fysisk aktivitet (Álvarez-Bueno et al., 2017).

For metakognisjon viste delanalysen at forsterket fysisk aktivitet styrker eksekutive funksjoner (høyere orden), mens kun beriket fysisk aktivitet har en effekt på «*bevissthet om læring*» (Álvarez-Bueno et al., 2017).

Teorikapittelet inneholder en selvstendig enhet teori fra forskningsfeltet, og kan leses separat. Metode, resultat og diskusjonskapittelet er en redegjørelse av mitt arbeid i dette systematiske reviewet og bør lese sammenhengende.

3.0 Metode

Utgangspunktet for denne undersøkelsen var å identifisere relevante studier. Primærstudier ble generert ved systematiske databasesøk. Bakgrunnen for valg av rapportform – systematisk review blir forklart. Variablene som brukes for fysisk aktivitet og kognisjon og akademisk prestasjon, blir konkretisert og beskrevet. Flytdiagram gir en oversikt med identifikasjon, screening og valg av inkluderte og ekskluderte artikler. En del av metoden består i å beskrive hvordan studier blir kategorisert og kodet, hvordan metodologisk kvalitet vurderes og anvendes. Metodologisk kvalitet og gradert evidens blir brukt som et mål på studiepreferanse eller hvilke måleresultater som har størst troverdighet. Metode utdyper regler for hvordan syntesen presenteres i et Harvestplott og hvordan funn i konkurrerende hypoteser blir vurdert etter «best evidens».

3.1 Inklusjonskriterier

Rapporten tar for seg effekten av fysisk aktivitet på kognisjon og akademisk prestasjon. Inklusjonskriteriene innbefatter valg av populasjon/utvalg, intervensjonsgruppe versus

kontroll- eller sammenligningsgruppe, resultat (utfall) og studiedesign. Studiene blir systematisk beskrevet etter malen «population, intervention, comparison, outcome, study design» PICOS (Higgins, Green & Cochrane collaboration, 2009).

3.1.1 Valg av populasjon og utvalg

Det var i utgangspunktet ønskelig med populasjoner som er like eller lignende. Utvalget ble bestemt å skulle bestå av forsøkspersoner i alder 18 år eller yngre og innenfor normal utvikling. Fysisk aktivitetsintervensjoner foregår ofte med utvalg fra skoler og klasser på barnetrinn, mellomtrinn, ungdomsskole eller gymnas. Forsøk skulle finne sted i den ordinære skolehverdagen, men inkluderte studier har også aktivitet etter skoletid og i friminutter.

3.1.2 Valg av intervensjonsaktivitet

Et sentralt inklusjonskriterium ble at intervensjonen skulle inneholde fysisk aktivitet. Fysisk aktivitet ble undersøkt ved søk for en rekke underkategorier: aerob aktivitet, løping, hoppetau og spill med intensitet. Motorisk aktivitet og ferdighetsutvikling, tradisjonell kroppsøving, fritt spill, basketball, volleyball, badminton, bordtennis, grunnleggende ferdighet, koordinasjon, balanse, lateralitet og objekt kontroll har også tidligere vært undersøkt i litteraturen (Vazou, Pesce, Lakes & Smiley-Oyen, 2016). Treningsfrekvens og tid per økt (dosering) av fysisk aktivitet blir en sentral del av beskrivelsen. Ingen studier ble ekskludert på grunn av varighet.

3.1.3 Kontroll eller sammenligningsaktivitet

Kontroll- eller sammenligningsaktiviteten er til en viss grad gitt ved denne typen forskningsprosjekter. Kontrollgruppen har enten ordinær kroppsøving eller ordinær undervisning. Elevene befinner seg i en ordinær skolehverdag. Alle har ordinær kroppsøving i løpet av uken. Alle har ordinær undervisning i akademiske fag og alle har friminutter med ulik aktivitet.

3.1.4 Valg av utfall

Et sentralt inklusjonskriterium ble at utfall skulle inneholde målinger av kognitive funksjoner og/eller akademisk prestasjon (Donnelly et al., 2016). Tidligere er utfall undersøkt for en rekke underkategorier: kognitiv psykologisk respons, eksekutive funksjoner, akademisk måloppnåelse, planlegging, kreativitet, arbeidsminne, romlig oppfatning og tankegang,

matematikk kompetanse, lesekompetanse, staving, reaksjonstid og nøyaktighet, oppmerksomhet, rekapitulering, kognitive skifter, språk og perseveranstendens (evne til opprettholdelse av aktivitet) (Vazou et al., 2016).

3.1.5 Valg av studiedesign

Studiedesign med betegnelsen randomisert eksperiment (Randomized controlled trials – RCT) og kluster randomiserte eksperimenter (cluster randomized controlled trials - CRT) ble inkludert. RCT ble i utgangspunktet valgt som et ledd i utviklingen av forskningsfeltet med et ønske om å kartlegge studier med tilstrekkelig metodologisk kvalitet og evidens. RCT er i utgangspunktet en klinisk konstruksjon, mens fysisk aktivitetsintervensjoner bygger på utvalg som er inndelt i klasser (dvs. utvalget har en klusternatur). Følgelig ble også CRT en viktig del av undersøkelsen. Ifølge Singh, Uijtdewillingen, Twisk, Van Mechelen & Chinapaw (2012) finns en sammenheng mellom høy metodologisk kvalitet og signifikant positive resultater for sammenhengen mellom fysisk aktivitet og kognisjon. Flere RCT- eller CRT-design legger til rette for en sterk evidens (Berghmans et al., 1998).

3.2 Søkeprosessen

En rekke innledende testsøk ble foretatt. Tittelsøk, termsøk og boolske søk ble prøvd ut (jmf. 8.2 Vedlegg 2: Tverrfaglig databasesøk). Søkestreng ble utarbeidet. Medical subject headings (MeSH) ble undersøkt med hensyn til presisjon i søk. MeSH brukes som indekserte begrep, men viste seg ikke spesielt nyttig i disse ikke-medisinske søkene.

3.2.1 Valg av søkeord

Treff i databasene skulle baseres på at alle kategorier var representert i artiklene. Søkestrengen i hver kategori S1 – S4 ble satt sammen med logisk «OG». Treff i databasene skulle baseres på minst ett av søkeordene i hver søkestreng. Hver kategori (eksempel S3) ble satt sammen av søkeord med logisk «ELLER». Søkeord ble videreutviklet etter relevante artikler og rammeverk for strukturerte spørsmål. Søkestrengene ble tilpasset inklusjonskriteriene (populasjon, intervensjon/kontroll, utfall og forskningsdesign):

S1: infancy, childhood, children, adolescents, pupil, young people.

S2: Activity, aerobic exercise, physical activity, physical fitness, exercise, movement skills, physical education, motor skills training, motor skills.

S3: learning, school health, executive function, cognition, cognitive function, cognitive performance, cognitive ability, cognitive engagement, intelligence, academic achievement, academic performance, scholastic achievement, grading.

S4: randomized controlled trials, RCT, randomized control trials, random control trial.

Tabell 2 viser maskinkode fra en linje i et systematisk søk. Søkestrengen S3 er gitt som et boolsk uttrykk. Søket genereres fra en rekke databaser. Antall treff er 2,9 millioner for S3 utfall.

Tabell 2 Søkestreng S3 som består av søkeord - underkategorier av resultat (utfall). Søkeordene er satt sammen logisk med «ELLER». Antall treff for resultater(utfall) er 2919828.

	Query (Søkestreng)	Limiters/Expanders (Begrensing i streng)	Last Run Via (Søk i følgende databaser)	Results (Antall treff)
S3	learning or school health or executive function or executive functioning or cognition or cognitive function or cognitive performance or cognitive ability or cognitive engagement or intelligence or academic achievement or academic performance or scholastic achievement or grading	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - SPORTDiscus; CINAHL Plus with Full Text; MEDLINE; Academic Search Complete; Health and Psychosocial Instruments; ERIC	2919828

3.2.2 Tverrfaglig databasesøk

Det ble foretatt et tverrfaglig databasesøk med bruk av flere databaser samtidig. Følgende databaser ble brukt: SPORTDiscus, CINAHL Plus with Full Text, MEDLINE, Academic Search Complete, Health and Psychosocial Instruments, og ERIC under EBSCO host.

Disse databasene inneholder typisk relevant forskningslitteratur innen idrett og helse.

Totalt antall treff etter logisk begrensning (S1 AND S2 AND S3 AND S4) var 1807. Det systematiske søket ble foretatt i løpet av uke 24 og 25, 2017. Tabell 3 viser søkeresultatene etter logisk begrensning, peer reviewed begrensning og etter tidsbegrensning (2007 – 2017) (se også 8.2 Vedlegg 2: Tverrfaglig databasesøk).

Tabell 3: Tverrfaglig databasesøk, Resultatet viste 1807 treff etter S1 - S4 var logisk begrenset, tidsbestemt, engelskspråklig, og peer reviewed.

Query (Boolsk sammensatt søkestreng)	Limiters/Expanders (begrensinger)	Last Run Via (Søk i følgende databaser)	Results (Antall treff)
S1 AND S2 AND S3 AND S4	Limiters - Peer Reviewed; Published Date: 20070101-20171231 Narrow by Language: - english Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database – SPORTDiscus; CINAHL Plus with Full Text; MEDLINE; Academic Search Complete; Health and Psychosocial Instruments; ERIC	1807

3.2.3 Systematiske søk i enkeltdatabaser

Siden hver database har ulike kriterier for eksklusjon ble det i tillegg, foretatt separate søk i enkeltdatabaser. Følgende databaser ble gjenstand for nye søk: Medline, SPORTDiscus og ERIC. Sluttresultatet (etter Endnote) for Medline var det samme som for det tverrfaglige databasesøket, 32 treff. I SPORTDiscus fremkom 14 treff, ingen supplerende til Medline. I ERIC fremkom 3 treff, 2 tilleggsstudier, i forhold til Medline. Totalt antall eligerte artikler etter separate søk var 34. Søk i enkeltdatabaser ble foretatt i løpet av uke 32. Tabell 4 viser maskinkode for en linje systematisk databasesøk for søkestreng S2 fysisk aktivitet. Søkestrengen var satt sammen boolsk, og søket ble gjennomført i databasen ERIC. Antall treff ble 300.000.

Tabell 4: Databasesøk i en database – ERIC, resultatet viser 300979 treff for kategori S2 fysisk aktivitet.

#	Query	Limiters/Expanders	Last Run Via	Results
S2	Activity or aerobic exercise or physical activity or physical fitness or exercise or movement skills or physical education or motor skills training or motor skills or physical	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - ERIC	300979

3.3 Utvelgelse av litteraturreferanser ved PRISMA's flytdiagram

Endnote er en applikasjon som tilrettelegger for personlig lagring og redigering av litteraturreferanser. I Endnote ble alle importerte referanser screenet for duplikater (resterende 1295). Figur 4 under viser flytskjema med reduksjon av artikler. Artikkeltitler ble screenet og ikke-

eligerte artikler ble ekskludert. Utvelgelsesprosedyren ble effektivisert ved å søke en rekke søkeord (primært patologiske tilstander) for en rask eksklusjon av artikler. Artikler som inneholdt: health, patient, infection, disabilities, injury, vaccination, antibiotic, obesity, overweight, fat, pregnancy, pelvic, diet, nutrient, supplementation, food, fracture, polio, arthritis, asthma, autism, acne, ADHD, insulin, pain, therapy, cardiovascular, apnea, oncology, malaria, hiv, alcohol, tobacco, smoking, sleep, meta-analysis og Cochrane review, ble ekskludert.

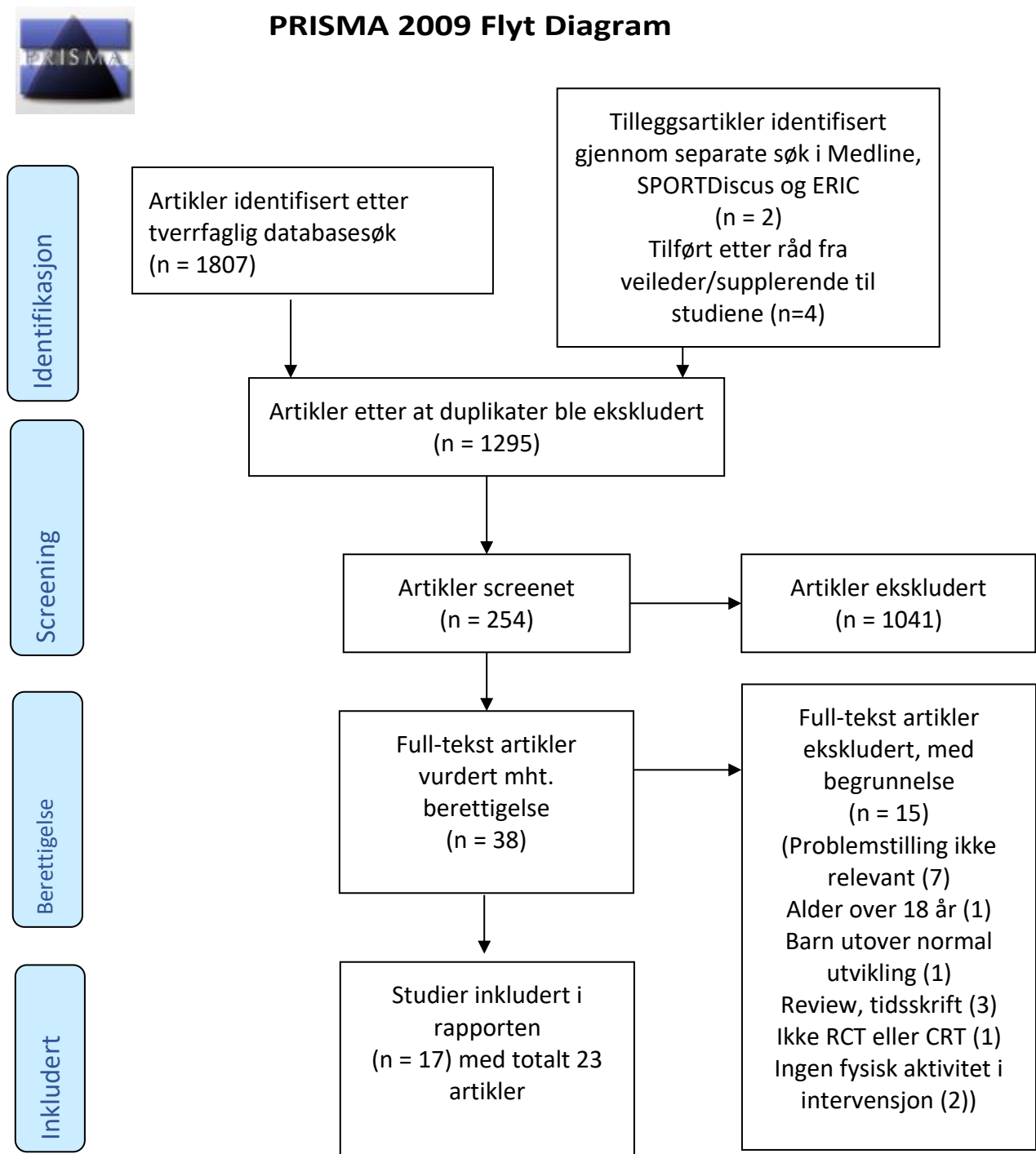
Resterende artikler (254) ble gjennomgått manuelt med hensyn til tittel og sammendrag. Referanser som ikke besvarte problemstillingen ble ekskludert. I tillegg ble 4 artikler som ikke ble funnet i det systematiske databasesøket inkludert. Årsaken til at disse artiklene ikke ble inkludert i det systematiske søket kan være at søkeord ikke matchet eller at de logiske begrensningene ekskluderte disse artiklene. En artikkel ble inkludert etter råd fra veileder. Tre artikler ble inkludert som supplement til eksisterende studier, og var pilotstudier eller artikler som ga en metodologisk forklaring på forskningsdesign.

For resterende (38) litteraturreferanser ble hele teksten gjennomgått. Av disse 38 artiklene ble 15 artikler ekskludert av ulike grunner. Begrunnelser for eksklusjon (antall referanser) var som følger:

- Problemstillingen var ikke relevant (7)
- Utvalget i alder over 18 år (1)
- Utvalget med barn utover normal utvikling (1)
- Review studier, eller populærvitenskapelige tidsskrift (3)
- Forskningsdesignet var ikke RCT eller CRT (1)
- Intervensjonen inneholdt ikke fysisk aktivitet (2)

Rapporten ender opp med 23 artikler i 17 inkluderte studier (forskningsprosjekter). Blant 17 forskningsprosjekter var 9 studiedesign randomiserte kontrollerte eksperimenter (randomized controlled trials, RCT), mens 8 studiedesign ble gjennomført som kluster randomiserte eksperimenter (cluster randomized trials, CRT).

PRISMA 2009 Flyt Diagram er et standardisert flytdiagram og gir en oversikt over eksklusjon og inklusjon av studier. Antallet studiene er presentert for tverrfaglig databasesøk og enkeltvise databasesøk.



Figur 4: Flytdiagram for systematisk databasesøk.

3.4 Beskrivelse av inkluderte studier

17 studier tilfredsstilte inklusjonskriteriene. Det ble foretatt valg av hvordan hver studie skulle beskrives i tabell. Den deskriptive informasjonen skulle gi et innblikk i det vesentlige ved studiene.

3.4.1 Deskriptiv informasjon

En sammenligning av alle studier viste at de hadde ulik varighet. På bakgrunn av en ide om treningstilpasning til fysisk aktivitet, ble studiene sortert etter varighet. Det høye antallet studier gjorde det nødvendig å dele opp informasjonen i flere tabeller. Deskriptiv tabell inneholder studier med intervensjoner som varer i 1 uke – 3mnd., 3 – 6 mnd., 6 – 12 mnd., og 12 - 24 måneder.

Tabell 5 viser en linje fra deskriptiv tabell (jmf. 4.1 Beskrivelse av inkluderte studier).

Enkeltstudier ble sortert etter strukturen PICOS (eng: population, intervention, comparison, outcomes, study design) (Booth et al., 2016), der hver studie (St1 – St17) beskrives med forfatter, utgivelsesår og forskningsdesign. Inkluderte studier er publisert i perioden 2007 – 2017 og gir et oppdatert bilde av forskningsfeltet. I noen tilfeller har hver studie generert flere publikasjoner. Referanselisten utdyper hver litteraturreferanse ytterligere (jmf. 7.0 Referanser).

Tabell 5: Beskrivelse av ett studie med utvalg, type treningsintervensjon og kontrollgruppe, varighet, kognitive tester og måleresultater.

Studie	Utvalg	Type Treningsintervensjon I og sammenlignings- eller kontrollgruppe C	Varighet av I og C	Kognitive tester	Resultater
St12 Resaland et al. (2015; 2016), Aadland et al. (2017) CRT	N=1129 rekr., 5 trinn, 10 år. N=620, I N=582, C	Aerob og akademisk vs kroppsøving I: fysisk aktivitet + ASK ASK: fysisk aktiv undervisnings- leksjon, (løping, stafett, hinderløype og aktiv lek). aktive friminutt med fysisk aktivitet, hjemmelekse fysisk aktivitet. C: kroppsøving + fysisk aktivitet.	7 mnd. I tot. 300min/uke ASK: 165 min/uke 3*30 min/uke + 5 min/ dag + 10 min/dag, C: 135 min/uke,	AP, EF, SCWT, inhibering VF, kognitiv fleksibilitet TMT, kognitiv fleksibilitet WISC-IV, arbeidsminne	I: 0 på AP - lesing, regning, engelsk, nasjonale prøver (UDIR). Analyse per protokoll: I: + på EF sammensatt. I: + VF, TMT, kognitiv fleksibilitet Høyest progresjon for deltagere med lav baseline.

I tabell 5 ble utvalget i hver studie beskrevet med antall deltagere (N = 1129 rekruttert).

Artiklene gir informasjon om antallet rekrutterte- eller analyserte deltagere eller begge. I noen tilfeller oppgis antallet allokerede deltagere i de ulike gruppene (intervensjons- eller

kontrollgruppe). Tabell 5 viser deltagernes alder presentert som et aldersinkrement eller som gjennomsnittsalder \pm standardavvik eller begge. Alle studiene var skolestudier. Trinntilknytning ble oppgitt i artikler, der klassetrinn fremgikk i primærstudiene.

Tabell 5 viser type treningsintervensjoner og kontrollaktivitet for hver studie, beskrevet stikkordsbasert med innhold. Studien hadde både en og to intervensjoner i tillegg til kontrollaktivitet. Hensikten var å få frem essensen i «behandlingen» som er kvantitativ eller kvalitativ eller begge deler (jmf. 2.4.2 Kvantitative og kvalitative, akutte og kroniske intervensjoner).

Tabell 5 viser også varigheten av intervensjon og kontrollaktivitet beskrevet med total tidsbruk. Artikkelen gir informasjon om tidsbruk per økt, treningsfrekvens per uke og totalt antall uker. Intervensjon og kontroll hadde samme tidsbruk, men ulik aktivitet.

Tabell 5 viser kognitive tester beskrevet for hver studie. Tabellen beskriver de kognitive testene eller de akademiske testene med forkortelser. Under hver tabell fremgår en forklaring på forkortelser. I 8.1 Vedlegg 1: Glossarium – begreper, forkortelser og administrering av testinstrumenter, finnes en oversikt over forkortelser og administrering av testinstrumenter i hver primærstudie.

Tabell 5 viser resultater beskrevet for hver studie. Artikkelen gir informasjon om hvordan intervensjonen står seg i forhold til kontrollaktivitet. Resultatene er oppgitt som I: +, I: 0 eller I: –, dersom intervensjonen hhv. har en bedre effekt, lik effekt eller dårligere effekt sammenlignet med kontrollaktiviteten. I tillegg fremgår hvilke kognitive eller akademiske tester som har gitt utfallet. I noen tilfeller fremgår en ytterligere forklaring på resultatene.

3.4.2 Metode for koding av inkluderte studier

Vazou et al. (2016) og andre utviklet systematisk koding for identifisering av fremtredende egenskaper ved innholdet i intervensjon- og kontrollaktivitet (Vazou et al., 2016; Best, 2010; Cliff et al., 2016). En tilsvarende koding ble utviklet og presentert i denne rapporten. Koden «**forsterket intensitet vs akademisk**» ble brukt i St1 etter sammenligning av intervensjon og kontrollgruppe (intervensjonsulikhet). «Forsterket intensitet» betyr at intervensjonen inneholdt en økt fysisk aktivitet med økning i intensitet. I tabell 6 nedenfor fremgår grunnlaget for hvordan kodene fremkom for studie 1, 2, og 3 (St1, St2, St3).

St1 har en intervensjonsgruppe og en kontrollgruppe. Kodene forklarer intervensjonsulikhet, der intervensjonen (I) har en aerob fysisk aktivitet med forsterket intensitet i forhold til kontrollgruppen. Kontrollgruppen (C) i St1 går direkte til kognitiv testing før ordinær kroppsøving. Kontrollgruppen kommer fra en ordinær skolehverdag (akademisk). St 2 har 3 grupper (I₁, I₂ og C) som sammenlignes. Dette gir tre koder for intervensjonsulikhet. Intervensjon 1 (I₁, 3 trinn) inneholder aerob fysisk aktivitet med forsterket frekvens i forhold til kontrollgruppen. Intervensjon 2 (I₂, 5 trinn) inneholder aerob fysisk aktivitet med forsterket frekvens i forhold til kontrollgruppen. Kontrollgruppen ble satt til å lese treningslitteratur (akademisk). Sammenligning av I₁ og I₂ er en undersøkelse av alderseffekten (5 trinn vs. 3 trinn). St3 har tre grupper (I₁, I₂ og C) som sammenlignes. Dette gir tre koder for intervensjonsulikhet. Intervensjon 1 (I₁) inneholder aerob fysisk aktivitet med forsterket frekvens i forhold til kontrollgruppen. Intervensjon 2 (I₂, *2) inneholder forsterket frekvens i forhold til kontrollgruppen. Kontrollgruppe ble satt til å gjøre simulerte skoleoppgaver (akademisk). Dette danner grunnlaget for kodingen: **Forsterket frekvens vs akademisk** for St3. Disse tre studiene undersøker effekten av en forsterket aerob fysisk aktivitet, og effekten av en kvantitativ akutt intervensjon (jmf. 2.5.2 Kvantitative akutte intervensjoner). Effekten måles på ulike kognitive funksjoner, primært eksekutive funksjoner.

Tabell 6: Grunnlaget for sammenligning av intervensjon og kontrollaktivitet for St1, St2 og St3. Koder for intervensjonsulikhet i St1, St2 og St 3.

	St1. koder: Forsterket intensitet vs. Akademisk Forsterket intensitet vs. Akademisk (etter 24 timer)	St2. koder: Forsterket frekvens vs. Akademisk Forsterket frekvens vs. Akademisk Alder: 5 trinn vs. 3 trinn	St3. koder: Forsterket frekvens vs. Akademisk Forsterket frekvens * 2 vs. Akademisk Forsterket frekvens * 2 vs Forsterket frekvens * 1.
Intervensjon	I: PACER, Progressive Aerobic Cardiovascular Endurance test. Progressiv intensitet, stadig raskere løp repetert over 20 m (piper test), inntil utmattelse (6:40 – 20 min). Deretter minnetest.	I₁: femte trinn: Jogging moderat intensitet (60 -70 % HR maks) på en bane i 30 min. Hvile (20 – 25 min) deretter test. I₂: tredje trinn: Jogging moderat intensitet (60 -70 % HR maks) på en bane i 30 min. Hvile (20 – 25 min) deretter test.	I₁: fysisk aktivitet, 20 min økt på morgenen. I₂: fysisk aktivitet, 20min på morgenen og deretter 20 min midtveis på morgenen. Akutt fysisk aktivitet, moderat intensitet.
Kontrollgruppe	C: Minnetest før ordinær kroppsøving. Ingen behandling. (friminutt og undervisning timen før).	C: lesing av treningsrelatert litteratur, stillesitting i klasserom i 30 min.	C: ingen fysisk aktivitet. Simulerte skoleoppgaver, stillesitting hele morgenen.

Kognitivt beriket vs. Akademisk er en kode som essensielt handler om fysisk aktivitet som er tilført et kognitivt stimulerende innhold. Intervensjonene som beskriver kognitivt beriket

favner et bredt spekter av aktiviteter (jmf. Vedlegg 1 8.1.1 Begreper). Resultatkapittelet viser tabeller med koding av inkluderte studier (jmf. 4.1 Beskrivelse av inkluderte studier).

3.5 Metodologisk kvalitet

Denne rapporten følger definisjonen til Verhagen for metodologisk kvalitet. Metodologisk kvalitet gir et mål på sannsynligheten for at eksperimentelle design genererer resultater uten feil, som er tilstrekkelig presise og som tilrettelegger for en klinisk anvendelse (Verhagen i Maher, Sheerington, Herbert, Moseley & Elkins, 2003).

3.5.1 Valg av sjekklister for vurdering av metodologisk kvalitet

Innledningsvis ble flere sjekklister for metodologisk kvalitet undersøkt. En systematisk review (Singh et al., 2012) undersøkte kvaliteten i artikler som omhandlet fysisk aktivitet og skoleprestasjon. Singh et al. (2012) brukte en sjekklister av Tooth, Ware, Bain, Purdie & Dobson (2005), for vurdering av metodologisk kvalitet i longitudinelle observasjonsstudier. Singh et al. (2012) valgte å kvalitetsvurdere alle inkluderte studier etter en standard.

Tooth et al. (2005) undersøker ekstern validitet som viser i hvilken grad funn fra et utvalg kan generaliseres til en målpopulasjon. Tooth et al. (2005) undersøker avvik i målinger for intervensjon og utfall. Konfundere kartlegges. En konfunder er en variabel som skyldes avvik ved seleksjon av individer til et utvalg (eks. dropout). Sjekklisten søker blant annet å besvare spørsmålet: er antallet deltagere rettferdiggjort gjennom statistisk styrke (power beregning)?

En tilsvarende sjekklister er utviklet av Downs and Black som undersøker artikkelens rapportering, eksternvaliditet, internvaliditet og statistisk styrke (Downs & Black, 1998). Sjekklisten er omfattende og tilpasset undersøkelser av randomiserte kontrollerte eksperimenter (RCT).

I denne studien ble forskningsdesign avgjørende for valg av sjekklister for vurdering av metodologisk kvalitet. Sjekklister for randomiserte kontrollerte eksperimenter (RCT) av Maher et al. (2003) ble valgt (jmf. 4.3.1 Metodologisk kvalitet for inkluderte RCT). Sjekklisten fokuserer på: eligeringskriterier, randomisert allokering, blinding, dropout, baselinemålinger, intervensjon og kontroll, presentasjon av resultater og behandlingseffekt. Sjekklister for kluster randomiserte eksperimenter (CRT) av Campbell, Elbourne & Altman (2004) ble valgt (jmf. 4.3.2 Metodologisk kvalitet for inkluderte CRT). Sjekklisten inneholder kontrollpunkter som

er spesielt tilpasset kluster randomiserte eksperimenter. Sjekklisten tar gjennomgående for seg kluster vs individ, randomisering, resultater og diskusjon. Sjekklistene er oversatt til norsk og kontrollert av veileder. I vurderingen av metodologisk kvalitet er det rom for feil. Dersom primærstudien mangler rapportering eller man ikke nevner nøkkelkriterier som eksempel, «intention-to-treat analysis» eller «concealed allocation» vil kvalitet bli vurdert lavere enn den faktisk er.

3.5.2 Anvendelsen av metodologisk kvalitet

I Ifølge Maher et al. (2003) kan resultatene fra en kvalitetsvurdering benyttes på flere måter. I systematiske review kan studier med lav metodologisk kvalitet ekskluderes fra analysen (Herbert & Gabriel, 2002).

Metodologisk kvalitet kan også brukes til en vektning av studier og kategorisere etter tilstrekkelig eller lav kvalitet uten å ekskludere studier (Berghmans et al., 1998). Konsekvensen av en kvalitetsvurdering er at konklusjonen kan endres i metaanalyser (kvantitativ metode) eller i metasynteser (kvalitativ metode – eidenstilnærming).

Kvalitetsanalyser kan også anvendes utenfor systematiske review. Kvalitetsbetraktninger kan inngå i planlegging eller rapportering av forskningsprosjekter, eller inngå som en del av vurderingen ved klinisk implementering av resultater (Berghmans et al., 1998).

3.5.3 Kategorisert metodologisk kvalitet og gradert evidens

I denne rapporten er mangfoldet i undersøkelser avgjørende for valg av en kategorisert metodologisk kvalitet og gradert evidens. På den måten kan måleresultater for visse studier presenteres med større grad av troverdighet. Metodologisk kvalitet ble kategorisert med enten «tilstrekkelig» eller «lav» kvalitet. Gradert evidens kan indikere hvilke undersøkelser man kan betrakte som robuste og hvilke undersøkelser som er tentative (jmf. 4.4.6 Syntesens robusthet).

Nivået av evidens er avgjørende for den overordnede konklusjonen i studien. RCT ble klassifisert som hos Berghmans et al. (1998) med «tilstrekkelig» metodologisk kvalitet for kvalitetsscore $\geq 60\%$ eller «lav» metodologisk kvalitet for kvalitetsscore $< 60\%$ (Berghmans et al., 1998).

Nivået av evidens (Berghmans et al., 1998) ble rangert som følger:

- Nivå 1, *sterk evidens* krever flere klinisk relevante RCT med tilstrekkelig metodologisk kvalitet som peker i en bestemt retning.
- Nivå 2, *moderat evidens* krever ett klinisk relevant RCT med tilstrekkelig metodologisk kvalitet og flere RCT med lav kvalitet.
- Nivå 3, *svak evidens* krever ett klinisk relevant RCT med tilstrekkelig metodologisk kvalitet eller flere RCT med lav metodologisk kvalitet.
- Nivå 4, *ingen evidens* krever ingen klinisk relevante RCT og kan også inneholde selvmotsigende måleresultater.

I denne studien betyr «klinisk relevans» at studien har påvist en progresjon i kognitiv funksjon eller akademisk prestasjon.

3.6 Typen review

Review kan sammenstilles på ulike måter (Booth, Sutton & Papaioannou, 2016). I løpet av masterprosessen ble blant annet et metodologisk review vurdert. Etter at data forelå for inkluderte studier ble både en meta-analyse og et systematisk review tatt med i betraktningen. Valget falt på et systematisk review (Booth et al., 2016) som innbar en kategorisering av publikasjoner, en data ekstrahering, en kvalitetsvurdering og en narrativ syntese med presentasjon av data fra et Harvestplot (jmf. 4.4 Syntese av inkluderte studier – et narrativ). Avgjørende for valg av et systematisk review, (og ikke en meta-analyse) var effektstørrelse som ikke entydig fremgår i alle inkluderte studier. Det systematiske reviewet inneholder en vurdering av validitet (Booth et al., 2016) med fokus på eksternvaliditet og internvaliditet. En sensitivitetsanalyse (Booth et al., 2016) utforsker variabler som kan influere på konklusjonen i reviewet. Mangler i primærstudiene og reviewet (Booth et al., 2016) blir systematisk gjennomgått ved at ulike former for avvik blir analysert og diskutert.

3.6.1 Presentasjon av syntese

Ifølge Booth et al. (2016) er et Harvestplot en metode for å sammenstille forskning med ulike design og utfall for å gi en tilgjengelig oversikt (jmf. 4.4 Syntese av inkluderte studier – et narrativ). Harvestplottet kan presentere studienes egnethet, kvalitet og måleresultater. Tabell 7 viser en ramme for studier med positiv, ingen eller negativ gradient (funn har positiv, ingen eller negativ effekt). For hver studie avdekkes informasjon om utfall, studiens egnethet og kvalitetsvurdering. Leser kan selv vurdere hvor flesteparten av undersøkelsene er foretatt,

hvilke studier som er best egnet og hvor studier med høyeste kvalitet befinner seg (Oglivie et al., 2008).

Ifølge Booth et al. (2016) gir Harvestplottet en ny innovativ tilnærming til oppsummering av resultater i et systematisk review som undersøker effekten av intervensjonsulikheter.

Prosedyrene som følger i utviklingen av et Harvestplot er som følger:

1. definere hypotesene som skal testes (Eks. hypotesen: Funn gir positiv gradient)
2. allokere studiene til hypotesene som egner seg best (Eks. Funn viser ingen gradient).
3. plote fordelingen av evidens (funn).
4. skrive et narrativ – syntese som tolker Harvestplottet.

Tre hypoteser fremgår i tabell 7. Hypotesene: «Effekten har en positiv gradient, ingen gradient eller en negativ gradient». Alle undersøkelser er basert på sammenligning av to størrelser (Eks. Forsterket frekvens vs. Kroppsøving). Hver størrelse er målt før og etter for intervensjonen og kontroll. Dette gir to pluss to punkter. Mellom hvert punktpar kan man trekke en linje (lineær modell). Linjens stigningstall forteller om en type intervensjon er bedre enn en annen. Både som absoluttverdi men også som en gradient (stigningstall) som utvikles over tid.

Tabell 7: Harvestplot med eksempler på inkluderte studier.

Koder for typen ulikheter ved intervensjoner	Negativ gradient (effekt)	Ingen gradient	Positiv gradient
Forsterket frekvens vs. kroppsøving		⑤ ⑤	
Forsterket intensitet vs. Kognitiv beriket	⑦		

Hver kategori inneholder to enheter som sammenlignes. Tabell 7 viser kategorien «Forsterket frekvens versus kroppsøving». Funn viser at sammenligning av fysisk aktivitet med forsterket frekvens er like effektiv som ordinær kroppsøving for kognitiv funksjon og akademisk prestasjon (⑤ ⑤). Rekkefølgen av kategorienhetene er avgjørende for plassering i tabellen. Eksempel «Forsterket intensitet versus Kognitiv beriket» viser en negativ gradient for kognitiv funksjon (⑦). Dette betyr at første enhet (Forsterket intensitet) i intervensjonsulikheter har gitt en svakere effekt enn siste enhet (Kognitiv beriket).

Harvestplottet egner seg som en teknikk for å presentere funn fra et systematisk review hvor studier måler divergerende og komplekse utfall. Metoden kan brukes til å summere forskningsresultater fra effekten av intervensjoner. Harvestplottet viser også metodologisk kvalitet. Metodologisk kvalitet ble vurdert og vektet etter en skala fra 0 – 100% (jmf. 4.3 Metodologisk kvalitet). Tallverdien for metodologisk kvalitet er videreført for Harvestplottet som en verdi mellom 0 - 10. Tallverdien 0 – 5 indikerer en lav metodologisk kvalitet, mens en tallverdi 6 – 10 indikerer en tilstrekkelig metodologisk kvalitet (jmf. 3.5.3 Kategorisert metodologisk kvalitet og gradert evidens).

Harvestplottet avdekker måleresultatene for hver undersøkelse. Medierte gradienter viser målinger av medierte resultater, det vil si en eller annen form for kognitiv funksjon (blank sirkel), mens endelige måleresultater avdekker akademisk prestasjon (svart sirkel).

Harvestplottet benyttes for å sammenfatte resultatene fra akutt og kroniske studier, for store ($N > 1000$) og små ($N < 100$ per gruppe) studier. Syntesen brukes til å utvikle en modell for sammenhengen mellom intervensjonsulikhet og måleresultater. Mønstre i syntesen undersøkes og heterogenitet kartlegges.

3.6.2 Diskusjon av begrensing i syntesens robusthet

Resultatene kan avdekke eventuelle svake punkter i syntesens robusthet. Diskusjonskapittelet (jmf. 5.4 Begrensinger i syntesens robusthet) omfatter en behandling av ikke-signifikante resultater, to tilfeller av konkurrerende hypoteser og tre tilfeller av sterk evidens.

Sammenfallende funn med sterk evidens behandles med en vurdering av heterogenitet for om mulig å avdekke forskjeller mellom undersøkelser som kan redusere evidensnivå.

Konkurrerende hypoteser er funn som står direkte mot hverandre. For å avgjøre hvilke funn som har beste evidens brukes prinsippet om «best evidens». Prinsippet er formulert av Slavin (1995) og får sin anvendelse når en studie ikke umiddelbart kan erstatte et annet. Prinsippet er hentet fra rettsalen med bevisvurdering – «Et bevis kan erstattes av beste tilgjengelige bevis». I rettsalen har man fastsatt en rekkefølge for kriterier som avgjør bevisbyrden. Mens i forskningen er rekkefølgen av vurderingskriterier ikke fastlagt. I denne rapporten blir følgende kriterier brukt for å avgjøre hvilke funn som kan erstatte motsatsen: metodologisk kvalitet, evidensnivå for studiedesign, deretter brukes primærstudienes rapportering av begrensninger ved eget forskningsprosjekt. Denne rekkefølgen av kriterier blir avgjørende for hvilke funn som kan beholdes og hvilke funn som kan forkastes.

4.0 Resultater

Inkluderte studier blir beskrevet og presentert. Studiene blir kodet og vurdert for metodologisk kvalitet. Data transformeres til synteseplanet og beskrives narrativt. En empiri-tro modell ble utviklet. Syntesen undersøkes for mønstre og ulike karakteristika. Metodologisk kvalitet brukes i vurderingen av syntesens robusthet. Store studier presenteres som indikasjon på generelle resultater og små studier presenteres som indikator på publikasjonsavvik.

4.1 Beskrivelse av inkluderte studier

Fysiske aktivitetsintervensjoner sammenfattes med antagelsen om den nevrobiologiske hypotesen eller den psykososiale hypotesen (2.1.1 Den nevrobiologiske hypotesen; 2.1.2 Den psykososiale hypotesen). Studiene deles i akutt og kronisk inngripen. Kroniske intervensjoner deles videre opp etter studiens varighet. Intervensjonsvarighet har betydning ved antagelse om at fysisk aktivitet vil gi en fysisk og kognitiv/akademisk tilpasning som er tidsavhengig. Forkortelser og administrering av testinstrumenter er beskrevet i vedlegg 1 (jmf. 8.1 Vedlegg 1: Glossarium – begreper, forkortelser og administrering av testinstrumenter).

4.1.1 Akutte intervensjoner

Litteratursøket identifiserte 4 studier (St1 – St4) med akutte treningsintervensjoner for barn fra 9 – 13 år. Akutte treningsintervensjoner inkluderte Pacertest (løp mellom 2 punkter med stadig økende hastighet), jogging med moderat intensitet, og fysisk aktivitet med moderat til høy intensitet (MVPA). En studie fokuserte på forskjellen mellom elever på tredje trinn og femte trinn (St2). En studie undersøkte forskjellen mellom 1 og 2 økter (St3). En siste studie så på effekten av humanistiske fag testet med en oppmerksomhetstest, sammenlignet med en høyintensitets økt og med en blandet økt med motorisk ferdighet og fysisk aktivitet samt problemløsning (St4).

Tabell 8: Oppsummerte funn for RCT som er utført med akutt treningsøkt og som undersøker effekten på eksekutiv funksjon og kognisjon.

Studie	Utvalg	Type Treningsintervensjon I og sammenlignings- eller kontrollgruppe C	Varighet av I og C	Kognitive tester	Resultater
St1 Etnier et al. (2014) RCT	N= 43 rekr., 6 trinn 11-12 år N=19, I N=24, C	I: Akutt fysisk aktivitet i kroppsøvingstimen, 5 min oppvarming, PACER task - an aerobic fitness assessment deretter RAVLT. C: RAVLT før vanlig kroppsøving	I: 6:42 - 20:00 min, individuell varighet C: 45 min	RAVLT, minnetest, gjentatt etter 24 timer.	I: + læring av ordliste, rekapitulering av ord etter en kort pause. I = C etter 24 timer.
St2 Chen et al. (2014) RCT	N=87 anal., N= 34, 3 trinn N=17, I M _{alder} =9.24±0.44 N=17, C M _{alder} =9.12±0.33 N=53, 5 trinn N=22, I M _{alder} =11.07±0.27 N=27, C M _{alder} =11.14±0.35	I ₁ : 3 trinn, Akutt fysisk aktivitet, jogging moderat intensitet. I ₂ : 5 trinn, akutt fysisk aktivitet, jogging. C: ingen økt (vanlig skole).	I: 30 min	EFT, inhibering 2 -back, oppdatering av arbeidsminne M-OT, skifte av setting ved baseline og etter I.	I ₁ og I ₂ : + EF - inhibering, arbeidsminne, skifte av setting. I ₂ > I ₁ , unntak skifte av setting.
St3 Altenburg et al. (2016) RCT	N= 62 rekr., 56 anal., 10 -13 år N=17, I ₁ M _{alder} =11.4(0.8), I ₁ N=20, I ₂ M _{alder} =11.7(0.7), I ₂ N=19, C M _{alder} =11.6(0.9),	I ₁ : En økt fysisk aktivitet moderat intensitet, etter 90 min sitting. I ₂ : 2 økter fysisk aktivitet moderat intensitet, før og etter 90 min sitting. C: sitting med skoleoppgaver	I ₁ : 20 min, I ₂ : 20 min før + 20 min etter	SS – substest av TEA-Ch, Ved baseline, etter 20 min, 110 min, 130 min og 220 min.	I ₂ : + på oppmerksomhet I ₂ > I ₁ og C, pga. repetert fysisk aktivitet.
St4 Galotta et al. (2015) CRT	N=116 rekr., 8, 9, 10, 11 år 3, 4, 5 trinn N=31, I ₁ N=46, I ₂ N=39, C	I ₁ : Akutt fysisk anstrengelse, kroppsøvingstime, 15 min oppvarming, 30 min moderat – anstrengende aktivitet (MVPA, HR>139 slag/min), 5 min nedtrapping og strekk. I ₂ : blandet kognitiv og fysisk anstrengelse, koordinasjon og bevegelsesbasert problemløsning beslutningsoppgaver, reaksjon, tilpasset kraft, rom orientering. C: kognitiv anstrengelse, humanistisk fag.	I ₁ : 50 min I ₂ : 50 min C: 50 min	d2-a. ved baseline, etter 50 min og 100 min.	C > I ₁ > I ₂ . alle + umiddelbar og forsinket oppmerksomhetskapasitet, fysisk aktivitet fremmer oppmerksomhet.

Notis: RAVLT = Rey Auditory Verbal Learning Test. EF = Eksekutiv funksjon. EFT = modifisert Eriksen flanker test. 2-back = modifisert visual «N-back» (2 - back) task. M-OT = more-odd task (for skifte av setting). SS = Sky Search. TEA-Ch = Test of selective attention in children. D2- a = d2-test of attention.

Pacertesten (St1) hadde individuell varighet fra 6:42 min til 20 min, ellers hadde intervensjonene en totaltid mellom 20 og 50 min (St2 – St4). Signifikante kognitive resultater for eksekutiv funksjon ble påvist for inhibering, arbeidsminne og skifte av setting sammenlignet med ordinær skolehverdag - akademisk. Femtetrinns elever scoret bedre enn tredjetrinns elever. En studie viste høyere oppmerksomhets-score ved repeterte økter

sammenlignet med en økt og med ordinær skoletime. En studie viste bedre score på Pacertest sammenlignet med vanlig kroppsøving. En test viste bedre score på oppmerksomhet umiddelbart etter økt og 50 minutter senere sammenlignet med baseline. Imidlertid viste oppmerksomhetstesten best resultat etter ordinær skoletime med humanistiske fag.

4.1.2 Kroniske intervensjoner med varighet 1 uke – 3 mnd.

Litteratursøket identifiserte 4 studier (St5 -St8) med kroniske treningsintervensjoner fra 1 uke til 3 måneder for barn fra 8 – 16 år. Korte kroniske treningsintervensjoner inkluderte fysisk aktivitet i klasserommet, fysisk aktivitet og lagspill med høy intensitet, hurtigløp, hopp og jumping-jacks, styrketrening og sirkeltrening. Yoga med puste- og avslappingsteknikker ble beskrevet som kognitivt stimulerende økter (St8). En studie ble designet som en RCT – «crossover design» med 2 ukers varighet (St5).

Tabell 9: Oppsummerte funn fra RCT utført med kronisk treningsøkt (1 uke – 3 mnd.) og som undersøker effekten på eksekutiv funksjon og kognisjon.

Studie	Utvalg	Type Treningsintervensjon I og sammenlignings- eller kontrollgruppe C	Varighet av I og C	Kognitive tester	Resultater
St5 Hill et al. (2010) RCT - crossover	N=1224 rekr., 8, 9, 10, 11 år, 4, 5, 6, 7 trinn	IA: Fysisk aktivitet, lærerstyrt klasseromsprogram (strekking, aerob fysisk trening, løp på stedet, hopping til musikk) i uke1 og ingen fysisk aktivitet i uke 2. IB: Crossover av intervensjonsprogram.	2 uker IA: 75min, 15min/dag i uke 1, 0min i uke 2 IB: 0min uke1, 15min/dag i uke 2.	Daglige psykomotoriske tester, PSA, SOT, SST, DSB, DSE	IB: + på psykomotoriske tester, for fysisk aktivitet i andre uke. Varierte for type test og alder på deltagere.
St6 Schmidt et al. (2015) CRT	N=181 anal., 10 – 12 år M _{alder} = 11.35(0.6) N=69, I ₁ N=57, I ₂ N=55, C	I ₁ : lagspill, høy intensitet, krevende kognitivt. Storskala stimuli av EF. I ₂ : Aerob trening, høy intensitet, lite krevende kognitivt. Marathon. C: kontrollert situasjon, lav intensitet, lite krevende kognitivt.	6 uker I: 2økter/uke,	EF: Nsn-b, - oppdatering, C-a FT, - inhibering, MB FT, - skifte av setting. Ved baseline og etter intervensjon.	I ₁ > I ₂ = C I ₁ : 0 på EF – oppdatering. I ₁ : 0 på EF – inhibering.. I ₁ : + på EF - skifte av setting.
St7 Costigan et al. (2015; 2016) RCT	N=65 anal., M _{alder} =15.8±0.6 I ₁ : n=21 I ₂ : n=22 C: n=22	I ₁ : aerob trening, HIT i løpet av treningsøkt, hurtigløp, jumping jacks, hopping I ₂ : Aerob trening HIT og styrketrening, hurtigløp, jumping jacks, hopping og squats, push-ups. Sirkeltrening. C: ordinær kroppsøving	8 uker. I ₁ : 8-10 min, 3 økter/uke,	EF, TMT - visuell oppmerksomhet, hurtighet, prosessering, og mental fleksibilitet. FS, K10, PS-DQ ved baseline og etter intervensjon	Ikke signifikante resultater I ₁ : (+) på EF, (+) på FS - Psykologisk velvære I ₂ : (+) på EF, (+) på FS – psykologisk velvære, (+) på PS-DQ. Sign. resultater. (+) for K10, emosjonell tilstand.

St8 Telles et al. (2013) RCT	N=98 rekr., 8 – 13 år I = 49 fysisk aktivitet, 23 jenter, M _{alder} =10.5±1.3 C = 49 yoga, 15 jenter, M _{alder} =10.4±1.2	I: fysisk aktivitet, bevegelse, jogging på stedet, rask repetert bevegelse, bøy og tøy, stafettløp, lek C: Yoga, pusteteknikk, løse opp, fysisk positur, avslappingsteknikker	3 mnd. I: 45 min/dag, 5 dager/uke, C: samme tidsbruk,	EPB, fysisk fitness, SCWT, kognitiv prestasjon, BSEI, selvpfatning, LA, lydighet, akademisk prestasjon, oppmerksomhet, punktighet, atferd med venner og lærere. Ved baseline og etter 3 mnd.	I: + Selvpfatning i sosial setting og I: + kognitiv prestasjon.
---------------------------------------	---	---	--	--	---

Notis: PSA = paced serial addition. SOT = size ordering test. LS= listening span. DSB= digit-span backwards. DSE = digit-symbol encoding. NSn-b = nonspatial n-back task. C-a FT = child-adapted Flanker task. MB FT = mixed block Flanker task. EF = eksekutiv funksjon. TMT = trail making test. FS = Flourishing Scale. K10 = The Kessler Psychological Distress Scale. PS-DQ = Physical Self-Description Questionnaire. EPB = Eurofit physical battery. SCWT = Stroop color-word task for children. BSEI = Battles self-esteem inventory. LA = Lærervurdert atferd. - lydighet, akademisk prestasjon, oppmerksomhet, punktighet, atferd med venner og lærere.

En studie undersøkte forskjellen mellom høy intensitet - kognitivt krevende og høy intensitet - lite kognitivt krevende og lav intensitet – lite kognitivt krevende (St6). En studie så på effekten av fysisk aktivitet med høy intensitet og høy intensitet/styrketrening sammenlignet med ordinær kroppsøving (St7). En siste studie sammenlignet effekten av kognitivt stimulerende aktivitet - Yoga og fysisk aktivitet (St8). Intervensjonene hadde en varighet mellom 8 og 45 min per økt og total varighet mellom 2 uker og 3 måneder. Signifikante kognitive resultater for eksekutiv funksjon (skifte av setting) ble påvist ved høy intensitets kognitivt krevende lagspill sammenlignet med kontrollert situasjon (St6). Crossover studiet (St5) viste høyere score på en rekke psykomotoriske tester for gruppen som hadde intervensjonen i uke 2 (skyldes tilvenning til testene). En studie viste høyere oppmerksomhets-score ved repeterte økter sammenlignet med en økt og med ordinær skoletime. En studie (St8) viste høyere score på *selvpfatning i sosial setting* etter fysisk aktivitet sammenlignet med Yoga.

4.1.3 Kroniske intervensjoner med varighet 3 – 6 mnd.

Litteratursøket identifiserte 3 studier (St9 -St11) med kroniske treningsintervensjoner fra 3 til 6 måneder for barn fra 12 – 16 år. Kroniske treningsintervensjoner inkluderte fysisk aktivitet med høy intensitet, fysisk aktivitet i forbindelse akademisk fag, i friminutt, under transport, strukturert (lærerstyrt) fysisk aktivitet og selvvalgt (elevstyrt) fysisk aktivitet (St11). En studie sammenlignet dobbel treningsfrekvens moderat intensitet- og høy intensitets fysisk aktivitet med ordinær kroppsøving (St9). En studie undersøkte forskjellen mellom fysisk aktivitet

under akademiske fag og ordinær kroppsøving (St10). En siste studie (St11) sammenlignet effekten av strukturert fysisk aktivitet og ustrukturert fysisk aktivitet.

Tabell 10: Oppsummerte funn fra RCT utført med kronisk treningsøkt (3 – 6 mnd.) og som undersøker effekten på eksekutiv funksjon og kognisjon.

Studie	Utvalg	Type Treningsintervensjon I og sammenlignings- eller kontrollgruppe C	Varighet for I og C	Kognitive tester	Resultater
St9 Ardoy et al. (2014) RCT	N=67 rekr., 54 anal. 12–14år M _{alder} =13.0(0.1) G= 43 J=24 n =18, I ₁ n=26, I ₂ n=23, C	I ₁ : Fysisk aktivitet, 2* kroppsøving med samme mål, innhold, og pedagogisk strategi som i kroppsøving. I ₂ : fysisk aktivitet, høy intensitet > 120 BPM, oppvarming, økende intensitet ved motivering og fysisk utfordring. C: vanlig kroppsøving	4 mnd. I ₁ : 4 økter/uke I ₂ : 4 økter/uke C: 2 økter/uke	IGF-M - ikke verbal og verbal evne, abstrakt resonnement, spatial evne, verbalt resonnement, numerisk evne. AP, skolekarakterer.	I ₂ : + for IGF-M alle disipliner, unntak verbalt resonnement, og AP. I ₂ > I ₁ . I ₁ og C: ingen forskjell
St10 Tarp et al. (2016), Bugge et al. (2014) CRT	N=632 anal., 12-14 år, 6 og 7 trinn M _{alder} = 12.9 (0.6) N=194, I N=438, C	I: fysisk aktivitet under akademiske fag, Fysisk aktivitet i friminutt, Aktiv skoletransport, C: vanlig kroppsøving	20 uker 60 min/dag	EF, EFT - måler nøyaktighet og reaksjonstid. AP, Matematikk	I: 0 på EF. I:0 på AP, matematikk. C + på EFT reaksjonstid.
St11 Subramanian et al. (2015) RCT	N=347 anal. (79% fullførte), 14 (12–16) median (min-maks) Strukturert gr.: n=172 n=36, I ₁ trente g=19, j=17 n=136, I ₂ utrente g=77, j=59 Ustrukturert gr.: n=175 n=36, trente g=19, j=17 n=139, utrente g=75, j=64.	I ₁ : strukturert fysisk aktivitet for trente, anstrengende fysisk aktivitet, styrke av muskel og beinbygning, strekking. + treningsprog. fra egen idrett. I ₂ : strukturert fysisk aktivitet for utrente. samme treningsfrekvens som trente. I ₃ : ustrukturert trente, selvvalgt aktivitet, fotball, basket, hopping, hagearbeid. + treningsprogram fra egen idrett. I ₄ : ustrukturert fysisk aktivitet, utrente, selvvalgt aktivitet i samme tidsrom.	6 mnd. I ₁ : 60 min, 3 økter/uke, + egen idrett. I ₂ : 2*60 min, 6 økt/uke, I ₃ : 60 min, 3 økter/uke, + egen idrett. I ₄ : 2* 60 min, 6 økt/uke	LCT, TTA og TTB, RFFT	I ₁ og I ₂ : + på LCT, TTB, RFFT i forhold til I ₃ og I ₄ . I ₁ og I ₂ : strukturert trening bedre pga. kognitivt innhold i intervensjon.

Notis: BPM = beats per minute, IGF-M = Spanish Overall and Fractional Intelligence Test. AP = akademisk prestasjon. EF = eksekutiv funksjon. EFT = Eriksen flanker test. LCT = Two target letter cancelation test. TTA og TTB = Trail Making Test A and B. RFFT = Ruff figural Fluency test.

Øktene hadde en varighet på 60 min med 2 – 5 økter /uke i 4 til 6 måneder. Signifikant kognitive resultater for eksekutiv funksjon (reaksjonstid) ble påvist ved ordinær kroppsøving sammenlignet fysisk aktivitet under akademiske fag (St10). En studie viste høyere intelligensscore ved dobbel treningsfrekvens ved høy intensitet sammenlignet med ordinær kroppsøving

(St9). En studie viste høyere score på kognitiv funksjon etter strukturert fysisk aktivitet sammenlignet med ustrukturert fysisk aktivitet (St11).

4.1.4 Kroniske intervensjoner med varighet 6 – 12 mnd.

Litteratursøket identifiserte 3 studier (St12 -St14) med treningsintervensjoner fra 6 til 12 måneder for barn fra 7 – 10 år. Kroniske treningsintervensjoner inkluderte fysisk aktivitet med ASK (Active Smarter Kids) med løping, stafett, hinderløype og aktiv lek, fysisk aktivitet i akademisk klasse med jumping jacks og problemløsning, og klasseromsintegret fysisk aktivitet. En studie (St12) sammenlignet ASK undervisningsleksjoner med ordinær kroppsøving. En studie (St13) undersøkte forskjellen mellom fysisk aktivitet under akademiske fag og ordinær akademisk undervisning. En siste studie (St14) sammenlignet effekten av klasseromsintegret fysisk aktivitet med ordinær matematikkundervisning.

Tabell 11: Oppsummerte funn fra RCT utført med kronisk treningsøkt (6 – 12 mnd.) og som undersøker effekten på eksekutiv funksjon og kognisjon.

Studie	Utvalg	Type Treningsintervensjon I og sammenlignings- eller kontrollgruppe C	Varighet av I og C	Kognitive tester	Resultater
St12 Resaland et al. (2015), (2016), Aadland et al. (2017) CRT	N=1129 rekr., 5 trinn, 10 år. N=620, I N=582, C	I: fysisk aktivitet + ASK ASK: fysisk aktiv undervisningsleksjon, (løping, stafett, hinderløype og aktiv lek). aktive friminutt med fysisk aktivitet, hjemmelekse fysisk aktivitet. C: kroppsøving + fysisk aktivitet.	7 mnd. I: tot. 300min/uke ASK: 165 min/uke 3*30 min/uke + 5 min/ dag + 10 min/dag, C: 135 min/uke,	AP, EF, SCWT, inhibering VF, kognitiv fleksibilitet TMT, kognitiv fleksibilitet WISC-IV, arbeidsminne	I: 0 på AP - lesing, regning, engelsk, nasjonale prøver (UDIR). Analyse per protokoll: I: + på EF sammensatt. I: + VF, TMT, kognitiv fleksibilitet Høyest progresjon for deltagere med lav baseline.
St13 Fedewa et al. (2015) RCT	N = 460 rekr., 3, 4, 5 trinn N=156, I N=304, C	I: fysisk aktivitet i akademisk klasse, jumping jacks etter matematikk oppgave eller plukke opp kort som er spredt rundt i klasserommet etter instruksjon med farge, tall osv. C: vanlig undervisning.	8 mnd. I: 20 min/dag, 5 dager /uke, C: 0	Ravens`s SPM, IQ test MAP, nasjonal standardisert test ved baseline og etter 8 mnd.	I: 0, SPM - intelligens I og C: +, for MAP matematikk, lesing.
St14 Have et al. (2016) Have et al. (2016) upubl. CRT	N=505 rekr., 1 trinn, M _{alder} = 7.2 ±0.3 N=294, I N=211, C	I: Klasseromsintegret fysisk aktivitet, aktiv matte. Enhver kroppsbevegelse, minst 15 min per time. C: matematikk	9 mnd. Minst 15min/ time, 6 *45 min/uke,	MG, EF - EFT modifisert, BRIEF – kartlegging av EF TTCT - kreativitet,	I: +, MG. I: 0, EF.

Notis: MG = 45-minute standardized math test. EF = Eksekutiv funksjon. EFT = Erickson Flanker task. SCWT = Stroop color-word task for children. VF = Verbal fluency, TMT = The Trail making test. WISC-IV = Wechsler Intelligence Scale for Children, 4th ed. SPM = Raven`s Standard Progressive Matrics. MAP = Measures of Academic Progress. EFT = Erikson flanker test. BRIEF = behavior rating inventory of executive function. TTCT = Torrance tests of creative Thinking.

Øktene hadde en varighet på 15 – 30 min med 3 – 6 økter /uke i 7 til 9 måneder. En studie (St12) viste signifikant kognitive resultater for eksekutiv funksjon. En studie (St13) viste høyere score på akademisk prestasjon ved klasseromsintegreert fysisk aktivitet, men resultatene var tilsvarende også for vanlig undervisning. En sist studie (St14) viste bedret akademisk prestasjon, men ingen progresjon i eksekutiv funksjon ved sammenligning av fysisk aktiv matematikk og ordinær matematikk undervisning.

4.1.5 Kroniske intervensjoner med varighet 12 – 24 mnd.

Litteratursøket identifiserte 3 studier (St15 -St17) med treningsintervensjoner fra 12 til 24 måneder for barn fra 8 – 13 år. Lange kroniske treningsintervensjoner inkluderte skolebasert fysisk aktivitet med hopping, dansing, enkel styrketrening, moderat til anstrengende fysisk aktivitet i akademisk undervisning, høy intensitetstrening, fysiske tester og selvvalgt organisert idrett. En studie sammenlignet skolebasert fysisk aktivitet og vanlig kroppsøving (St15). En studie undersøkte forskjellen mellom integrert fysisk aktivitet i problemløsning og ordinær klasseromsundervisning (St16). En siste studie sammenlignet effekten av ulike intervensjoner med ordinær kroppsøving (St17).

Tabell 12: Oppsummerte funn fra RCT utført med kronisk treningsøkt (12 – 24 mnd.) og som undersøker effekten på eksekutiv funksjon og kognisjon.

Studie	Utvalg	Type Treningsintervensjon I og sammenlignings- eller kontrollgruppe C	Varighet av I og C	Kognitive tester	Resultater
St15 Ahamed et al. (2007) CRT	N=396 rekr., 288 anal., M _{alder} =10.2(0.6) n = 99/115, I (4/5 trinn) n = 34/39, C (4/5 trinn)	I: skolebasert fysisk aktivitet: hoppe, danse, styrketrening med hånd- gripere og treningsbånd. C: vanlig kroppsøving	16 mnd. I: 139 min/uke, C: 92 min/uke.	AP: CAT-3, matematikk, lesing og språk	I: 0 på CAT-3
St16 Mullender-Wijnsma et al. (2016), De Greff et al. (2016) CRT	N = 499 rekr., 2 og 3 trinn M _{alder} =8.0(0.72), I M _{alder} =8.2(0.74), C I=249 C=250	I: F & V lesson, fysisk aktiv (moderat til anstrengende intensitet) akademisk undervisning, matematikk og staving, eks. 8 hopp på stedet for å løse multiplikasjon "2 × 4". C: vanlig klasseromsundervisning	22 + 22 uker I: 20 -30 min, 3 leksjoner /uke, over 2 år,	matematikk - hurtighet og ferdighet språk - lesing og staving. ved baseline, etter første og andre år.	I: + for matematikk – hurtighet og ferdighet I: + for språk, staving. Fysisk aktivitet ga 4 mnd. mer læring enn for C.

St17 Quinto Romani and Klausen (2017) CRT	N=1157 rekr., 925 anal., 6 trinn, 11-13 år N=283, I ₁ N=201, I ₂ N=210, I ₃ N=463, C	Prosjekt 3A 2008 - 2010: 4 ulike intervensjoner: I ₁ : høy intensitets trening, 75-85% av HF maks I ₂ : fysiske tester for potensiell atferdsendring. I ₃ : selvalgt organisert idrett, moderat intensitet. C: kroppsøving	24 mnd. I ₁ : 20min, 2økter/uke, I ₂ : 30-45min /mnd. I ₃ : 90 – 135 min.	GPA i 9 trinn. 2011/12, deltagere GPA i 9 trinn. 2008, generelle elever Ved baseline og etter 24 mnd.	I: 0 for GPA. I ₁ : + for GPA - språk karakterer (spesielt baseline inaktive). I ₂ : 0 I ₃ : 0 eksaminasjonskarakter og lese- karakter.
--	--	---	---	--	--

Notis: F & V lessons = physically active math and language lessons. AP = Akademisk prestasjon. CAT-3 = Canadian Achievement test. EF = eksekutiv funksjon. GPA= grade point average.

Øktene hadde en varighet på 20 – 45 min med 1 – 3 økter /uke i 16 til 24 måneder.

Signifikante kognitive resultater for eksekutiv funksjon ble ikke målt. En studie viste høyere score på akademisk prestasjon ved klasseromsintegreert fysisk aktivitet sammenlignet med vanlig kroppsøving. En studie viste høyere score på akademisk prestasjon sammenlignet med ordinær klasseromsundervisning. En studie viste høyere score på akademisk prestasjon for høyintensitetstrening sammenlignet med kroppsøving (St17).

4.2 Koding av inkluderte studier

I Metodekapittelet (jmf. 3.4.2 Metode for koding av inkluderte studier) ble prosedyren for koding av inkluderte studier forklart. Kodene ble satt sammen etter sammenligning av intervensjonsgrupper og kontrollgrupper. Bakgrunnen for denne prosedyren er en ide om å få frem intervensjonens reelle funksjon i forhold til kontrollgruppe også med tanke på at deltagerne til daglig fungerer i en skolehverdag med ordinær undervisning og kroppsøving. Tabellen under viser kodene for inkluderte akutt studier.

Tabell 13: Viser kodene for inkluderte akuttstudier.

Studie	Kode for intervensjonsulikheter – akutt intervensjoner.	Type Treningsintervensjon I og sammenlignings- eller kontrollgruppe C
St1 Etnier et al. (2014) RCT	Forsterket intensitet vs. Akademisk Forsterket intensitet vs. Akademisk (etter 24 timer)	I: Akutt fysisk aktivitet i kroppsøvingstimen, 5 min oppvarming, PACER task - an aerobic fitness assessment deretter RAVLT. C: RAVLT før vanlig kroppsøving
St2 Chen et al. (2014) RCT	Forsterket frekvens vs. Akademisk Forsterket frekvens vs. Akademisk Alder (5 trinn vs. 3 trinn)	I1: 3 trinn, Akutt fysisk aktivitet, jogging moderat intensitet. I2: 5 trinn, akutt fysisk aktivitet, jogging. C: ingen økt (vanlig skole).
St3 Altenburg et al. (2016) RCT	Forsterket frekvens vs. Akademisk Forsterket frekvens *2 vs. Akademisk	I1: En økt fysisk aktivitet moderat intensitet, etter 90 min sitting. I2: 2 øker fysisk aktivitet moderat intensitet, før og etter 90 min sitting. C: sitting med skoleoppgaver
St4 Galotta et al. (2015) CRT	Forsterket intensitet vs. Akademisk Kognitivt beriket vs. Akademisk Forsterket intensitet vs. Kognitivt beriket	I1: Akutt fysisk anstrengelse, kroppsøvingstid, 15 min oppvarming, 30 min moderat – anstrengende aktivitet (MVPA, HR>139 slag/min), 5 min nedtrapping og strekk. I2: blandet kognitiv og fysisk anstrengelse, koordinasjon og bevegelsesbasert problemløsning beslutningsoppgaver, reaksjon, tilpasset kraft, rom orientering. C: kognitiv anstrengelse, humanistisk fag.

Eksempel, Etnier, Labban, Piepmeier, Davis & Henning (2014) undersøkte effekten av en intervensjon med økt intensitet. Koden «forsterket intensitet vs. Akademisk» uttrykker at deltagerne fikk et tillegg i intensitet ved intervensjonen. Undersøkelsen ble gjentatt etter 24 timer (Koden: Forsterket intensitet vs. Akademisk etter 24 timer). Tabell 14 viser koding for inkluderte kroniske studier.

Tabell 14: viser kodene for inkluderte kroniske studier.

Studie	Kode for intervensjonsulikheter – kronisk intervensjoner.	Type Treningsintervensjon I og sammenlignings- eller kontrollgruppe C
St5 Hill et al. (2010) RCT -crossover	Kognitivt beriket vs. Akademisk	IA: Fysisk aktivitet, lærerstyrt klasseromsprogram (strekking, aerob fysisk trening, løp på stedet, hopping til musikk) i uke 1 og ingen fysisk aktivitet i uke 2. IB: Crossover av intervensjonsprogram.
St6 Schmidt et al. (2015) CRT	Kognitivt beriket vs. Akademisk Forsterket intensitet vs. Akademisk Forsterket intensitet vs. kognitivt beriket	I1: lagspill, høy intensitet, krevende kognitivt. Storskala stimuli av EF. I2: Aerob trening, høy intensitet, lite krevende kognitivt. Marathon. C: kontrollert situasjon, lav intensitet, lite krevende kognitivt.
St7 Costigan et al. (2015; 2016) RCT	Forsterket intensitet vs. Kroppsøving Forsterket intensitet vs. Kroppsøving Type aerob vs. Aerob + styrketrening	I1: aerob trening, HIT i løpet av treningsøkt, hurtigløp, jumping jacks, hopping I2: Aerob trening HIT og styrketrening, hurtigløp, jumping jacks, hopping og squats, push-ups. Sirkeltrening. C: ordinær kroppsøving
St8 Telles et al. (2013) RCT	Forsterket frekvens vs. Kognitivt beriket	I: fysisk aktivitet, bevegelighet, jogging på stedet, rask repetert bevegelse, bøy og tøy, stafettløp, lek C: Yoga, pusteteknikk, løse opp, fysisk positur, avslappingsteknikker
St9 Ardoy et al. (2014) RCT	Forsterket frekvens vs. Kroppsøving Forsterket intensitet vs. Kroppsøving Forsterket intensitet vs. Forsterket frekvens	I1: Fysisk aktivitet, 2* kroppsøving med samme mål, innhold, og pedagogisk strategi som i kroppsøving. I2: fysisk aktivitet, høy intensitet > 120 BPM, oppvarming, økende intensitet ved motivering og fysisk utfordring.

		C: vanlig kroppsøving
St10 Tarp et al. (2016), Bugge et al. (2014) CRT	Kognitivt beriket vs. Kroppsøving	I: fysisk aktivitet under akademiske fag, Fysisk aktivitet i friminutt, Aktiv skoletransport, C: vanlig kroppsøving
St11 Subramanian et al. (2015) RCT	Strukturert vs. Ustrukturert	I1: strukturert fysisk aktivitet for trente, anstrengende fysisk aktivitet, styrke av muskel og beinbygning, strekking. + treningsprog. fra egen idrett. I2: strukturert fysisk aktivitet for utrente. samme treningsfrekvens som trente. I3: ustrukturert trente, selvvalgt aktivitet, fotball, basket, hopping, hagearbeid. + treningsprogram fra egen idrett. I4: ustrukturert fysisk aktivitet, utrente, selvvalgt aktivitet i samme tidsrom.
St12 Resaland et al. (2015), (2016), Aadland et al. (2017) CRT	Kognitivt beriket vs. Kroppsøving	I: fysisk aktivitet + ASK ASK: fysisk aktiv undervisningsleksjon, (løping, stafett, hinderløype og aktiv lek). aktive friminutt med fysisk aktivitet, hjemmeleke fysisk aktivitet. C: kroppsøving + fysisk aktivitet.
St13 Fedewa et al. (2015) RCT	Kognitivt beriket vs. akademisk	I: fysisk aktivitet i akademisk klasse, jumping jacks etter matematikk oppgave eller plukke opp kort som er spredt rundt i klasserommet etter instruksjon med farge, tall osv. C: vanlig undervisning.
St14 Have et al. (2016) Have et al. (2016) upubl. CRT	Kognitivt beriket vs. akademisk	I: Klasseromsintegret fysisk aktivitet, aktiv matte. Enhver kroppsbevegelse, minst 15 min per time. C: matematikk
St15 Ahamed et al. (2007) CRT	Forsterket frekvens vs. kroppsøving	I: skolebasert fysisk aktivitet: hoppe, danse, styrketrening med hånd- gripere og trenings- bånd. C: vanlig kroppsøving
St16 Mullender- Wijnsma et al. (2016), De Greff et al. (2016) CRT	Kognitivt beriket vs. akademisk	I: F & V lesson, fysisk aktiv (moderat til anstrengende intensitet) akademisk undervisning, matematikk og staving, eks. 8 hopp på stedet for å løse multiplikasjon "2 x 4". C: vanlig klasseromsundervisning
St17 Quinto Romani and Klausen (2017) CRT	Forsterket intensitet vs. Kroppsøving Ustrukturert vs. Kroppsøving Fysisk tester for atferdsendring	Prosjekt 3A 2008 - 2010: 4 ulike intervensjoner: I1: høy intensitets trening, 75-85% av HF maks I2: fysiske tester for potensiell atferdsendring. I3: selvvalgt organisert idrett, moderat intensitet. C: kroppsøving

Kodene er resultat av en kategorisering og brukes videre i synteseplanet hvor data-tilgjengeligheten bedres. Fysisk aktivitetsintervensjoner innføres som et tillegg til elevenes timeplan. Eksempel, Ahamed et al. (2007) undersøkte en treningsintervensjon som tillegg til ordinær skole. Koden «forsterket frekvens versus kroppsøving», uttrykker at deltagerne fikk et tillegg i mengde og ikke intensitet.

4.3 Metodologisk kvalitet

Et sentralt element var å undersøke den metodologiske kvaliteten i inkluderte studier. I denne undersøkelsen blir forskningsdesign delt slik det fremgår i primærstudiene. Studier som er rapportert som RCT blir vurdert etter en RCT standard. Studier som blir beskrevet som CRT blir vurdert etter en CRT standard.

4.3.1 Metodologisk kvalitet for inkluderte RCT

En testgjennomgang av sjekklisten ble utført for RCT studier etter Maher (Maher et al., 2003). Resultatene fra kvalitetsvurderingen er summen av 11 kriterier formulert som kontrollspørsmål. Det ble gitt 1 poeng for «ja» besvarelse av spørsmål, 0 poeng for «nei» eller «uvisst/ukjent». Kvalitetsscoren presenteres som en prosentvis andel av totalscore per studie og per kontrollspørsmål. Kvaliteten for RCT varierer mellom 54.5% (lav metodologisk kvalitet) og 90.9% (tilstrekkelig metodologisk kvalitet). Syv av ni studier blir vurdert å ha tilstrekkelig metodologisk kvalitet (60% eller høyere). Studiene er ordnet etter oppnådd metodologisk kvalitet (se tabellen under).

Tabell 15: Kvalitetsvurdering av randomiserte kontrollerte eksperimenter (RCT). Total score per kontrollspørsmål og total score per studie (%) (St1, St2, St3, St5, St7, St8, St9, St11 og St13).

	poeng ja=1, nei og ukjent=0.	St8	St7	St1	St2	St3	St5	St11	St9	St13	Total score
1.	Er Eligerings- (utvelgelse) kriterier spesifisert? Populasjonen som er kilde til deltagerne og eligeringskriterier for deltagelse rapporteres i studien.	1	1	1	1	1	1	1	1	1	9/9 100%
2.	Er deltagere tilfeldig allokert (tildelt) grupper? Det kreves at rapporten uttrykker at allokeringen var tilfeldig (randomisert). Detaljene i allokeringen behøver ikke være beskrevet.	1	1	1	1	1	0	1	0	1	6/9 66.7%
3.	Er allokeringen skjult? Den skal være skjult for allokeringspersonell (blindet) mens allokering foregår.	1	1	1	1	1	1	1	1	0	8/9 88.8%
4.	Er de fleste målingene med betydning for utfall, like eller lignende ved baseline? Studien må rapportere minst en nøkkelvariabel (variabel som har effekt/manglende effekt på utfall) som har betydning for utfall ved baseline. Vesentlig er at måleresultatene ikke endres som følge av ulike baselineverdier. Det er tilstrekkelig at baselineverdier kun presenteres for deltagere som fullfører studien.	1	1	1	1	1	0	1	1	1	8/9 88.9%
5.	Er deltagerne blindet for behandlingen? Deltagerne skal ikke ha kjennskap om hvilke grupper de er allokert til, og kan sannsynligvis ikke skille mellom intervensjon og kontrollaktivitet.	0	0	0	0	0	0	0	0	0	0/9 0%
6.	Er forskningspersonell blindet for deltageridentifisering i grupper og behandling som avgis? Har forskningspersonell	1	0	0	0	0	1	0	0	0	2/9 22.2%

	(treningslærer) kjennskap til hvilken treningsintervensjon som blir gitt?										
7.	<i>Er forskningspersonell som driver datainnsamling blindet for deltageridentifisering i grupper.</i> Eks. ved selvrappotering har ikke forskningspersonell kjennskap til hvem som har besvart testen.	1	1	0	0	0	0	0	0	0	2/9 22.2%
8.	<i>Var drop out mer enn 15%?</i> Kriteriet er oppfylt dersom minstekravet ved baseline og oppfølgingsmålinger er minst 85% deltagelse.	1	1	1	1	1	1	0	0	0	6/9 66.7%
9.	<i>Fikk alle deltagerne behandling og kontroll som opprinnelig planlagt?</i> Kriteriet er tilfredsstillt dersom rapporten uttrykker at alle deltagerne fikk behandling og kontroll som allokert. Eks. data kan analyseres etter opprinnelig plan («Intention to treat») dersom mengden trening ikke er som planlagt. Analyse basert på intention to treat, betyr at deltagerne blir analysert som om de får planlagt behandling (intervensjon og kontroll).	1	1	1	0	0	1	1	1	1	7/9 77.8%
10	<i>Er resultatene fra den statistiske sammenligningen av gruppene rapportert for minst en sentral utgangsvariabel.</i> Statistisk sammenligning av 2 eller flere grupper, kan bety sammenligning mellom 2 behandlinger, eller mellom behandling og sammenligningsgruppe med en kontrollallokering.	1	1	1	1	1	1	1	1	1	9/9 100%
11	<i>Presenterer studien målinger for sentraltendens og spredningsmål for minst en sentral utgangsvariabel?</i> Sentraltendensen skal gi et mål for effekten av behandlingen. Behandlingseffekten kan oppgis som differansen mellom grupper eller utfall for hver gruppe. Spredningsmål inkluderer standardavvik, standardfeil, konfidensintervall, interkvartil rekkevidde (eller annet kvartil range). Presentasjon kan gjøres grafisk eller med søylediagram så lenge innholdet er tydelig. For kategoridata må antallet deltagere presenteres for hver kategori i hver gruppe.	1	1	1	1	1	1	1	1	1	9/9 100%
	% score av totalscore (11)	90.9	81.8	72.7	63.6	63.6	63.6	63.6	54.5	54.5	

Ifølge Maher et al. (2003) er sjekklisten reliabilitetstestet. Ulikt testpersonell har i 94% av kontrollspørsmålene gitt en poengscore som samsvarer innen 2 poeng eller mindre (enige eller avviker med opptil 2 poeng). Av kontrollspørsmål 3 fremgår at 88.8% av RCT studiene ble gjennomført med skjult allokering. Videre at studie 3 (St3) har 63.6% av kontrollkriteriene tilfredsstillt, men mangler blinding av deltagerne, treningslærer og forskningspersonell som driver datainnsamling. Kontrollspørsmål 5, 6, og 7 som omhandler blinding får hhv 0, 22.2 og 22.2 % altså ingen eller lav kvalitetsscore. Manglende blinding illustrerer en prinsipiell forskjell mellom kliniske studier (medikamentell behandling vs placebo) og fysisk aktivitetsintervensjoner (fysisk aktivitet vs kontrollaktivitet).

4.3.2 Metodologisk kvalitet for inkluderte CRT

Kluster randomiserte eksperimenter (Campbell et al., 2004) er designmessig mer kompliserte enn individuelt randomiserte eksperimenter. Det kreves flere deltagere for å oppnå samme statistiske styrke (power). Analysen er mer komplisert. Kluster randomiserte eksperimenter krever at man tar hensyn til korrelasjonen mellom deltagere i hver kluster (individene i en klasse er ikke uavhengige). *Effektiv utvalgsstørrelse ved klusterdesign* er lavere enn total utvalgsstørrelse ved individuell randomisering. Reduksjonen i effektiv utvalgsstørrelse er avhengig av gjennomsnittlig klusterstørrelse og graden av korrelasjon mellom individer i klusterne, kjent som intrakluster- (intra-klasse) korrelasjonskoeffisient (ρ). Intrakluster-korrelasjonskoeffisient er andelen av total varians i utfall som kan forklares med forskjeller mellom ulike kluster. Power beregnes ved å multiplisere individuell utvalgsstørrelse n , med designeffekt $(1 + (m - 1) \cdot \rho)$, der m er gjennomsnittlig klusterstørrelse og ρ er intrakluster-korrelasjonskoeffisient (Campbell et al., 2004).

Kluster randomiseres vanligvis alle på en gang i motsetning til RCT hvor individene trekkes tilfeldig en av gangen (Campbell et al., 2004). Etter gjennomført randomisering må hver enkelt samtykke til delttagelse. Dersom individer ikke samtykker, åpner muligheten for seleksjonsavvik (skjevhet i antall drop out fra hver gruppe, med fare for redusert internvaliditet). Dersom deltagerne blir inkludert uten å bli spurt om samtykke får forskerteamet et etisk dilemma.

En rekke undersøkelser har påvist mangler ved gjennomføringen av kluster randomiserte eksperimenter. I følge Campell (Campbell et al., 2004) har 2 reviews avdekket at kun 19% av klusteregenskapene ved designet, ble tatt hensyn til under planleggingen av forsøket. Et Review har videre avdekket at 70% av klusterdesign ikke tilfredsstillt kravene til klasseinndeling i CRT. Legitimering eller rasjonale (Donner i Campbell et al., 2004) for bruken av kluster design fremkom i 4 av 16 undersøkte CRT. Beregning av effektiv utvalgsstørrelse fremgikk i 3 av 16 CRT. I artikkelens analyse tok 8 av 16 studier hensyn til egenskapene ved klusterdesign (Isaakidis & Ionnaidis i Campbell et al., 2004).

Metodologisk kvalitet virker å være et område som er i positiv utvikling. I denne rapporten avdekkes at 62.5% av undersøkte studier har beregnet effektiv utvalgsstørrelse. Videre at 62.5% av studiene har beregnet behandlingseffekt med usikkerhet (95% konfidensintervall). Studiene har ikke eksplisitt rapportert et rasjonale (kun 12.5%) for bruken av kluster

randomisert eksperiment som forskningsdesign. Kvaliteten for CRT varierer mellom 59.0% (lav metodologisk kvalitet) og 86.4% (tilstrekkelig metodologisk kvalitet). Tilstrekkelig metodologisk kvalitet (60% eller høyere) ble funnet i 6 av 8 studier. Studiene ble ordnet etter oppnådd metodologisk kvalitet. I hvilken grad kriteriene oppfylles ble vurdert skjønnsmessig og individuelt.

Tabell 16: Kvalitetsvurdering av kluster randomiserte eksperimenter (CRT). Total score per kontrollspørsmål og total score per studie (%) (St4, St6, St10, St12, St14, St15, St16, St17).

	poeng ja=1, nei og ukjent=0.	St12	St16	St10	St6	St14	St4	St15	St17	Total score
	Tittel eller sammendrag									
1.	<i>Er forskningsdesign beskrevet? Hvordan er deltagere allokert til intervensjonen (randomisering). Kriteriet tilfredsstilles dersom studien spesifiserer at allokering er basert på klustere?</i>	1	1	1	1	1	1	1	1	8/8 100%
	Introduksjon									
2..	<i>Legitimeres kluster randomiseringen? Kriteriet er tilfredsstillt hvis det er gitt en vitenskapelig begrunnelse for hvorfor studien bruker kluster randomisering.</i>	0	0	1	0	0	0	0	0	1/8 12.5%
	Metode									
3.	<i>Spesifiseres eligerings- (utvelgelse) kriterier av deltagere og klustere, og fra hvilken setting og lokalisering foregår datainnsamlingen?</i>	1	1	1	1	1	1	1	1	8/8 100%
4.	<i>Har planlagt forskningsforsøk en beskrivelse av intervensjonen for hver gruppe, med en spesifisering av om intervensjon er tilsiktet enkelt deltager, kluster eller begge, og hvordan og når intervensjonen blir gjennomført?</i>	1	1	1	1	1	1	1	1	8/8 100%
5.	<i>Inneholder hypotesen/hensikten med undersøkelsen også en spesifisering av om den er tilsiktet individ- eller klusternivå eller begge?</i>	1	1	1	1	0	1	1	0	6/8 75%
6.	<i>Er resultatene tydelig beskrevet med definert primær- og sekundærutfall? Det må spesifiseres om resultatene gjelder individ-, klusternivå eller begge, og hvilke metoder som kan bedre kvaliteten i målingene (eks. multiple målinger, forberedende øvelser for observatører).</i>	1	1	1	1	1	1	1	1	8/8 100%
7.	<i>Blir effektiv utvalgsstørrelse for klusterdesign beregnet? (inkludert beregningsmåte, antallet klustere, klusterstørrelse, intrakluster-korrelasjonskoeffisient (ICC, eller k) med en indikasjon på usikkerhet i beregningen). Forklares foreløpige analyser?</i>	1	1	1	1	1	0	0	0	5/8 62.5%
	Randomisering									
8.	<i>Hvordan gjennomføres allokeringsssekkefølge (tildeling), inkludert blokkinnndeling, stratifisering og matching?</i>	1	1	1	1	1	1	1	1	8/8 100%
9.	<i>Gjennomføres blinding under allokering inntil intervensjonsgruppen er bestemt?</i>	1	1	1	1	1	1	1	1	8/8 100%
10	<i>Blir det spesifisert hvem som genererer allokeringsssekkefølgen, hvem som fordeler deltagerne, hvem som melder på deltagere?</i>	1	1	0	0	1	1	1	1	6/8 75%

11.	<i>Blir deltagere, treningslærere og forskningspersonell som vurderer måleresultater blindet for behandling og måleresultater i forhold til fordelte deltagere? Var blindingen vellykket?</i>	0	0	0	0	0	0	0	0	0/8 0%
12.	<i>Beskrives de statistiske metodene for å sammenligne grupper med primærresultater? Blir det i analysen tatt hensyn til beregning av effektiv utvalgsstørrelse for kluster? Beskrives delanalyser, som subgruppeanalyse og justert analyse?</i>	1	1	1	1	1	0	0	0	5/8 62.5%
	Resultater									
13	<i>Blir flyten av klustere og deltagere presentert? Et diagram anbefales som forklarer antallet klustere, og deltagere, som tilfeldig allokeres, mottar behandling og fullfører intervensjon. Blir primærutfall analysert med beskrivelse av avvik fra planlagt protokoll med begrunnelse?</i>	1	1	1	0	1	0	1	1	6/8 75%
14	<i>Beskrives datoer som definerer rekrutteringsperiode og oppfølging?</i>	1	1	1	1	1	1	1	1	8/8 100%
15	<i>Beskrives baseline informasjon for hver gruppe på individ- og klusternivå?</i>	1	1	1	1	1	1	1	1	8/8 100%
16	<i>Inkluderes antallet klustere og individer i hver gruppe som analyseres? Presenteres analysen med absolutte tall (10/20 ikke 50%)? Og baseres analysen på opprinnelig plan «Intention to treat»?</i>	0	0	0	0	0	0	0	0	0/8 0%
17	<i>Presenteres en resultatoppsummering for hvert primær- og sekundærresultat med beskrivelse av gruppe på individ- og kluster nivå? Rapporteres beregnet effektstørrelsen med usikkerhet (95% konfidensintervall) og intrakorrelasjons- koeffisient (ICC eller k) for hvert primærresultat?</i>	1	1	0	1	1	1	0	0	5/8 62.5%
18	<i>Adresseres samsvar med andre analyser, inkludert delanalyser og justerings analyser, både planlagte og utforskende analyser?</i>	1	1	1	1	1	1	0	0	6/8 75%
19	<i>Rapporteres alle viktige motsetninger i resultatene eller bivirkning av intervensjonen?</i>	1	1	1	1	1	1	1	1	8/8 100%
	Diskusjon									
20	<i>Rapporteres en tolkning av resultatene? Deriblant hypotesen i studien, potensielle feilkilder, unøyaktighet eller farene ved at analyser og resultater forekommer flere ganger.</i>	1	1	1	1	1	1	0	0	6/8 75%
21	<i>Blir (relevante) eksperimentelle funn beskrevet i forhold til generaliserbarhet (ekstern validitet), for individer og klustere?</i>	1	1	1	1	0	0	1	1	6/8 75%
22	<i>Blir resultatene gjenstand for generell tolkning i forbindelse med evidensen gitt av studien?</i>	1	1	1	0	0	0	0	1	4/8 50%
	% score av totalscore (22)	86.4	86.4	81.8	72.7	72.7	63.3	59.0	59.0	

Av kontrollspørsmål 9 fremgår at skjult allokering er gjennomført i 100% av studiene. Studie 12 (St12) har høyest metodologisk kvalitet (86.4%), men mangler blinding av deltagere og treningslærer som gjennomfører øktene. Metodologiske problemstillinger utdypes i diskusjonskapittelet (jmf. 5.1 Vurdering av validitet).

4.4 Syntese av inkluderte studier – et narrativ

Syntesen er inndelt i akutte og kroniske fysisk aktivitetsintervensjoner, og viser en kompilering av deskriptive data fra hovedfunn i primærstudier. Informasjon fra Harvestplot (jmf. 3.6.1 Presentasjon av syntese) besvarer en rekke spørsmål når det gjelder sammenligning av ulike intervensjoner og kontrollgrupper. Syntesen skiller ikke mellom RCT og CRT design. En forbindelse til primærstudiene eksisterer via koding (tabell 13 –14) og beskrivelse av studier (tabell 8 -12) og via kvalitetsvurderinger av enkeltstudier (tabell 15 - 16).

4.4.1 Akutte fysiske aktivitetsintervensjoner

Ved syntese av intervensjonsulikhet for akutte studier ble det avdekket syv kategorier koder (jmf. 4.2 Koding av inkluderte studier). Ved syntese ble funn organisert i tre kategorier (negativ gradient, ingen gradient, positiv gradient). Antatt «beste» intervensjon ble plassert først, og fikk en positiv gradient dersom undersøkelsen bekreftet at intervensjonen ga et bedre resultat enn kontrollgruppen. Hvert funn ble kategorisert som en mediert måling (blank sirkel) eller endelig funn (svart sirkel). Den medierte målingen viser alle former for målinger av kognitive funksjoner i inkluderte studier. Endelige funn (svart sirkel) viser alle former for måling av akademisk prestasjon (jmf. modell i 2.2 Teoretiske modeller for betydningen av fysisk aktivitet på kognitive funksjoner og akademisk prestasjon). Videre ble hvert funn merket med tilhørende metodologisk kvalitetsscore (tall i sirkel). Tabell 17 viser at alle funn er fremkommet som medierte gradienter. Hovedtyngden av funn dreier seg om undersøkelse av sammenhengen mellom forsterket frekvens og ordinær undervisning. Tre funn viser en positiv gradient. Disse tre studiene viser en fordel for fysisk aktivitet med forsterket frekvens (økt treningsmengde) sammenlignet med ordinær undervisning. Et funn avdekker ingen gradient (likt) mellom fysisk aktivitet med forsterket frekvens og ordinær undervisning. Kategorien forsterket intensitet versus akademisk ga to funn. En undersøkelse avdekket en positiv mediert gradient for fordel fysisk aktivitet med forsterket intensitet sammenlignet med ordinær undervisning. Målingen ble gjentatt etter 24 timer og avdekket ingen gradient (kategori: Forsterket intensitet vs. Akademisk etter 24 timer). Et funn viste også negativ gradient for kategorien Forsterket intensitet vs. Akademisk. Som betyr en målt fordel for ordinær undervisning sammenlignet med fysisk aktivitet med forsterket intensitet. Alle akuttstudiene hadde en tilstrekkelig metodologisk kvalitet (6 eller bedre, skala 1 – 10). Syntesen avdekker videre at femte trinns elever har en bedre effekt enn tredje trinns elever, en dobling av økter har bedre effekt, og at forsterket intensitet har en fordel sammenlignet med

en kognitivt beriket økt. Kognitivt beriket fysisk aktivitet er satt sammen for å øke kognitivt stimuli.

Tabell 17: Harvestplot med syntese av akutt fysisk aktivitetsintervensjoner.

Koder for typen ulikheter ved intervensjoner	Negativ gradient (effekt)	Ingen gradient	Positiv gradient
Kognitivt beriket vs. akademisk	⑥		
Forsterket intensitet vs. akademisk	⑥		⑦
Forsterket frekvens vs. akademisk		⑥	⑥⑥⑥
Forsterket intensitet vs. akademisk etter 24 timer		⑦	
Alder 5 trinn vs. 3 trinn			⑥
Forsterket frekvens *2 vs. forsterket frekvens			⑥
Forsterket intensitet vs. kognitivt beriket			⑥

Resultatene blir utdypet i diskusjonskapittelet (jmf. 5.2 Progresjon i kognitiv funksjon og akademisk prestasjon).

4.4.2 Kroniske fysiske aktivitetsintervensjoner

Ved syntese av intervensjonsulikheter for kroniske studier ble det avdekket tolv kategorier koder (jmf. 4.2 Koding av inkluderte studier). Ved syntese ble funn organisert i tre kategorier (negativ gradient, ingen gradient, positiv gradient). Antatt «beste» intervensjon ble plassert først, og fikk en positiv gradient dersom undersøkelsen bekreftet at intervensjonen ga et bedre resultat enn kontrollgruppen. Hvert funn ble kategorisert som en mediert (blank sirkel) eller endelig resultat (svart sirkel) som viser alle former for måling av akademisk prestasjon.

Videre ble hvert funn merket med tilhørende metodologisk kvalitetsscore (tall i sirkel). Tabell 18 viser at funn er fremkommet målt som kognitiv funksjon eller som akademisk prestasjon. Hovedtyngden av funn dreier seg om undersøkelse av sammenhengen mellom kognitivt beriket og ordinær undervisning. Fire funn viser en positiv gradient. Tre funn viser ingen gradient. Disse syv undersøkelsene viser enten en fordel for kognitivt beriket fysisk aktivitet (kognitivt stimulert) eller ingen effekt, sammenlignet med ordinær undervisning. Kategorien kognitivt beriket versus kroppsøving ble undersøkt i fem tilfeller. En undersøkelse avdekket

en positiv gradient for fordel kognitivt beriket fysisk aktivitet sammenlignet med ordinær kroppsøving. Tre funn viste ingen gradient. Av disse tre var to undersøkelser målinger av akademisk prestasjon. Et funn viste en negativ mediert gradient, og et funn viste en positiv mediert gradient. Kategorien forsterket intensitet versus kroppsøving ble undersøkt i fire tilfeller. Alle undersøkelsene ga en positiv gradient for fordel fysisk aktivitet med forsterket intensitet sammenlignet med ordinær kroppsøving. To av disse resultatene var ikke signifikant (i parentes). Kategorien forsterket frekvens versus ordinær kroppsøving ble undersøkt i to tilfeller. Begge undersøkelsene ga ingen gradient. Dette betyr at forsterket mengde fysisk aktivitet ga den samme effekten som ordinær kroppsøving. Kategorien forsterket intensitet versus ordinær undervisning ble undersøkt i ett tilfelle og ga ingen gradient. En positiv gradient ble vist i enkeltfunn for kategoriene: forsterket frekvens versus kognitiv beriket, forsterket intensitet versus forsterket frekvens og strukturert fysisk aktivitet versus ustrukturert fysisk aktivitet. Ingen gradient ble vist i enkeltfunn for kategoriene: type fysisk aktivitet aerob versus aerob og styrketrening (ikke-signifikante måleresultater), ustrukturert fysisk aktivitet versus ordinær kroppsøving og ved å innføre fysiske tester for å måle effekten ved atferdsendring. En undersøkelse som sammenlignet forsterket intensitet og kognitiv beriket fysisk aktivitet (⑦*) viste en negativ gradient. Dette betyr at kognitiv beriket fysisk aktivitet ga bedre effekt enn forsterket intensitet for en mediert måling (kognitiv funksjon). Kroniske studier hadde både lav (5, skala 1 - 10) og tilstrekkelig metodologisk kvalitet (6 – 9, skala 1 – 10).

Tabell 18: Harvestplot med syntese av kroniske fysisk aktivitetsintervensjoner.

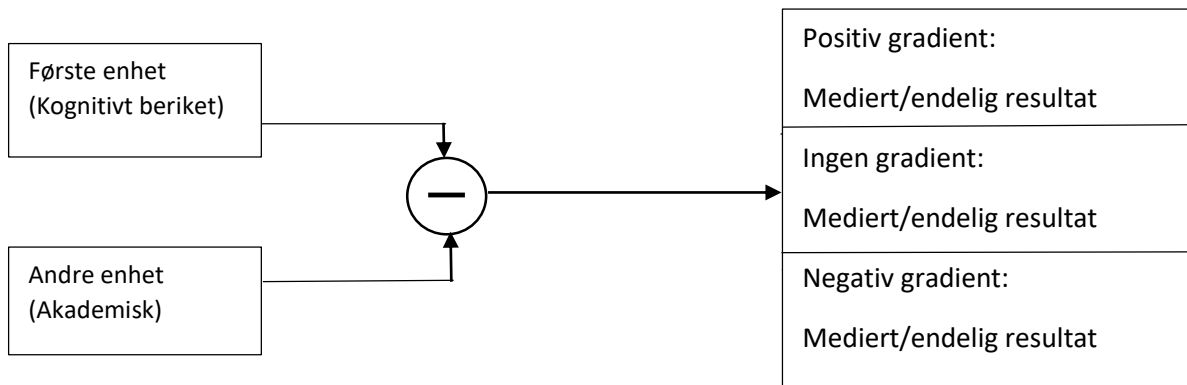
Koder for typen ulikheter ved intervensjoner	Negativ gradient (effekt)	Ingen gradient	Positiv gradient
Kognitivt beriket vs. akademisk		⑤ ⑤ ⑦	⑥ ⑦ ⑦ ⑧
Kognitivt beriket vs. kroppsøving	⑧	⑧ ⑧ ⑧	⑧
Forsterket intensitet vs. akademisk		⑦	
Forsterket intensitet vs. kroppsøving			(⑧)(⑧) ⑤ ⑤
Forsterket frekvens vs. kroppsøving		⑤ ⑤	
Forsterket intensitet vs. Kognitivt beriket	⑦*		
Forsterket frekvens vs. kognitivt beriket			⑨
Forsterket intensitet vs. forsterket frekvens			⑤
Type, aerob vs. aerob og styrke		(⑧)	
Strukturert vs. ustrukturert			⑥
Ustrukturert vs. kroppsøving		⑤	
Fysiske tester for atferdsendring		⑤	

Resultatene blir utdypet i diskusjonskapittelet (jmf. 5.2 Progresjon i kognitiv funksjon og akademisk prestasjon).

4.4.3 Utvikling av en teoretisk modell om hvordan intervensjonen fungerer

Teorikapittelet presenterer en rekke intuitive modeller for sammenhengen mellom fysisk aktivitet og kognisjon eller akademisk prestasjon (jmf. 2.2 Teoretiske modeller for betydningen av fysisk aktivitet på kognitive funksjoner og akademisk prestasjon). Modellene er godt etablert. Ulike forskere bruker ulike modeller avhengig av hensikten med undersøkelsene (Lubans et al., 2016; Voss et al., 2013; Howie & Pate, 2012; Diamond, 2013). Det kan likevel hevdes at empiriske data ikke alltid passer til de etablerte modellene. Man kan få ulike måleresultater for tilsynelatende samme stimuli. Systematiske review gir et kraftig verktøy for utvikling av modeller. Sammenslutningen av primærstudier i en syntese gir muligheten for utvikling av en modell som kan forklare sammenhengen mellom stimuli og måleresultater (Popay et al., 2006). Modellen (Figur 5) gir et direkte ekstrakt av syntesen for

en kategori, eksempel «Kognitivt beriket versus akademisk». Differansen mellom enhetene i intervensjonsulikhet gir et mediert eller endelig måleresultat som kan ha positiv, ingen eller negativ gradient. Dette betyr at intervensjonen kan fungere bedre, likt eller dårligere enn kontrollgruppen ved undersøkelse av kognitiv funksjon eller akademisk prestasjon (Figur 5).



Figur 5: Modell av sammenhengen mellom intervensjonsulikhet og resultat. Modellen er utviklet på bakgrunn av empiriske resultater.

Modellen tar hensyn til kjeden av kausale sammenhenger fra design, intervensjonsulikhet til medierende og endelige resultater. Modellen er opptatt av hvordan, hvorfor og for hvem intervensjonen fungerer. Modellen kan bidra til tolkning av funn, og til vurdering av anvendelsen av undersøkelsene som er oppsummert systematisk (Popay et al., 2006).

4.4.4 Identifisering av mønstre i syntesen

Syntesen av studier er delt i akutte og kroniske fysisk aktivitetsintervensjoner i tråd med teorikapittelet (jmf. 2.4.2 Kvantitative og kvalitative, akutte og kroniske intervensjoner). Kategorier som inneholder forsterket intensitet eller forsterket frekvens er kvantitative intervensjoner. Kategorier som inneholder kognitivt beriket er kvalitative intervensjoner etter teorien.

Hovedtyngden av resultater har positiv gradient sammenlignet med ingen eller negativ gradient. Kategoriene inneholder flere funn med positiv og negativ gradient. Dette refererer ikke selvmotsigende resultater, men konkurrerende hypoteser (Ogilvie et al., 2011) (jmf. 5.4 Begrensinger i syntesens robusthet). Blant flere kategorier som inneholder en intervensjon sammenlignet med en kontrollgruppe eksisterer likevel en tendens til positiv gradient eller ingen gradient for intervensjonen.

I forskjellige kategorier er det også slik at resultatene peker i forskjellige retninger. Det er ikke noe urovekkende ved at resultatene peker i ulik retning ved en oppsummering i et systematisk review (Booth et al., 2016). Man tenker seg at forskningen sannsynligvis vil inneholde en bredde av funn, og at gjennomsnittet kanskje kan være av interesse. Syntesen avdekker at hver kategori kan inneholde funn med en viss tetthet, det vil si at flere undersøkelser har sett på samme problemstilling. En viss tetthet av funn kan også være interessant for videre tolkning.

Studiene er presentert med en vurdering av metodologisk kvalitet. Ingen funn er ekskludert på bakgrunn av lav kvalitet, men metodologisk kvalitet kan også tas med i betraktningen av robusthet i evidens. I tillegg er visse studier representert uten signifikante resultater. En del av slutningen man trekker av syntesen, bør også inneholde en merknad om ikke-signifikante resultater (jmf. 5.4 Begrensinger i syntesesens robusthet). Eksempel: Kronisk «Forsterket intensitet versus kroppøving» har 4 funn som enten er ikke-signifikante eller har lav metodologisk kvalitet. Disse resultatene har et lavere evidens nivå og får sannsynligvis ikke samme oppmerksomhet som signifikante resultater med tilstrekkelig metodologisk kvalitet.

4.4.5 Heterogenitet i synteseplanet

Alle inkluderte studier har tilfredsstillende inklusjonskriteriene. Inkluderte studier er kvantitative primærstudier som er relativt homogene når det gjelder forskningsdesign og deltagere (barn og ungdom). Heterogenitet eksisterer i detaljene i primærstudier (baseline karakteristika, individ, intervensjon, varighet og utfall), men også som en følge av ulikheter i syntesen. For følgende områder eksisterer heterogenitet i synteseplanet:

1. Akutte og kroniske fysiske aktivitetsintervensjoner
2. Kvantitative og kvalitative intervensjoner
3. Strukturert og ustrukturerte treningsøkter
4. Ulike kategorier intervensjonsulikheter (intervensjon vs. kontroll, intervensjon 1 vs. intervensjon 2).
5. Hypotesen om positiv, ingen eller negativ gradient for utfall.
6. Utfall, mediert og endelig.
7. Metodologisk kvalitet blant studier (lav og tilstrekkelig metodologisk kvalitet).
8. Signifikante og ikke-signifikante resultater.

I den narrative syntesen endres ikke resultater som følge av heterogenitet i synteseplanet. Resultatene presenteres delt eller som grupperte funn. Metodologisk kvalitet blir tatt hensyn til i vurderingen av robusthet.

4.4.6 Syntesens robusthet

For å kunne si noe om syntesens robusthet kan man benytte gradert evidens. Gradert evidens representerer en vurdering av styrken i evidens (jmf. 3.5.3 Kategorisert metodologisk kvalitet og gradert evidens). Gradert evidens krever at måleresultatene har en klinisk relevans. I synteseplanet er resultatene presentert enten som en mediert måling av kognitiv funksjon eller en endelig akademisk prestasjon. Opprinnelsen til begrepet gradert evidens er medisinske eksperimenter. Begrepet «klinisk relevans» kan ikke videreføres med samme betydning som for medisinsk forskning. Man kan videreutvikle og anvende begrepet «klinisk relevans» i en pedagogisk sammenheng. Dersom begrepet «klinisk relevans» defineres som en progresjon i akademisk prestasjon, kan syntesen benyttes til å identifisere akademisk prestasjon med positiv gradient. Akuttstudiene inneholder ikke målinger av akademisk prestasjon. Blant kroniske fysisk aktivitetsintervensjoner inneholder to kategorier målinger av akademisk prestasjon: Have et al. (2016) og Mullender Wijsma et al. (2016) fant med sterk evidens progresjon i akademisk prestasjon for kognitiv beriket fysisk aktivitet versus ordinær undervisning (tabell 19). Quino Romani & Klausen (2017) fant progresjon i akademisk prestasjon for forsterket intensitet versus kroppsøving. Manglende robusthet i evidens skyldes lav metodologisk kvalitet (Tabell 19).

Tabell 19: Progresjon i akademisk prestasjon for kroniske fysisk aktivitetsintervensjoner

<i>Typen intervensjonsulikhet</i>	<i>Evidensnivå</i>	<i>Intervensjonsulikhet</i>	<i>Metodologisk vurdering</i>	<i>Utfall</i>
Kvalitativ	Sterk	Kognitivt beriket versus akademisk,	tilstrekkelig	progresjon i akademisk prestasjon for to undersøkelser (St14, St16).
Kvantitativ	Ingen	Forsterket intensitet versus kroppsøving,	lav	progresjon i akademisk prestasjon for en undersøkelse (St17).

Dersom begrepet «klinisk relevans» utvides til progresjon av kognitive funksjoner som et forstadium til progresjon i akademisk prestasjon, kan syntesen benyttes til å identifisere målinger av kognitive funksjoner med positiv gradient. Akutte fysisk aktivitetsintervensjoner

inneholder syv kategorier målinger av kognitiv funksjon rangert etter gradert evidens: Chen et al. (2014) og Altenburg et al. (2016) fant sterk evidens for progresjon i kognitiv funksjon for forsterket frekvens versus ordinær undervisning (Tabell 20). Svak evidens ble funnet i en rekke undersøkelser. Redusert robusthet skyldes enkeltstående funn. Altenburg et al. (2016) fant progresjon i kognitiv funksjon for forsterket frekvens (dobling av treningsmengde) versus forsterket frekvens (enkel treningsmengde) (Tabell 20).

Tabell 20: Progresjon i kognitiv funksjon for akutt fysisk aktivitetsintervensjoner

<i>Typen intervensjonsulikhet</i>	<i>Evidensnivå</i>	<i>Intervensjonsulikhet</i>	<i>Metodologisk vurdering</i>	<i>Utfall</i>
Kvantitativ og kvalitativ	Sterk	Forsterket frekvens versus akademisk	tilstrekkelig	progresjon i kognitiv funksjon for tre undersøkelser (St2, St3).
Kvantitativ	Svak	Forsterket frekvens *2 versus forsterket frekvens *1	tilstrekkelig	progresjon i kognitiv funksjon for en undersøkelse (St3).
Kvantitativ	Svak	Alder (trinn 5 versus trinn 3)	tilstrekkelig	progresjon i kognitiv funksjon for en undersøkelse (St2).
Kvantitativ og kvalitativ	Svak	Forsterket intensitet versus kognitiv beriket	tilstrekkelig	progresjon i kognitiv funksjon for en undersøkelse (St4).
Kvalitativ	Svak	Akademisk versus kognitivt beriket	tilstrekkelig	progresjon i kognitiv funksjon for en undersøkelse (St4).
Kvantitativ og kvalitativ	Svak	Forsterket intensitet versus akademisk	tilstrekkelig	progresjon i kognitiv funksjon for en undersøkelse. Konkurrerende hypotese (St1).
Kvalitativ og kvantitativ	Svak	Akademisk versus forsterket intensitet	tilstrekkelig	progresjon i kognitiv funksjon for en undersøkelse. Konkurrerende hypotese (St4).

Etnier et al. (2014) og Galotta et al. (2015) fant konkurrerende hypoteser for progresjon i kognitiv funksjon. Forsterket intensitet versus akademisk konkurrerer med akademisk versus

forsterket intensitet (Tabell 20). I diskusjonskapittelet diskuteres begrensinger i robusthet (jmf. 5.4 Begrensinger i syntesesens robusthet).

Kroniske fysisk aktivitetsintervensjoner inneholder åtte kategorier med progresjon i kognitiv funksjon rangert etter gradert evidens. Hill et al. (2010) og Schmidt et al. (2015) fant med sterk evidens for progresjon i kognitiv funksjon for kognitivt beriket fysisk aktivitet versus ordinær undervisning. Costigan et al. (2015) og Ardoy et al. (2014) fant med sterk evidens for progresjon i kognitiv funksjon for forsterket intensitet versus kroppsøving. Studiene hadde både tilstrekkelig og svak metodologisk kvalitet, og ikke-signifikante og signifikante resultater (Tabell 21). Svak evidens ble funnet i en rekke undersøkelser. Redusert robusthet skyldes enkeltstående funn.

Tabell 21: Progresjon i kognitiv funksjon for kronisk fysisk aktivitetsintervensjoner.

<i>Typen intervensjonsulikhet</i>	<i>Evidensnivå</i>	<i>Intervensjonsulikhet</i>	<i>Metodologisk vurdering</i>	<i>Utfall</i>
Kvalitativ	Sterk	Kognitivt beriket versus akademisk	Tilstrekkelig	progresjon i kognitiv funksjon for to undersøkelser (St5, St6).
Kvantitativ	Sterk	Forsterket intensitet versus kroppsøving	Tilstrekkelig/svak	progresjon i kognitiv funksjon for tre undersøkelser. (St7 Ikke-signifikante resultater i to, St9)
Kvantitativ og kvalitativ	Svak	Kognitiv beriket versus forsterket intensitet	Tilstrekkelig	progresjon i kognitiv funksjon for en undersøkelse (St6).
Kvantitativ og kvalitativ	Svak	Forsterket frekvens versus kognitivt beriket	Tilstrekkelig	progresjon i kognitiv funksjon for en undersøkelse (St8).
Kvantitativ	Svak	Forsterket intensitet versus forsterket frekvens	Svak	Progresjon i kognitiv funksjon for en undersøkelse (St9).
Kvantitativ og kvalitativ	Svak	Strukturert versus ustrukturert	Tilstrekkelig	progresjon i kognitiv funksjon for en undersøkelse (St11).
Kvalitativ og kvantitativ	Svak	Kognitivt beriket versus kroppsøving	tilstrekkelig	progresjon i kognitiv funksjon for en

				undersøkelse (St12). Konkurrerende hypotese.
Kvantitativ og kvalitativ	Svak	Kroppsøving versus kognitiv beriket	Tilstrekkelig	progresjon i kognitiv funksjon for en undersøkelse (St10). Konkurrerende hypotese.

Resaland et al. (2015) og Tarp et al. (2016) fant konkurrerende hypoteser for progresjon i kognitiv funksjon. Kognitivt beriket versus kroppsøving konkurrerer med kroppsøving versus kognitivt beriket (Tabell 21). I diskusjonskapittelet avgjøres konkurrerende hypoteser (jmf. 5.4 Begrensinger i syntesens robusthet).

4.4.7 En «sann» effektstørrelse

En sammenstilling av studier med høy utvalgsstørrelse kan indikere hvor den «sanne» verdien for effektstørrelse ligger. Funnplot anvendes i metaanalyser for å plassere studier langs aksene effektstørrelse og utvalgsstørrelse. Toppunktet i et funnplot kan gi en tolkning av den «sanne» effektstørrelsen for denne typen studier. Tilsvarende, men ufullstendig betraktning kan gjøres ved et Harvestplottet (tabell 22). Syntesen presenterer 3 studier med $N > 1000$ og 3 studier med $N \approx 500$, for store studier. Grunlaget for beregnet utvalgsstørrelse er gitt i diskusjonskapittelet (jmf. 5.7.5 Hvilke utfall brukes til å beregne utvalgsstørrelse i primærresultater?).

Tabell 22: Store studier ($N > 1000$ og $N \approx 500$) akutte og kroniske fysisk aktivitetsintervensjoner.

Koder for typen ulikheter ved intervensjoner	Negativ gradient (effekt)	Ingen gradient	Positiv gradient
Kognitivt beriket vs akademisk		⑦	⑥ ⑦ ⑧
Kognitivt beriket vs kroppsøving	⑧	⑧ ⑧ ⑧	⑧
Forsterket intensitet vs kroppsøving			⑤
Ustrukturert vs kroppsøving		⑤	

Seks store studier er representert i syntesen (Hill et al., 2010; Resaland et al., 2016; Aadland et al., 2017; Quinto Romani & Klausen, 2017; Tarp et al., 2014; Have et al., 2016;

Mullender-Wijnsma et al., 2016). Akademisk prestasjon er målt i 6 tilfeller, mens mediert funn (kognitiv funksjon) er målt fem ganger. Syntesen avdekker sterk evidens for progresjon i akademisk prestasjon for kategorien «Kognitiv beriket vs akademisk». Syntesen avdekker sterk evidens for ingen effekt på akademiske prestasjon for kategorien «kognitivt beriket vs kroppsøving». Undersøkelsene har tilstrekkelig metodologisk kvalitet. Videre avdekkes en svak evidens for enkeltstående mediert funn (kognitiv funksjon). Undersøkelsen av «Forsterket intensitet vs. Kroppsøving» og «Ustrukturert vs. Kroppsøving» gir ingen evidens som følge av lav metodologisk kvalitet.

4.4.8 Undersøkelse av publikasjonsavvik

I metaanalyser gir Funnplot en indikasjon på publikasjonsavvik. Man forventer en symmetrisk pyramide med studier jevnfordelt på begge sider av en generalisert linje for effektstørrelse. Ved å undersøke symmetri blant studier kan skjevheter fremkomme langs aksene – effektstørrelse. Tilsvarende undersøkelse kan gjennomføres ved hjelp av et Harvestplot. Syntesen av 8 småstudier (N<100 deltagere per gruppe) vises i Tabell 23. For undersøkelser med sammenligning av intervensjon og kontrollgruppe har alle resultater positiv, ingen eller negativ gradient. Kun 3 funn er representert med negativ gradient. Dette leder til antagelsen om at syntesen sannsynligvis inneholder publikasjonsavvik. Sannsynlighet for publikasjonsavvik diskuteres i diskusjonskapittelet (jmf. 5.8.2 Kan Harvestplot indikere publikasjonsavvik?).

Tabell 23: Småstudier akutte og kroniske fysisk aktivitetsintervensjoner N < 100 deltagere per gruppe.

Koder for typen ulikheter ved intervensjoner	Negativ gradient (effekt)	Ingen gradient	Positiv gradient
Kognitivt beriket vs. akademisk	⑥		⑦
Forsterket intensitet vs akademisk	⑥	⑦	⑦
Forsterket frekvens vs akademisk		⑥	⑥ ⑥ ⑥
Forsterket intensitet vs kroppsøving			(⑧)(⑧) ⑤
Forsterket frekvens vs kroppsøving		⑤	
Forsterket intensitet vs akademisk etter 24 timer		⑦	
Alder 5 trinn vs 3 trinn			⑥
Forsterket frekvens *2 vs forsterket frekvens			⑥
Type, aerob vs aerob og styrke		(⑧)	
Forsterket frekvens vs kognitivt beriket			⑨
Forsterket intensitet vs forsterket frekvens			⑤
Forsterket intensitet vs. kognitivt beriket	⑦		⑥

Valget av en grenseverdi på 100 deltagere per gruppe er hentet fra Resaland et al. (2015) (jmf 5.7.5 Hvilke utfall brukes til å beregne utvalgsstørrelse i primærresultater?).

5.0 Diskusjon

I denne delen diskuteres og utforskes funn og mangler ved undersøkelsene som er presentert i resultatdel. Diskusjonskapittelet følger malen til Booth et al. (2016) som formell tilnærming til det systematiske litteraturreviewet. Validitetsvurderingen tar for seg ekstern validitet, utvalgsstørrelse og internvaliditet for inkluderte studier. Funn som har gitt progresjon i kognitiv funksjon og akademisk prestasjon blir presentert. Anbefalinger om bestemte typer fysisk aktivitet fremmes på bakgrunn av robusthet i evidens. Analysen har som mål å vurdere styrken i evidens for å kunne trekke slutninger om resultatene. Hensikten er videre å kunne generalisere resultatene til en målpopulasjon eller til en valgt kontekst. Holdbarhet i funn for konkurrerende hypoteser avgjøres ved prinsippet om «best evidens». Prinsippet om «best

evidens» kombinerer teknikker som tar hensyn til robusthet ved syntesen og grunnlaget for analysen (Slavin, 1995). En sensitivitetsanalyse utforsker konsekvensen av manglende, uklare eller åpenbart uriktige, metoder eller data i inkluderte studier (Booth et al., 2016).

Sensitivitetsanalysen kan danne et utgangspunkt for hvordan enkelte primærstudier gir implikasjoner for reviewet. Gjennom en rekke spørsmål identifiseres manglende rapportering og avvik i søk og utvelgelse av studier, mangler ved måleresultater i primærstudier og publikasjonsavvik for reviewet. Kapittelet avsluttes med styrker og begrensninger i primærstudier og i reviewet.

5.1 Vurdering av validitet

Hovedhypotese 2 (Primærstudiene har tilstrekkelig metodologisk kvalitet) er oppfylt for 76.5% av inkluderte studier. Inkluderte studier hadde RCT og CRT design. Derfor ble to sjekklister for vurdering av metodologisk kvalitet tatt i bruk. Begge listene fikk sitt «cut off» ved 60%. Studier med kvalitetsvurdering $\geq 60\%$ fikk betegnelsen «tilstrekkelig kvalitet». Studier med kvalitetsvurdering $< 60\%$ fikk «lav kvalitet». Sjekklisene for RCT og CRT var tilpasset forskningsdesignet og hadde hhv. 11 og 22 metodologisk kriterier (jmf. 4.3 Metodologisk kvalitet). I syntesene ble den metodologiske kvalitetsvurderingen videreført (0 - 10 poeng tilsvarer 0 – 100%) med inndeling i lav og tilstrekkelig metodologisk kvalitet (4.4 Syntese av inkluderte studier – et narrativ). Ingen av sjekklisene er spesielt tilpasset studier som undersøker sammenhengen mellom fysisk aktivitetsintervensjoner og kognisjon eller akademisk prestasjon.

5.1.1 Eksternvaliditet

I hvilken grad primærstudienes resultater kan generaliseres utover studiens kontekst slik den ble gjennomført beskrives av *eksternvaliditet*. Galotta et al. (2015) og Have et al. (2016) manglet rapportering som omhandlet generalisering av resultater. Tolking av evidens er et sentralt kriterium ved vurdering av eksternvaliditet. CRT studiene (Galotta et al., 2015; Schmidt et al., 2015; Have et al., 2016) manglet tolkning av evidens ved funn i studiene. Eksternvaliditet er et globalt begrep som bør undersøkes for en rekke domener. Ifølge Dekkers et al. (2009) kan eksternvaliditet vurderes spesielt i forbindelse med:

- i. Utvalgets berettigelseskriterier

Primærstudiens utvalg bør som følge av berettigelseskriterier, være representative for populasjonen. Rekrutterte deltagere som samtykker og ikke samtykker til deltagelse har ulike karakteristika. Drop out blir et sentralt kriterium for vurdering av eksternvaliditet. Drop out >

15% eksisterer for 17.6% av studiene (3 av 17). RCT studiene (Arday et al., 2014; Subramanian et al., 2015; Fedewa et al., 2015), som har drop out mer enn 15%, rapporterer ikke karakteristika for rekrutterte deltagere som ble ekskludert eller ikke samtykket til deltagelse. Disse studiene har et problem med generalisering. En innkjøringsperiode kan ekskludere deltagere som ikke egner seg for undersøkelsen (Dekkers et al., 2010). Altenburg et al. (2016) antyder at enkelte deltagere hadde vanskeligheter med å være aktive gjennom hele 20 minutter. Chen et al. (2014) iverksatte en detaljert baselinevurdering og ekskluderte deltagere som ikke var i stand til å gjennomføre forsøket.

ii. Midlertidighet, etniske, sosioøkonomiske og geografiske forhold

Ifølge Dekkers et al. (2009) er målpopulasjonen per definisjon ulik primærstudien utvalg (deltagerne) ved ulike populasjoner. Muligheten for generalisering begrenses ved endrede geografiske og sosioøkonomiske forhold, ved midlertidighet (endring i forskningen) og for ulike etniske populasjoner. Midlertidighet gjelder fordi forskningsfeltet utvikles. Få studier er rene reproduksjoner. I beste er fall er studiene replikasjoner med lignende forskningsdesign. Studien er gjennomført i vestlige land med noen unntak (Chen et al., 2014; Telles et al., 2013; Subramanian et al., 2015). Disse studiene har et problem med generalisering til vestlige land, Europa, USA og Australia.

iii. Forhold utover utvalgets berettigelseskriterier

I randomiserte eksperimenter brukes ofte et strengt alderskriterium. Generalisering utenfor utvalgets alderssammensetning bør baseres på erfaring og biologisk sannsynlighet (Dekkers et al., 2009). Hill et al. (2010) antyder at analysen avdekker en aldersmoderator som kan gi ulike måleresultater. Hill et al. (2010) målte trinnvis (trinn 4, 5, 6, 7) ulik kognitiv funksjon. Chen et al. (2014) målte en bedring av eksekutiv funksjonen fra trinn 3 (9 år) til trinn 5 (11 år). Dekkers et al. (2009) nevner multiple lidelser som et domene som krever nærmere undersøkelse før generalisering. For fysisk aktivitetsintervensjoner kan tilsvarende avvik i baselineverdier vanskeliggjøre generalisering.

iv. Anvendelighet av resultatene

Måleresultater kan påvirkes av ferdighetene til treningslærer og av settingen hvor forsøket foregår (økologisk validitet) (Dekkers et al., 2009). Variasjon i instruksjonen kan svekke muligheten for generaliserte resultater. Chen et al. (2014) antyder at instruksjonen kan ha vært farget av forsøket. Forsøkssettingen kan påvirke effekten av intervensjonen. Subramanian et al. (2015) undersøkte strukturert og ustrukturert aktivitet som innebærer ulik instruksjon og forskjellige settinger. Forsøkssetting dreier seg om personell som benyttes under forsøket, kontrollfrekvenser og materiell som brukes i forbindelse med intervensjonen. Modererende

variabler ble kontrollert ved lik fordeling av kjønn, alder, trent/utrent i intervensjon og kontrollgruppe (Subramanian et al., 2015). Forsøkene foregår i skolen, både i klasserom, gymsal, men også utendørs. Kontrollmetode og frekvens er nevnt som et gjennomgående problem (jmf. 5.4 Begrensinger i syntesens robusthet). Administrative beslutninger kan ha innflytelse på effekten av fysisk aktivitetsintervensjoner (Dekkers et al., 2010).

v. Den reaktive effekt av eksperimenter

Deltagere er klar over at de deltar i et eksperiment. Vissheten om deltagelse kan påvirke hvordan de responderer på eksperimentell aktivitet. Dette har betydning for studiens generalisering. Internvaliditet – avvik (Downs & Black, 1998) omtaler blinding av deltagere. Manglende blinding av deltagere for intervensjonen og manglende tiltak for å hindre smitteeffekt forekom i alle RCT og CRT design. Fysisk aktivitetsintervensjoner er ikke mulig å holde skjult for deltagerne. Man bør anta at deltagerne snakker sammen om forsøket. Følgelig vil i tillegg en «smitteeffekt» sannsynligvis inntre. Blinding av treningslærer under gjennomføring er ikke mulig, og blinding av forskningspersonell under datainnsamling ble gjennomført i 2 av 9 RCT studier. Smitteeffekten er sannsynligvis noe redusert i større studier hvor elevene befinner seg ved ulike skoler. Resaland et al. (2015) medgir at blinding av deltagere var umulig som følge av eksperimentets natur. Studien hadde opprettet en ledergruppe som hadde formell kunnskap om gruppetildelingen. Forsker som behandlet data og statistikk var blindet i forhold til randomisert gruppetildeling inntil analysen var gjennomført. Alle 8 CRT studier har et generelt problem med blinding av deltagere for intervensjonen og smitteeffekt.

5.1.2 Beregning av utvalgsstørrelse

En presis utvalgsstørrelse sannsynliggjør et presist utvalg ved generalisering fra utvalg til målpopulasjon. Andelen studier som rapporterer en beregnet utvalgsstørrelse er 29.4% (5 av 17). I synteseplanet ble tre studier identifisert som store studier ($N > 1000$). Hill et al. (2010), Resaland et al. (2015) og Quinto Romani & Klausen (2017) har oppfylt kravet til utvalgsstørrelse som er et av flere sjekkpunkter ved generalisering. Hill et al. (2010) og Resaland et al. (2015) har tilstrekkelig metodologisk kvalitet (hhv. 63.6 % og 86.4%). I tillegg ble tre studier identifisert med $N \approx 500$ deltagere. Disse studiene har en beregnet utvalgsstørrelse, men med reduserte krav til type 2 feil ($1 - \beta = 0.80$) og antatt moderat intrakorrelasjonskoeffisient. Tarp et al. (2016), Have et al. (2016) og Mullender-Wijnsma et al. (2016) oppfyller kravet om beregnet utvalgsstørrelse, som er et av flere sjekkpunkter ved

generalisering. Studiene har tilstrekkelig metodologisk kvalitet (hhv. 81.8%, 72.7% og 86.4%).

Store studier gir en «sann» effektstørrelse for sammenhengen mellom fysisk aktivitet og akademisk prestasjon (jmf. 4.4.7 En «sann» effektstørrelse). Kategorien kognitivt beriket versus ordinær undervisning dokumenterer sterk evidens for progresjon i akademisk prestasjon ved kronisk kvalitativ fysisk aktivitetsintervensjon. Ifølge Have et al. (2016) og Mullender- Wijnsma et al. (2016) er gradienten positiv. Kategorien kognitivt beriket versus kroppsøving dokumenterer sterk evidens for ingen effekt i akademisk prestasjon ved kronisk kvalitativ fysisk aktivitetsintervensjon. Ifølge Tarp et al. (2016) og Resaland et al. (2016) er gradienten null. Motsetningen i resultatet kan forklares på flere måter. Ulike kontrollgrupper kan forklare forskjellen. Tarp et al. (2016) forklarer selv manglende effekt på akademisk prestasjon ved:

1. Kortvarig intervensjon (kun 20 uker).
2. Ulikhet i baselineverdier ble registrert. Randomiseringen kan være årsak (mulig konfunder).
3. Målingene av akademisk prestasjon ble gjennomført med en (1) matetikkprøve. En matematikktest gir ikke tilstrekkelig målefølsomhet. Man trenger flere målinger og en standardisering av testene (Tarp et al., 2016).
4. Studien mangler registrering av drop out etter allokering og før baselinemålinger (mulig konfunder).
5. Kun 38% av akselerometerregistreringene ble protokollført. Akselerometer kan avdekke manglende kontrast mellom intervensjon og kontrollgruppe (mulig konfunder).
6. Rekrutterte skoler hadde på forhånd implementert et program for fysisk aktivitet. Dermed ble effekten som følge av eksperimentet mindre.

Resaland et al. (2016) forklarer selv manglende effekt på akademisk prestasjon ved:

1. Manglende kontrast mellom intervensjon og kontrollgruppe. Planlagt dose - respons var i utgangspunktet liten (konfunder).
2. Treningsprogrammet ble ikke fulgt. Treningslærer rapporterte gjennomført, 80% av planlagt fysisk aktivitet i intervensjonsgruppe og 120% av planlagt aktivitet i kontrollgruppe. Forsøket varte i 7 måneder.

3. Treningslærer ble gitt stor grad av frihet i trenergjerningen, men mentale og koordinative utfordringer (elevenes oppfatning) ble ikke målt (konfunder).
4. Baselineverdier for akademisk prestasjon var ulike. Delanalyse avdekket en positiv effekt for deltagere med svak akademisk prestasjon, som indikerer at det akademiske nivået er moderator for denne type forsøk.
5. Manglende kontroll med forsøksomgivelsen (begrenser muligheten for generalisering).

Manglende effekt kan skyldes en rekke metodologiske problemstillinger (systematiske og tilfeldige), men også forhold som er særegne for hver studie. Det er verdt å merke seg at studiene har tilfredsstillende metodologiske kvalitet, men har i tillegg begrensninger som ikke registreres i en standardisert sjekklister.

Have et al. (2016) forklarer progresjon og mulige begrensninger med:

1. En ny matematikkprøve ble tatt i bruk. Testen er ikke reliabilitetstestet.
2. Kun 55% av aktivitetsregistreringen ble protokollført. Kun 6% av hvert spørreskjema ble besvart.
3. Lærere ble gitt stor grad av frihet til å implementere fysisk aktivitet i undervisningen. Det kvalitative innholdet, typen aktivitet og kompleksiteten i den sosiale konteksten ble ikke protokollført.
4. I løpet av forsøket var skolen i lock out (25 dager), hvor deltagerne ikke gjennomførte intervensjon og kontroll. Forsøket varte i 9 måneder.

Mullender-Wijnsma et al. (2016) forklarer progresjon og mulige begrensninger med:

1. Treningslærerne var spesialister på treningsprogrammet (F&V lessons). Dette kan ha gitt særlige gode resultater første året. Forsøket varte 20 uker pluss 20 uker.
2. Child academic monitoring system (CAMS), monitorering av deltagerne ble administrert av skolen selv, og kan ha påvirket måleresultatene.

Sannsynlig årsak til dokumentert positiv gradient er at begge studiene har gjennomført fysisk aktivitet med direkte relasjon til matematikk i en god sosial setting med dyktige treningslærere. Forsøkene er ikke gjennomført med kontinuitet.

5.1.3 Internvaliditet

Internvaliditeten i en studie avgjøres ved å undersøke systematiske avvik i resultatene. En studie med tilfeldige avvik kan fortsatt ha høy internvaliditet (Dekkers et al., 2009). En høy

internvaliditet sannsynliggjør at den fysiske aktivitetsintervensjonen er årsak til den observerte effekten. En lav internvaliditet sannsynliggjør at konfundere er årsak til den observerte effekten. Internvaliditet er en global konstruksjon som fokuserer på kausalitet, og som bør undersøkes for en rekke domener. Kriterier som kan true internvaliditet er ifølge Bryman (2008):

- i. Rekkefølge, den kausale retning (årsak kommer før effekt).

Domenet rekkefølge ivaretas av RCT designet.

- ii. Seleksjon, som følge av *forhåndsforskjeller mellom intervensjon og kontrollgruppe*.

Domenet seleksjon, er ideelt sett ivarettatt av RCT design ved at kontrollgruppen kan fungere som referanseverdi. Kvalitetsvurderingene avdekket vesentlige svakheter ved seleksjon. Blant 4 RCT studier (Hill et al., 2010; Ardoy et al., 2014; Chen et al., 2014; Altenburg et al., 2016) eksisterer ulike former for seleksjonsavvik (internvaliditet – konfundere) (Downs & Black, 1998). Et kriterium som er sentralt for vurdering av konfundere er relatert til randomisert allokering av intervensjon og kontrollgruppe. Fordelen med randomisert allokering er at forskjellene mellom gruppene er utelukkende tilfeldige. Ideelt foregår randomisering til RCT design ved tilfeldig utvelgelse av individer til intervensjon og kontrollgruppe. Ideelt foregår kluster randomisering ved å tilfeldig velge et antall skoler i en region. Deretter tilfeldig velge et antall individer fra hver skole til intervensjon og kontrollgruppe. Primærstudiene avviker i noen tilfeller fra denne fremgangsmåten. Andelen studier som mangler en tilfeldig allokering er 11.7% (2 av 17). Hill et al. (2010) gjennomførte randomisert utvelgelse av to klasser ved hver skole, men utvelgelse av skole var ikke beskrevet. Deltagere i hver klasse var heller ikke randomisert. Ardoy et al. (2014) gjennomførte en RCT, men utførte randomisert allokering på klassenivå for 67 rekrutterte. Blant 8 CRT-design viste kvalitetsvurderingen lav risiko for seleksjonsavvik (internvaliditet – konfundere) (Downs & Black, 1998). Alle 8 studier beskrev hvordan allokeringsrekkefølgen ble gjennomført og rapporterte skjult allokering inntil intervensjon og kontroll var bestemt. I de 8 studiene ble baseline informasjon rapportert for hver gruppe og informasjon om rekrutteringsperiode og oppfølging beskrevet. Mullender-Wijnsma et al. (2016) analyserte ulikhet i baselineverdier mellom intervensjon og kontrollgruppe ved en uavhengig t-test eller en X^2 -test. Et eksempel på en mulig konfunderende variabel fremgår i Tarp et al. (2016) hvor skolen hadde tilrettelagt et program for fysisk aktivitet allerede da undersøkelsen ble iverksatt. Tarp et al. (2016) hevder selv at programmet kan ha hatt en konfunderende effekt som svekker internvaliditet. Studien hadde tilstrekkelig metodologisk kvalitet (81.8%).

- iii. Testing, domenet testing omhandler muligheten for at deltagere tilvennes testene og

blir kjent med hensikten med forsøkene. Ved RCT design vil sannsynligvis tilvenning også finne sted i kontrollgruppen.

iv. Instrumentering, domenet instrumentering dreier seg om at endringer ved administrering av tester kan endre måleresultater for pre- og post-tester. Ved RCT design vil eventuelle endringer i gjennomføring også eksistere for kontrollgruppen.

v. Historie, domenet historie inneholder deltageres treningserfaring eller tidligere akademisk prestasjoner. Domenet historie er ideelt sett ivaretatt av RCT designet. Imidlertid målte Resaland et al. (2016) en bedre effekt for deltagere med lav kognitiv score (baseline). Dette betyr at deltageres kognitive og akademisk nivå sannsynligvis er en moderator for fysisk aktivitetsintervensjoner.

vi. Modning, domenet modning gjelder all utvikling som funksjon av tid. Modning er ideelt sett ivaretatt av RCT design ved antagelse om at fordelingen i modenhet er lik for hvert trinn og krever at analysen presenterer resultater kontrollert for alder.

vii. Drop out

Domenet drop out har innvirkning på internvaliditet og kausalitet ved at utvalgets karakteristika endres.

Tabellene nedenfor er en punktvis gjennomgang av rapporterte begrensinger relatert til internvaliditet i primærstudiene. Ikke alle studier har rapportert svakheter ved internvaliditet. Dette betyr ikke at studien er fri for merknader, men at studien ikke har diskutert problemstillingen. Noen studier har rapportert flere punkter som har betydning for intern validitet. I enkelte av disse studiene finns gjennomgående emner som er overførbare til fysisk aktivitetsintervensjoner generelt. Eks. Måleresultater fra kontrollgruppen utgjør ikke et sant nullpunkt (mangler et plasebo for kontrollgruppe).

Tabell 24: Merknader som kan ha betydning for internvaliditet i (akutt) primærstudier.

Studie	Kode for intervensjonsulikheter – akutt intervensjoner.	Merknader som har innflytelse på intern validitet, kausalitet.
St1 Etnier et al. (2014) RCT	Forsterket intensitet vs. Akademisk Forsterket intensitet vs. Akademisk (etter 24 timer)	-
St2 Chen et al. (2014) RCT	Forsterket frekvens vs. Akademisk Forsterket frekvens vs. Akademisk Alder (5 trinn vs. 3 trinn)	-
St3 Altenburg et al. (2016) RCT	Forsterket frekvens vs. Akademisk Forsterket frekvens *2 vs. Akademisk	Domenet testing innebærer tilvenning til testinstrument som kan gi en konfunder. Deltagere gjennomførte ikke aktivitet som planlagt.
St4 Galotta et al. (2015) CRT	Forsterket intensitet vs. Akademisk Kognitivt beriket vs. Akademisk Forsterket intensitet vs. Kognitivt beriket	Modning kan forklare måleresultater. Mangler kontrollgruppe, differansen i måleresultater er avhengig av et «nullpunkt».

Tabell 25: Merknader som kan ha betydning for internvaliditet i (kroniske) primærstudier.

Studie	Kode for intervensjonsulikheter – kronisk intervensjoner.	Merknader som har innflytelse på internvaliditet, kausalitet.
St5 Hill et al. (2010) RCT -crossover	Kognitivt beriket vs. Akademisk	Tidspunkt på dagen hvor aktiviteten gjennomføres påvirker spenningsnivået og effekten av fysisk aktivitet. «Tilvenning» til testene (5 + 5 tester), ikke til den enkelte test. Drop out fra enkelte tester som ikke ble gjennomført pga. vanskelighetsgrad.
St6 Schmidt et al. (2015) CRT	Kognitivt beriket vs. Akademisk Forsterket intensitet vs. Akademisk Forsterket intensitet s. kognitivt beriket	Studien skiller ikke mellom ulike motorisk stimuli. Detaljeringsgraden redusert. Eksperimentets varighet er kort. Domenet modning kan medføre ulike utslag for komponentene i eksekutiv funksjon. Inhibering er ferdig utviklet på eksperimenttidspunktet. Fitness-nivået (moderator) kan være ulikt for intervensjon og kontrollgruppe. Konfunderende at aktiviteten (differanse i dose og intensitet) ikke befinner seg innenfor instrumentets arbeidsområde. Mental mobilisering (innsats). Observasjon gir ikke en standardisert måling. Klasserandomisering er ikke individuell randomisering, seleksjon.
St7 Costigan et al. (2015; 2016) RCT	Forsterket intensitet vs. Kroppsøving Forsterket intensitet vs. Kroppsøving Type aerob vs. Aerob + styrketrening	Kun gjennomsnitt hjertefrekvens ble målt under aktivitet, ikke absoluttverdi ved høy intensitet.
St8 Telles et al. (2013) RCT	Forsterket frekvens vs. Kognitivt beriket	Mangler kontrollgruppe, men har sammenligningsgruppe. Eksperimentets varighet er kort.
St9 Ardoy et al. (2014) RCT	Forsterket frekvens vs. Kroppsøving Forsterket intensitet vs. Kroppsøving Forsterket intensitet vs. Forsterket frekvens	Mangler individuell randomisering, seleksjon.
St10 Tarp et al. (2016), Bugge et al. (2014) CRT	Kognitivt beriket vs. Kroppsøving	Eksperimentets varighet er kort. Skolen hadde på forhånd et program for fysisk aktivitet. Domenet historie medfører ulike baselineverdier som kan inneholde moderatorer. Domenet modning dreier seg om pubertetsutvikling og kan innvirke ulikt på ulike komponenter.

		Drop out mellom randomisering og baselinemålinger endrer utvalgets karakteristika.
St11 Subramanian et al. (2015) RCT	Strukturert vs. Ustrukturert	Krysskontaminasjon mellom grupper, smitte gir en konfunderende variabel. Dose - respons utenfor instrumentets arbeidsområde gir en konfunderende variabel.
St12 Resaland et al. (2015; 2016), Aadland et al. (2017) CRT	Kognitivt beriket vs. Kroppsøving	Dose - respons er for svak for instrumentets arbeidsområde, og kan gi en konfunder. Manglende kontroll med forsøksomgivelsen, åpner for en rekke konfundere. Drop out som skyldes per protokoll analyse (Aadland et al., 2017) kan endre utvalgets karakteristika.
St13 Fedewa et al. (2015) RCT	Kognitivt beriket vs. akademisk	Den sosiale strukturen i skolens hierarki kan introdusere konfunderende variabler.
St14 Have et al. (2016a) Have et al. (2016b) unpubl. CRT	Kognitivt beriket vs. akademisk	Manglende informasjon om lærer - elev relasjonen og kompleksiteten i den sosiale omgivelsen kan introdusere en rekke konfundere.
St15 Ahamed et al. (2007) CRT	Forsterket frekvens vs. kroppsøving	Domenet historie uttrykkes ved ulike baselineverdier som kan introdusere moderatorer. Eksperimentets varighet er kort.
St16 Mullender- Wijnsma et al. (2016), De Greff et al. (2016) CRT	Kognitiv beriket vs. akademisk	-
St17 Quinto Romani & Klausen (2017) CRT	Forsterket intensitet vs. Kroppsøving Ustrukturert vs. Kroppsøving Fysisk tester for atferdsendring	Eksperimentets varighet er kort (dersom effekten er avhengig av et lengre eksperiment). Innholdet i aktivitet ble endret underveis. Informasjonen om innholdet i intervensjonen er mangelfull. Datainnsamling foregikk lenge etter avsluttet eksperiment, muligheten for konfunderende variabler øker.

På bakgrunn av disse resultatene bør argumentasjonen for kausalitet fremmes med varsomhet. Ingen studier er i realiteten fri for merknader. Vurderingen av merknadens betydning for kausalitet bør legges til primærstudien.

5.2 Progresjon i kognitiv funksjon og akademisk prestasjon

Progresjon i kognitiv funksjon og akademisk prestasjon blir presentert i resultatdelen (jmf. 4.4.6 Syntesens robusthet) som et tilsvar til hypotese 1 (Fysiske aktivitetsintervensjoner vil styrke kognitiv funksjon (eks. eksekutiv funksjon) eller akademisk prestasjon sammenlignet med ordinær undervisning eller kroppsøving.). Kategorier av intervensjonsulikheter med progresjon blir ordnet etter evidensnivå.

5.2.1 Akutte kvantitative og kvalitative fysisk aktivitetsintervensjoner

Progresjon i *kognitiv funksjon* (positiv gradient) etter akutt fysisk aktivitetsintervensjoner er målt i 63.6 % av undersøkelsene (7 av 11). Ingen gradient avdekkes for 18.2% av undersøkelsene (2 av 11). Negativ gradient eksisterer for 18.2% av undersøkelsene (2 av 11) (jmf. Tabell 17 i 4.4.1 Akutte fysiske aktivitetsintervensjoner).

Hypotese 1 er oppfylt for kategorien *forsterket frekvens versus ordinær undervisning*. Resultatene dokumenterer sterk evidens for progresjon i kognitiv funksjon ved akutt kvantitativ og kvalitativ fysisk aktivitetsintervensjon. Chen et al. (2014) og Altenburg et al. (2016) fant en positiv effekt på mediert kognitiv funksjon. Disse studiene var basert på den neurobiologiske hypotesen, og målte hhv. eksekutiv funksjon (inhibering, arbeidsminne og skifte av setting) og selektiv oppmerksomhet, i samsvar med teori (jmf. 2.3.1 Eksekutive funksjoner) som plasserer disse målingene under styrt oppmerksomhet og kognitiv fleksibilitet. Sterk evidens gjør at funnene vurderes robuste.

Hypotese 1 er oppfylt for en rekke kategorier som dokumenterer svak evidens for kognitiv funksjon for akutte kvantitative og kvalitative fysisk aktivitetsintervensjoner. Etnier et al. (2014), Chen et al. (2014), Altenburg et al. (2016) og Galotta et al. (2015) er alle basert på den neurobiologiske hypotesen. Etnier et al. (2014) og Galotta et al. (2015) har konkurrerende hypoteser som ble vurdert etter prinsippet om «best evidens» (jmf. 5.4 Begrensinger i syntesens robusthet). Funn med svak evidens kan med fordel vurderes som tentative til tross for at studien er vurdert med tilstrekkelig metodologisk kvalitet.

5.2.2 Kroniske kvantitative og kvalitative fysisk aktivitetsintervensjoner

Progresjon i *kognitiv funksjon* (positiv gradient) etter kronisk fysisk aktivitetsintervensjon er målt i 38.5% av undersøkelsene (10 av 26, kategorien forsterket intensitet vs. kognitivt beriket ((7)*) ble sortert i motsatt rekkefølge). Ingen gradient avdekkes for 23% av undersøkelsen (6 av 26). Negativ gradient eksisterer for 3.8% av undersøkelsene (1 av 26). Progresjon i *akademisk prestasjon* (positiv gradient) etter kronisk fysisk aktivitetsintervensjon er målt i 11.5% av undersøkelsene (3 av 26). Ingen gradient avdekkes for 23% av undersøkelsen (6 av 26). Negativ gradient eksisterer for 0% av undersøkelsene (0 av 26) (jmf. Tabell 18 i 4.4.2 Kroniske fysiske aktivitetsintervensjoner).

Hypotese 1 er oppfylt for kategorien *kognitivt beriket versus ordinær undervisning*.

Resultatene dokumenterer sterk evidens for progresjon i kognitiv funksjon for kroniske kvalitative fysisk aktivitetsintervensjoner. Hill et al. (2010) og Schmidt et al. (2015) er basert på den nevrobiologiske hypotesen. Begge studiene måler mediert kognitiv funksjon, Hill et al. (2010) målte eksekutive funksjoner (informasjonsbearbeiding, arbeidsminne), korttidsminne og svekkelse i kognitiv funksjon. Schmidt et al. (2015) målte progresjon i eksekutiv funksjon (skifte av setting). Hill et al. (2010) (1224) og Schmidt et al. (2015) (181), har ulik utvalgsstørrelse, alder er hhv. 8 - 11 år og 10 - 12 år. Intervensjonene er heterogene i intensitet med hhv. moderat intensitet i klasserommet og høyintensitets lagspill. Evidensnivået kan bestrides som følge av heterogen intensitet. Imidlertid har Schmidt et al. (2015) funnet at kategorien *forsterket intensitet versus akademisk* ikke gir en effekt. Med forbehold om nevnte forhold ovenfor, beholdes vurderingen av sterk evidens.

Hypotese 1 er oppfylt for kategorien *Forsterket intensitet versus kroppsøving*. Resultatene dokumenterer sterk evidens for progresjon i kognitiv funksjon for kronisk kvantitative fysisk aktivitetsintervensjoner. Costigan et al. (2015; 2016) målte eksekutiv funksjon (kognitiv fleksibilitet, visuell oppmerksomhet, prosesseringshurtighet). Costigan et al. (2015) målte psykologisk velvære, psykologisk lidelse (negativt stress), persepsjon av suksess i relasjoner, selv-tillit, hensikt, optimisme og psykologisk tilstand (ble målt med et (1) spørsmål selvrappotert før og etter økt). Ved statistisk analyse ga differansen i effekt ikke signifikante resultater (95% konfidensintervall). Studien var basert på den nevrobiologiske- og den psykososiale hypotesen. Ardoy et al. (2014) målte kognitiv funksjon (intelligenstest). Studien er basert på den nevrobiologiske hypotesen, men har lav metodologisk kvalitet (54.5%). Med hensyn til nevnte forhold vurderes Costigan et al. (2015; 2016) med svak evidens for ingen gradient i kognitiv funksjon (tentativt funn), Ardoy et al. (2014) med ingen evidens for progresjon i kognitiv funksjon.

Hypotese 1 er oppfylt for en rekke kategorier som dokumenterer svak evidens for progresjon i kognitiv funksjon for kroniske kvantitative og kvalitative fysisk aktivitetsintervensjoner. Schmidt et al. (2015); Telles et al. (2013); Tarp et al. (2016); Subramanian et al. (2015); og Aadland et al. (2017)) var basert på den nevrobiologiske hypotesen. Unntaket er Telles et al. (2013) hvor den nevrobiologiske- og den psykososiale hypotesen er underliggende mekanisme. Tarp et al. (2016) og Aadland et al. (2017) har konkurrerende hypoteser som ble vurdert etter prinsippet om «best evidens» (jmf. 5.4 Begrensninger i syntesens robusthet).

Funn med svak evidens vurderes som tentative til tross for at studiene er vurdert med tilstrekkelig metodologisk kvalitet.

Hypotese 1 er oppfylt for kategorien *kognitivt beriket versus ordinær undervisning*.

Resultatene dokumenterer sterk evidens for progresjon i akademisk prestasjon ved kronisk kvalitativ fysisk aktivitetsintervensjon. Have et al. (2016a) og Mullender- Wijnsma et al. (2016) var basert på den nevrobiologiske hypotesen. Have et al. (2016c) målte matematikkferdighet som forutsatte tallforståelse, størrelse, relasjon, addisjon, subtraksjon og geometri. Mullender-Wijnsma et al. (2016) målte matematikkhurtighet og -ferdighet. En vesentlig forskjell fremgår ved undersøkt alder som er hhv. gjennomsnittsalder $M_{\text{alder}} = 7.2$ år (første trinn) og $M_{\text{alder}} = 8.0 - 8.2$ år (andre og tredje trinn). Evidensnivået kan bestrides som følge av heterogen alder, og kan sees i sammenheng med kognitiv utvikling (jmf. 2.2.2. Kognitiv læringsteori kan introdusere en aldersmoderator). Intervensjonen ble alderstilpasset i begge studier. Have et al. (2016a) og Mullender-Wijnsma et al. (2016) har $N \approx 500$ deltagere, og tilstrekkelig metodologisk kvalitet hhv. 72.7% og 86.4%. Med forbehold om nevnte forhold ovenfor, beholdes vurderingen av sterk evidens. Sterk evidens gjør at funnene vurderes robuste.

5.3 Anbefalinger for kroppsøving

Funn med sterk evidens har gitt robuste resultater og kan anbefales for videre utprøving i kroppsøvingstimen. Hypotese 3 (Det systematiske reviewet oppnår en robusthet i evidens som gir implikasjoner for anbefaling av bestemte fysiske aktivitetsintervensjoner.) er oppfylt for kategorien *kognitivt beriket versus ordinær undervisning*. Resultatene dokumenterer sterk evidens for progresjon i *kognitiv funksjon* for kroniske kvalitative fysisk aktivitetsintervensjoner. Hill et al. (2010) målte effekten av et lærerstyrt klasseromsprogram med løp på stedet ved moderat intensitet, hopping til musikk og strekking. Schmidt et al. (2015) målte effekten av et høyintensitets lagspill (gulvball, basketball) spesielt sammensatt for å stimulere eksekutiv funksjon. Programmet var kognitivt krevende med høy intensitet, prospektiv kontroll, kompleks øye-hånd koordinasjon og krav til målrettet atferd. Kognitiv mobilisering ble stimulert gjennom endring av regler underveis i spillet (fløytesignal og visuelt signal – rødt kort indikerte endring av regler). Intervensjonen stimulerte hurtig reaksjon, inhibering og skifte av setting i stor skala, og var tilrettelagt for idrettsspesifikk ferdighetsutvikling (pensum).

Hypotese 3 er oppfylt for kategorien *kognitivt beriket versus ordinær undervisning*.

Resultatene dokumenterer sterk evidens for progresjon i *akademisk prestasjon* ved kronisk kvalitativ fysisk aktivitetsintervensjon. Have et al. (2016c) målte effekten av et klasseroms-integrert fysisk aktivitetsprogram (aktiv matte). Programmet besto av at oppgaver ble løst ved kroppsbevegelse, minst 15 min per time. Mullender- Wijnsma et al. (2016) målte effekten av fysisk aktiv undervisning (F & V lesson). Oppgaver i matematikk og staving ble løst ved bevegelse, eksempelvis multiplikasjon “ 2×4 ” ble løst ved 8 hopp på stedet. Programmet hadde moderat til anstrengende intensitet.

5.4 Begrensinger i syntesens robusthet

Validitet skal kunne klargjøre om resultatene er «sanne», men validiteten er kun basert på visse kriterier for metodologisk kvalitet. Resultatene i syntesen (jmf. 4.4.6 Syntesens robusthet) kan i detalj, avdekke eventuelle svake punkter i robusthet. Siden robusthet vurderes for klinisk relevante resultater vil kun funn som i utgangspunktet har positiv eller negativ gradient bli inkludert. Undersøkelser med ingen gradient er studier hvor det ikke er funnet noen forskjell i intervensjonsulikhet.

Syntesen inneholder tre funn (Costigan et al., 2015) som rapporteres som ikke-signifikante positive funn for eksekutiv funksjon, psykologisk velvære og psykologisk «distress» (lidelse). Intervensjonsulikhet for Costigan et al. (2015) var forsterket intensitet versus kroppsøving (jmf. Tabell 21 i 4.4.6 Syntesens robusthet). Resultatene kan ha vært tilfeldige ved $\alpha = 0.05$.

Syntesen inneholder to tilfeller av konkurrerende hypoteser.

1. Etnier et al. (2014) og Galotta et al. (2015) har funnet progresjon i kognitiv funksjon for hhv. forsterket intensitet versus akademisk og akademisk versus forsterket intensitet (jmf. Tabell 20 i 4.4.6 Syntesens robusthet).
2. Aadland et al. (2017) og Tarp et al. (2016) har funnet progresjon i kognitiv funksjon for hhv. kognitivt beriket versus akademisk og akademisk versus kognitivt beriket (jmf. Tabell 21 i 4.4.6 Syntesens robusthet).

Konkurrerende hypoteser kan avklares ved prinsippet om «best evidens» (Slavin, 1995). Prinsippet har opprinnelse fra jus og sier at et bevis kan forkastes dersom det foreligger et bedre bevis. Dette prinsippet kan overføres til en vitenskapelig sammenheng og har konsekvens for studier med lav metodologisk kvalitet. Studiene kan forkastes til fordel for studier med tilstrekkelig metodologisk kvalitet. Etnier et al. (2014) og Galotta et al. (2015) har

relativt lik metodologisk kvalitet. Aadland et al. (2015) og Tarp et al. (2016) har også relativt lik metodologisk kvalitet. Videre kan studiedesign avgjøre konkurrerende hypoteser. Etnier et al. (2015) har gjennomført et RCT, mens Galotta et al. (2016) et CRT. Begge studier er småstudier med hhv. 43 deltagere fordelt på to grupper og 116 deltagere fordelt på tre grupper. Resaland et al. (2015) og Tarp et al. (2016) har begge gjennomført CRT med hhv. 1129 og 632 deltagere begge med en intervensjons- og kontrollgruppe. Konsensus kan oppnås for Etnier et al. (2014) og Galotta et al. (2014) ved at man aksepterer at småstudier har større variasjon. Disse utgjør bredden i eksempelvis et Funnelplokk (Booth et al., 2016). Ifølge Tarp et al. (2016) var utvalgte skoler allerede involvert i et program for fysisk aktivitet. Som en følge av programmet hevder Tarp et al. (2016) selv at undersøkelsen ikke er generaliserbar til alle skoler i Danmark. Ifølge Aadland et al. (2017) har deres studie en sterk eksternvaliditet og kan generaliseres til alle norske skoler og kanskje også til skoler i resten av verden. Aadland et al. (2017) vurderes med større robusthet enn Tarp et al. (2016) etter prinsippet om «best evidens». Man kan anta at Aadland et al. (2017) kan generaliseres til elever i alder 10 år. Ved å bruke prinsippet om «best evidens» ser man at rekkefølgen av testkriterier er avgjørende for hvilke studier som til sist, trekker det lengste strået. Ifølge Slavin (1995) eksisterer ikke en kriterieliste for vurdering av evidens i vitenskapelig sammenheng.

Syntesen inneholder tre tilfeller av sterk evidens.

1. De kroniske studiene Have et al. (2016) og Mullender- Wijnnsma et al. (2016) har funnet progresjon i akademisk prestasjon for *kognitivt beriket versus akademisk*.
2. Akuttstudiene Chen et al. (2014) og Altenburg et al. (2016) har funnet progresjon i kognitiv funksjon for *forsterket frekvens versus akademisk*.
3. De kroniske studiene Hill et al. (2010) og Schmidt et al. (2015) har funnet progresjon i kognitiv funksjon for *kognitivt beriket versus akademisk*.

Spørsmålet er om dette evidensnivået vil vedvare dersom detaljene i primærstudiene utforskes nærmere. Singh et al. (2012) hevder at heterogenitet eksisterer i detaljene i primærstudiene. For studier med fysisk aktivitetsintervensjoner undersøkte Singh et al. (2012) heterogenitet mellom studier ved å kartlegge studienes:

- 1) måter å vurdere eller måle deltagelse i fysisk aktivitet (ingen av studiene måler elevenes faktiske kognitive stimuli eller akademiske arbeidsinnsats)
- 2) måleinstrumenter for vurdering av mediert kognitiv funksjon og endelig akademisk prestasjon
- 3) utvalg (utvalgsstørrelse, kjønn, alder, sosioøkonomiske forhold og etnisitet)

4) studiedesign (RCT eller CRT)

Heterogenitet i henhold til disse punktene er viktig å inkludere fordi like resultater kan ha fremkommet på ulikt vis.

En vesentlig forskjell mellom Have et al. (2016) og Mullender- Wijnsma et al. (2016) er utvalgets alder som er hhv. Gjennomsnittsalder $M_{\text{alder}} = 7.2$ år (første trinn) og $M_{\text{alder}} = 8.0 - 8.2$ år (andre og tredje trinn). En diskusjon om evidensnivået bør reduseres som følge av heterogen alder, kan sees i sammenheng med kognitiv utvikling (jmf. 2.2.2. Kognitiv læringsteori kan introdusere en aldersmoderator). Intervensjonen ble alderstilpasset i begge studier. En vesentlig forskjell mellom Chen et al. (2014) og Altenburg et al. (2016) er utvalgets alder som er hhv. $M_{\text{alder}} = 9.2$ år og alder = 10 - 13 år. En diskusjon om evidensnivået bør reduseres som følge av heterogen alder, kan sees i sammenheng kognitiv utvikling (jmf. 2.2.2. Kognitiv læringsteori kan introdusere en aldersmoderator). Utvalgene har heterogen sosioøkonomisk bakgrunn. En vesentlig forskjell mellom Hill et al. (2010) og Schmidt et al. (2015) er utvalgsstørrelsen hhv. 1224 og 181 deltagere, alder er hhv. 8 - 11 år og 10 - 12 år. Intervensjonene er heterogene i intensitet med hhv. moderat intensitet i klasserommet og høyintensitets lagspill. Schmidt et al. (2015) har funnet at kategorien forsterket intensitet versus akademisk ikke gir en effekt. Med forbehold om nevnte forhold ovenfor beholdes vurderingen av sterk evidens.

5.5 Sensitivitetsanalyse

Sensitivitetsanalysen er en utforskning av resultatenes følsomhet for valg av primærstudier, metodologisk kvalitet, heterogenitet, «outliers», manglende resultater og eksklusjon av spesielle studier. Sensitivitetsanalysen (Booth et al., 2016) er karakterisert ved utforskning av spørsmålet «hva om?». Analysen inneholder en resultatvurdering med utgangspunkt i seks spørsmål.

5.5.1 Hva om studier ikke fredsstiller eligeringskriterier?

De inkluderte studiene ble vurdert og tilfredsstillte utvelgelseskriteriene. Studiene som ikke ble inkludert hadde følgende begrunnelser: Problemstillingen var ikke relevant (7), utvalget var i alder over 18 år (1), utvalget besto av barn utover normal utvikling (1), artikkelen hadde form som et review eller populærvitenskapelige tidsskrift (3), forskningsdesignet var ikke RCT eller CRT (1), intervensjonen inneholdt ikke fysisk aktivitet (2) (jmf. 3.3 Utvalgelse av litteraturreferanser ved PRISMA's flytdiagram). Konsekvensen av å inkludere disse studiene ville gitt en utvidelse av problemformuleringen. Endring i utvalg ville gitt implikasjoner for en

bredere målpopulasjon. En populærvitenskapelig tilnærming ville gitt manglende informasjon om metodisk prosedyre. Endring av forskningsdesign ville gitt lavere evidensnivå. Dersom intervensjonen ikke inneholder fysisk aktivitet vil også en bredere sammenligningsgrunnlag foreligge. Totalt sett ville inklusjon av ikke-eligerte studier gitt et høyere støynivå ved analyse og konklusjon.

5.5.2 Hva om studier med lav metodologisk kvalitet ekskluderes?

Studier med lav metodologisk kvalitet ble inkludert i syntesen (jmf. Tabell 18 i 4.4.2 Kroniske fysiske aktivitetsintervensjoner). Undersøkelsene blant de akutte studiene hadde alle tilstrekkelig metodologisk kvalitet. Blant de kroniske studiene hadde en del undersøkelser lav metodologisk kvalitet (9 av 26 funn). Konsekvensen av å ekskludere disse undersøkelsene ville redusert antall kategorier. Tolv kategorier kan reduseres ved at fire kategorier har lav metodologisk kvalitet (jmf. Tabell 18 i 4.4.2 Kroniske fysiske aktivitetsintervensjoner). Åtte kategorier intervensjonsulikhet kan ytterligere reduseres ved eksempelvis eksklusjon av ikke-signifikante resultater (2 kategorier ikke-signifikante). Metodologisk kvalitet var en del av evidensvurderingen (jmf. 4.4.6 Syntesens robusthet). Studier med lav metodologisk kvalitet kan per definisjon (jmf. 3.5.3 Kategorisert metodologisk kvalitet og gradert evidens) supplere studier med tilstrekkelig metodologisk kvalitet for evidensnivå 2, 3 og 4.

5.5.3 Hva om studier er heterogene?

Heterogenitet ble sammenfattet for syntesen (jmf. 4.4.5 Heterogenitet i synteseplanet). Heterogenitet har betydning for presentasjon av data. Data ble presentert i akutte og kroniske intervensjoner. I syntesen ble data presentert i ulike kategorier. Harvestplottet introduserte spesielle hensyn for ulike intervensjonsulikheter og hypotesen om positiv, ingen eller negativ gradient i utfall. Tolkningen av resultatene ble gjennomført for hvert enkelt funn.

Heterogenitetsvurderingen er spesielt nyttig i de tilfeller der undersøkelse peker på samme resultat og man mistenker at resultater har fremkommet på ulikt vis (jmf. 5.4 Begrensinger i syntesens robusthet). Chen et al. (2014) og Altenburg et al. (2016) fant en positiv effekt på kognitiv funksjon ved forsterket frekvens versus akademisk. Disse studiene målte hhv. positiv eksekutiv funksjon (inhibering, arbeidsminne og skifte av setting) og selektiv oppmerksomhet. Dette er i samsvar med teori (jmf. 2.3.1 Eksekutive funksjoner) som tolker

disse målingene som en del av styrt oppmerksomhet og kognitiv fleksibilitet. Resultatene er ulike, men ikke tilstrekkelig ulike til å hevde redusert robusthet i evidens.

5.5.4 Hva om ett resultat er veldig forskjellig fra resten?

Synteseplanet gir en utmerket oversikt over alle data, og ingen data i synteseplanet avviker spesielt fra resten av resultatene. En primærstudie av Costigan et al. (2015) fant en positiv effekt på psykologisk velvære ved forsterket intensitet. Tolkningen er forankret i den psykososiale hypotesen (jmf. 2.1.2 Den psykososiale hypotesen), vesentlig forskjellig fra den nevrobiologiske forklaringen. I synteseplanet behandles forklaringsmekanismer likt.

5.5.5 Hva om resultater mangler?

Manglende resultater kan eksistere av flere årsaker. Konsekvensen av manglende resultater kan være feilaktige konklusjoner. Søkeprosessen foregår i et begrenset tidsrom. Studier som publiseres etter og utover tidsbegrensningen på 10 år ekskluderes fra reviewet. Det finnes ikke en standard for hvordan primærstudiene skal presentere resultater, og forskningsfeltet er bundet av midlertidighet hvor stadig ny kunnskap med sterkere evidens avdekkes. Hvert primærstudie rapporterer et antall resultater. Dersom antallet er 20 eller flere kan man sannsynliggjøre liten grad av rapporteringsavvik. Videre kan Funnelplokk sannsynliggjøre liten grad av publikasjonsavvik ved symmetrisk fordeling av effektstørrelse i resultatene. Asymmetri er indikasjon på manglende resultater eller publikasjonsavvik. En kvalitativ fremstilling ved Harvestplokk gir et inntrykk av om småstudier ($N < 100$ per gruppe) er asymmetrisk omkring en «sann» effektstørrelse. I det videre blir eksistensen av manglende resultater undersøkt nærmere (jmf. 5.7 Identifisering av mangler i primærstudienes resultater og 5.8 Identifisering av mangler som kan relateres til publikasjonsavvik).

5.5.6 Hva om ett studie med spesiell innflytelse på resultatene ekskluderes?

Et eksempel på et funn med spesiell innflytelse finnes blant store studier (jmf. 4.4.7 En «sann» effektstørrelse). Hill et al. (2010) fant en positiv effekt på psykomotoriske tester ved kategorien kognitivt beriket versus akademisk i løpet av en 2 ukers RCT – crossover studie. Tilsvarende robusthet i evidens er ikke funnet for kognitiv funksjon (jmf. Tabell 21, 4.4.6 Syntesesens robusthet). Have et al. (2016c) fant progresjon i akademisk prestasjon, i en studie med 500 deltagere. Tilsvarende robusthet i evidens er ikke funnet for akademisk prestasjon (jmf. Tabell 19, 4.4.6 Syntesesens robusthet). Bortfall av Hill et al. studien og Have et al.

(2016c) ville redusere evidensnivået for undersøkelsene (jmf. 5.4 Begrensinger i syntesens robusthet).

5.6 Identifisering av mangler i søk og utvelgelse av studier

Ifølge Booth et al. (2016) bør et systematisk review rapportere en vurdering av validitet, en sensitivitetsanalyse og identifiserte mangler i datagrunnlaget. En totalvurdering av innsamlede data uttrykker hvordan primærstudiene kan influere på forskningsspørsmålet og den praktiske anvendelsen av forskningsresultatene.

5.6.1 Har alle berettigede studier blitt identifisert?

Søkeprosessen ble gjennomført i et begrenset tidsrom i to perioder, uke 24 - 25 og uke 32 2017. Supplerende litteratur har blitt vurdert også med tanke på inklusjon i rapporten. Enkelt studier har blitt inkluderte enten etter råd fra veileder eller som ufyllende informasjon til allerede inkluderte studier. Selv om rapporten har vært åpen gjennom masterprosjektet, har likevel ikke nye artikler blitt lagt til. Det ble besluttet å håndtere studieutvelgelse ved å akseptere artikler som ville ha kritisk betydning for konklusjonen (Randolf, 2009), og for øvrig akseptere at ingen søk er feilfrie. Ifølge Booth et al. (2016) gir Cochrane Collaboration ut en oppdatering av reviewer hvert andre år, dermed blir nye primærstudier godt ivaretatt.

5.6.2 Hvordan blir forskningsspørsmålet adressert i primærstudiene og blir spørsmålet besvart i konklusjonen?

Alle inkluderte studier har en relevant problemstilling. Studiene har til hensikt å bedre kunnskapen om effekten av akutt og kroniske fysisk aktivitet på kognitiv funksjon og akademisk prestasjon. Hensikten er i tillegg å vektlegge metodologisk kvalitet ved gjennomføringen av studier og om mulig å kunne fremme anbefalinger om bestemte typer fysisk aktivitet. I resultatkapittelet (jmf. 4.1 Beskrivelse av inkluderte studier) fremgår studienes praktiske adressering av problemstilling med tilhørende konklusjoner.

5.6.3 Ble primærstudier feilaktig akseptert eller ekskludert?

Inklusjonskriterier ble definerte for inkluderte studier. Metodekapittelet (jmf. 3.1 Inklusjonskriterier) beskriver inklusjonskriterier som omhandler populasjon, utvalg, intervensjonsaktivitet, kontrollaktivitet, utfall og studiedesign. I metodekapittelet (jmf. 3.3 Utvelgelse av litteraturreferanser ved PRISMA flytdiagram) beskrives hvordan tittel, sammendrag og hele teksten blir gjenstand for vurdering. Metodekapittelet beskriver hvordan

de ulike stegene i utvelgelsesprosessen har foregått. For å unngå avvik som relateres til feilaktig aksept eller eksklusjon av primærstudier har de inkluderte studiene blitt lest flere ganger. De ekskluderte studiene har blitt avskjediget med dokumentert begrunnelse (jmf. 3.3 Utvelgelse av litteraturreferanser ved PRISMA flytdiagram). Et systematisk ekstrakt av innsamlede data dokumenterer datagrunnlaget for reviewet (jmf. 4.1 Beskrivelse av inkluderte studier).

5.7 Identifisering av mangler i primærstudienes resultater

Manglende rapportering av måleresultater kan favorisere en studie eller en bestemt intervensjon fremfor en annen og dermed føre til feilaktige konklusjoner (Williamson & Gamble, 2005). Manglende resultater kan identifiseres på flere måter:

5.7.1 Har primærstudier unnlatt å rapportere sentrale resultater?

Sentrale resultater er rapportert i alle primærstudier, men det eksisterer ikke en standard for hva som bør rapporteres av primær og sekundær resultater. Eks. Have et al. (2016c) rapporterte progresjon i akademisk prestasjon. Måleinstrumentet for akademisk prestasjon var en 45 minutter lang standardisert matematikk-test. Mullender-Wijnsma et al. (2016) rapporterte progresjon i akademisk prestasjon. Måleinstrumentet for akademisk prestasjon var hurtighet og ferdighet i matematikk og lesing og staving innen språkfag.

5.7.2 Mangler reviewet måleresultater?

I reviewet er hovedfunn fra primærstudier rapportert. Syntesen har presentert utfall enten som medierte eller endelige resultater (kognitiv funksjon eller akademisk prestasjon). Sammenligningsgrunnlaget er blitt rapportert som en intervensjonsulikhet. Ett funn ble rapportert for studier med en intervensjonsgruppe og en kontrollgruppe. I Etnier et al. (2014) ble to funn rapportert. Der er målinger gjennomført umiddelbart etter intervensjon og 24 timer senere. Tre funn er rapportert i studier med to intervensjonsgrupper og en kontrollgruppe. Disse undersøkelsene utgjør hovedfunn i syntesen. Delanalyser i hver primærstudie er ikke blitt rapportert. Abstraksjonsnivået i syntesen med medierte og endelige resultater har gjort det mulig å sammenfatte alle primærstudier.

5.7.3 I hvilken grad rapporterer primærstudier medierte og endelig utfall?

Alle primærstudier inneholder enten måling av kognitiv funksjon eller akademisk prestasjon, men med ulik detaljeringsgrad og forskjellige måleinstrumenter (jmf. Vedlegg 1: 8.1.3

Administrering av testinstrumenter). Primærstudiene rapporterer ulike måleinstrumenter. Studiene rapporterer forskjellige kognitive funksjoner med delkomponenter for kognitiv funksjon eller akademisk prestasjon. Schmidt et al. (2015) rapporterer måling av eksekutiv funksjon (oppdatering, inhibering, skifte av setting). Måleinstrumentene var hhv. en «n-back task», en «Flanker task» og en «mixed block Flanker task». Detaljene i utfallsdata gir heterogenitet i primærstudiene, men sammenligningsgrunnlaget (tre intervensjonsulikheter) ble vurdert som det sentrale i vurderingen av det totale datagrunnlaget. Schmidt et al. (2015) som har påvist en positiv effekt ved «skifte av setting» har fått rapportert en positiv gradient for mediert kognitiv funksjon i synteseplanet.

Hill et al. (2010) rapporterte måling av psykomotoriske parametere. Måleinstrumentene var «paced serial addition task», «size ordering task», «listening span task», «digit span backward task» og «digit symbol encoding task» som måler informasjonsprosessering, arbeidsminne, korttidsminne og svekket kognitiv funksjon. Disse måler dels eksekutive domener, men også korttidshukommelse og endring i kognitiv funksjon. Hill et al. (2010) rapporterte en positiv effekt for alle tester i andre uke, og forklarer dette med tilvenning til testsituasjonen. Først i andre uke var testene vel kjent. I synteseplanet ble mediert kognitiv funksjon presentert som funn.

5.7.4 Hva er antallet tillagte funn rapportert i primærstudier?

Ifølge Booth et al. (2016) vil mer enn 20 rapporterte funn indikerer liten sannsynlighet for rapporteringsavvik i primærstudiene. Høy sannsynlighet for rapporteringsavvik fins ved 5 – 10 funn, og 10 – 20 funn indikerer en moderat sannsynlighet for rapporteringsavvik. Syntesen rapporterer 37 separate funn som gjelder for akutt og kronisk fysisk aktivitetsintervensjoner. Et overslag over totalt rapporterte funn i primærstudiene inkluderte deskriptiv statistikk og måleresultater. Akuttstudiene hadde gjennomsnittlig 43 tillagte funn og de kroniske studiene hadde gjennomsnittlig 144 tillagte funn. Williamson & Gamble (2005) gjør en grundig redegjørelse for rapporteringsavvik i metaanalyser.

5.7.5 Hvilke utfall brukes til å beregne utvalgsstørrelse i primærresultater?

Akuttstudiene dokumenterte ikke en beregnet utvalgsstørrelse. Akuttstudiene hadde en utvalgsstørrelse $N < 100$ per gruppe. Utgangspunktet for estimert utvalgsstørrelse var en RCT med $N = 206$ (fordelt på to grupper) (Resaland et al., 2015). Blant kroniske studier dokumenterte 7 av 13 en beregnet utvalgsstørrelse. Totalt hadde 59 % ikke dokumentert

beregnet utvalgsstørrelse. Resaland et al. (2015) beregnet utvalgsstørrelse for CRT basert på effektstørrelse Cohen's $D = 0.35$, $ICC = 0.15$, med $n = 16.2$ deltagere per klasse etter antatt 20% drop out. Resaland et al. (2015) beregnet en utvalgsstørrelse $N = 936$ (fordelt på 2 grupper) ved måling av akademisk prestasjon. Aadland et al. (2017) baserte sin påfølgende analyse av eksekutiv funksjon på Resaland et al. (2015) sin beregning av utvalgsstørrelse. Differansen mellom utvalgsstørrelser for RCT og CRT skyldes et relativt strengt krav ved type 2 feil og en forutsetning om høy korrelasjon ($ICC = 0.15$) mellom deltagere. Det er god grunn til å vurdere CRT design inkludert beregning av utvalgsstørrelse ved planlegging av fysisk aktivitetsintervensjoner. CRT design tar inn over seg forskningsoppsettets natur med deltagere som befinner seg i allerede eksisterende klasser. Primærstudiene viser store forskjeller ved beregnet utvalgsstørrelse, med tilsvarende stor variasjon i forutsetning for beregningen. Mullender-Wijnsma et al. (2016) forutsatte $\alpha = 0.05$, $1 - \beta = 0.80$, moderat $ICC = 0.10$, antatt effektstørrelse $= 0.44$, og beregnet minste utvalgsstørrelse $N \geq 500$ (≥ 20 deltagere i 25 grupper).

5.7.6 Er det samsvar i primærstudienes konferansebidrag og siste publiserte resultater?

Noen studier er sammensatt av flere artikler som beskriver planlegging og gjennomføring av forsøket (konferansebidrag), og videre artikler som beskriver resultater (endelige publikasjoner).

Nesten samtlige primærstudier erkjenner et problem med manglende kontroll av gjennomført fysisk aktivitet. Målinger av den fysiske aktiviteten er enten subjektiv selv-rapportert aktivitet eller objektive registreringer i et kort tidsrom i løpet av forsøkene. Analysedelen av eksperimentet kan tilpasses i de tilfeller hvor faktisk fysisk aktivitet er kjent. Hoveddelen av primærstudiene har ikke løst problemet ved å endre analyser underveis.

Et unntak er Aadland et al. (2017) som utførte en «per protokol» analyse for deltagere som fulgte planlagt treningsprotokoll. Resaland et al. (2015) beskrev hensikten med ASK programmet, studiedesign, intervensjon og metodiske detaljer. Resaland et al. (2016) rapporterte resultater med ingen effekt på akademisk prestasjon (nasjonale tester). I Aadland et al. (2017) ble ASK analysert med hensyn til eventuelle effekter på eksekutiv funksjon. Studien avdekker en signifikant forskjell mellom planlagt og faktisk gjennomført program. En «intention to treat» og en «per protokol» analyse ble gjennomført og sistnevnte påviste en progresjon i mediert kognitiv funksjon. Deltagere som enten hadde $< 80\%$ av

treningsmengden i intervensjonsgruppen eller >120% av treningsmengden i kontrollgruppen ble ekskludert. Innledende publikasjoner samsvarer med sluttresultatet. Underveis ble ulike statistiske analyser prøvd ut. Studien introduserer selv et problem knyttet til drop out som har betydning for generalisering og kausalitet.

5.7.7 Er syntesen analysert?

Syntesen er utforsket ved at det er utviklet en teoretisk modell for sammenhengen mellom intervensjonsulikhet og mediert eller endelig resultat. Modellen er tro mot de empiriske resultatene som foreligger i reviewet. Mønstre i syntesen er identifisert og analysert. Heterogenitet er kartlagt i synteseplanet, og i tillegg diskutert som en del av vurdering av robusthet i evidens. Syntesens robusthet er analysert med utgangspunkt i klinisk relevans.

5.7.8 Er funn og konklusjon presentert med en klar sammenheng?

Sammenhengen mellom funn og konklusjon blir tydelig ved at konklusjonen besvarer innledende forskningsspørsmål med tilhørende tre hovedhypoteser. Syntesens robusthet besvarer hypotese 1: Fysisk aktivitetsintervensjoner vil styrke kognitiv funksjon eller akademisk prestasjon sammenlignet med ordinær undervisning eller kroppsøving. Den metodologiske kvalitetsvurderingen og diskusjon av validitet besvarer hypotese 2: Primærstudiene har tilstrekkelig metodologisk kvalitet (> 60%). Diskusjonen av sterk evidens og syntesens robusthet gir svar på hypotese 3: Det systematiske reviewet oppnår en robusthet i evidens som gir implikasjoner for anbefaling av bestemte fysisk aktivitetsintervensjoner.

5.8 Identifisering av mangler som kan relateres til publikasjonsavvik

Publikasjonsavvik er definert som en familie av avvik som inkluderer språklig avvik, lokaliseringsavvik og databaseavvik (Booth et al., 2016). Ikke identifiserte publikasjoner kan identifiseres på flere måter:

5.8.1 Er alle rapporterte resultater signifikante?

Dersom alle resultater er signifikante kan ifølge Booth et al. (2016) signifikans være en indikasjon på selektiv rapportering. Syntesen viser at 8 % (3 av 37) funn er ikke-signifikante. Costigan et al. (2015) rapporterte tre ikke- signifikante funn. Primærstudiene rapporterte gjennomgående ikke-signifikante resultater for tillagte funn. Galotta et al. (2015) rapporterte ikke-signifikant forskjell for baselineverdier mellom jenter og gutter. Schmidt et al. (2015)

rapporterte ikke-signifikante forskjeller i alder og kjønn for deskriptive data. Dersom man antar at totaliteten av forskningsarbeider (signifikant og ikke-signifikante arbeider) er likefordelt kan en lav andel ikke-signifikante resultater indikere publikasjonsavvik.

5.8.2 Kan Harvestplot indikere publikasjonsavvik?

Harvestplottet for småstudier med $N < 100$ deltagere per gruppe er en tilnærming til Funnplotet. Funnplotet kan gi en indikasjon på publikasjonsavvik ved asymmetri. I Harvestplottet fordeles funn med negativ gradient, ingen gradient og positiv gradient med andel hhv. 14.3%, 23.8% og 61.9 % for småstudier. Dette viser en tydelig asymmetri ved at negativ gradient er underrepresentert i forhold til positiv gradient. En del av skjevfordelingen kan forklares med valg av rekkefølge for kategorier. Eksempel: Ardoy et al. (2014) fant en positiv effekt for forsterket intensitet versus forsterket frekvens. Forsterket intensitet vs. Forsterket frekvens har et funn med positiv gradient. Dette funnet kan rapporteres i motsatt rekkefølge som Forsterket frekvens vs. Forsterket intensitet, og på den måten gi en negativ gradient. Maksimal omfordeling på rekkefølge for 4 kategorier gir et balansert inntrykk (7 funn med negativ gradient og 9 funn med positiv gradient). Med forbehold om nevnte forhold kan ikke metoden benyttes til å fastslå publikasjonsavvik i inkluderte studier.

5.8.3 Har publikasjoner blitt tilført ved kontakt med forfattere?

Upublisert materiell kan tilføre reviewet informasjon som ellers ikke er tilgjengelig. Have et al. (2016a) har publisert et konferansebidrag: Rationale and design of a randomized controlled trial examining the effect of classroom-based physical activity on math achievement. Ved kontakt med forfatter Mona Have (per mail, September 2017) fremkom en ublisert doktoravhandling som inneholder den publiserte artikkelen samt en upublisert artikkel: A classroom-based physical activity intervention improves children`s math achievement – A randomized controlled trial. Resultatene fra det upubliserte materialet ble tillagt øvrige resultater i reviewet, og avviker ikke i forhold til de publiserte resultatene.

5.8.4 Andre bidrag til publikasjonsavvik

Ifølge Booth et al. (2016) kan en rekke problemstillinger diskuteres i forbindelse med rapporteringsavvik. Rapporten er begrenset ved at en del vurderinger er utelatt. Det gjelder vurderingen av om konferansebidrag har betydning for siste publikasjon. Er det samsvar i resultater, utvalgsstørrelse, statistiske metoder mellom konferansebidrag og siste publikasjon?

«Citations» av relevante artikler ble ekskludert, dermed eksisterer heller ikke «citation bias», det vi si publikasjonsavvik fra artikler med resultater som peker i samme retning. Databasene gir innpass til visse journaler mens andre journaler ekskluderes. En overvekt av etablerte journaler representerer databaseavvik. Primærstudienes Journaler ble heller ikke undersøkt nærmere. Studiene ble ikke undersøkt for eventuelle sponsorer/ bidragsytere. Dermed ble ikke «funding bias» avdekket. Grå litteratur består av publikasjoner som er selvstendige og ikke kontrolleres av kommersielle forleggere. Grå litteratur ble ikke undersøkt. Dermed ble det ikke avdekket om resultater fra grå litteratur er signifikant forskjellig fra inkluderte studier.

5.9 Styrker og begrensninger

Styrken ved reviewet er at syntesen er en kompilering av data til et abstraksjonsnivå hvor alle funn kan presenteres. Syntesen presenterer kognitiv funksjon og akademisk prestasjon med tilhørende metodologisk kvalitet. Detaljer i måleresultatene forblir i primærstudiene. Videre er det utviklet en modell som fungerer som en støtte til kausalsammenhengen mellom fysisk aktivitet og kognitiv funksjon og akademisk prestasjon. Modellen har støtte i empiriske resultater. Den nevrobiologiske hypotesen eller den psykososiale hypotesen støtter også kausalitet.

Begrensninger i reviewet er at kun 35%, 6 av 17 primærstudier er gjennomført med tilstrekkelig utvalgsstørrelse ($N \geq 500$). Utvalgsstørrelse utgjør et av flere sentrale kriterier ved vurdering av generalisering. Små utvalg ($N < 100$ per gruppe) benyttes i 47%, 8 av 17 studier. De kroniske studiene har relativt kort varighet. Kun 17.4% (3 av 17 studier) varer lengre enn 1 år. Reviewet har få målinger av akademisk prestasjon. Endelige funn (akademisk prestasjon) er rapportert for 24.3% (9 av 37 funn) av disse har kun 10.8% (4 av 37 funn) tilstrekkelig metodologisk kvalitet.

Styrken ved reviewet er inklusjonskriteriet som omhandler forskningsdesign. Kun RCT og CRT design ble inkludert. Den metodologiske kvaliteten er tilstrekkelig for en stor andel av primærstudiene. Det viser seg likevel at gjennomføringen av fysisk aktivitetsintervensjoner innebærer risiko for avvik i form av redusert eksterntvaliditet, utvalgsstørrelse og internvaliditet. Særlig uttalte kriterier ved vurdering av eksterntvaliditet er: drop out, midlertidighet, alder og den reaktive effekten ved gjennomføring av eksperimenter uten blinding. Kriterier som utpeker seg ved vurdering av internvaliditet er: seleksjonsavvik - konfundere, som følge av forhåndsforskjeller mellom intervensjon og kontrollgruppe. En

mulig konfunder kan introduseres ved manglende randomisering av deltagere. Drop out har innvirkning på kausalitet og generalisering ved at utvalgets karakteristika endres.

Sensitivitetsanalysen med identifisering av mangler har spesielt vektlagt mangler i søk og utvelgelse av studier, primærstudienes resultater og mangler som kan relateres til publikasjonsavvik. Analysen avdekker at primærstudiene er «sunne» på disse områdene.

Måleresultater i primærstudiene er ulike. Reviewet inneholder primærstudier med forskjellige intervensjoner for ulike utvalg barn og unge i ulike populasjoner. Måleinstrumenter og metoder er gjort på ulikt vis (Vedlegg 1: 8.1.3 Administrering av testinstrumenter). Måleinstrumentene introduserer begrensninger ved målefølsomhet for empiriske undersøkelser. Instrumentene fungerer innenfor et gitt «arbeidsområde». Heterogenitet blir adressert ved å presentere resultater med homogene kategorier intervensjonsulikhet. Inndeling i «homogene» kategorier eliminerer ikke heterogenitet i stimuli og utfall.

6.0 Konklusjon

Reviewet inkluderer kun studier med RCT og CRT design og oppnår på den måten høyeste nivå av evidens (Polit & Beck, 2014, s. 23). Koder er utviklet for å beskrive differansen mellom to eller tre allokerte grupper. Koder av intervensjonsulikhet danner kategorier som presenteres i syntesen. Kategoriene blir førende for hvilke slutninger som kan trekkes og hvilke anbefalinger som kan fremmes. Syntesen avbildes på et abstraksjonsnivå hvor alle funn kan presenteres. Videre er det utviklet en modell som fungerer som en støtte til kausalitet mellom stimuli og mediert eller endelig utfall. Modellen tar hensyn til empiriske resultater. Den nevrobiologiske hypotesen eller den psykososiale hypotesen støtter også kausalitet. Gradert evidens fremmes for å styrke undersøkelsens robusthet i evidens ved sammenstilling av klinisk relevante funn. Generalisering av resultater må gjøres for hvert primærstudie med forbehold om at generaliseringskriteriene er oppfylt. Kausalitet kan argumenteres for, men må vurderes med varsomhet i hvert enkelt primærstudie med forbehold om at kriteriene for internvaliditet er oppfylt (jmf. 5.1.3 Internvaliditet). Sensitivitetsanalysen med identifisering av mangler har spesielt vektlagt mangler i søk og utvelgelse av studier samt primærstudienes resultater og mangler som kan relateres til publikasjonsavvik. Analysen avdekker at primærstudiene er «sunne» på disse områdene.

Konklusjonen besvarer forskningsspørsmålet, hovedhypoteser (jmf. 1.2 Formulering og drøfting av problemstilling), og fremmer implikasjoner for fremtidige arbeider.

6.1 Hovedhypotese 1

Hovedhypotese 1: Fysiske aktivitetsintervensjoner vil styrke kognitiv funksjon (eks. eksekutiv funksjon) eller akademisk prestasjon sammenlignet med ordinær undervisning eller kroppsøving.

Hypotese 1 er oppfylt for progresjon i *kognitiv funksjon* (positiv gradient) etter akutt fysisk aktivitetsintervensjoner målt i 63.6% av undersøkelsene (7 av 11). Ingen gradient og negativ gradient er hver målt i 18.2% av undersøkelsene (2 av 11).

Hypotese 1 er oppfylt for progresjon i *kognitiv funksjon* (positiv gradient) etter kronisk fysisk aktivitetsintervensjon målt i 38.5% av undersøkelsene (10 av 26). Ingen gradient er målt i 23% av undersøkelsen (6 av 26). Negativ gradient er målt i 3.8% av undersøkelsene (1 av 26). Hypotese 1 er oppfylt for progresjon i *akademisk prestasjon* (positiv gradient) etter kronisk fysisk aktivitetsintervensjon målt i 11.5% av undersøkelsene (3 av 26). Ingen gradient er målt i 23% av undersøkelsene (6 av 26). Ingen undersøkelser har negativ gradient (0 av 26).

Hypotese 1 er oppfylt for en rekke kategorier intervensjonsulikhet. I diskusjonskapittelet (jmf. 5.2 Progresjon i kognitiv funksjon og akademisk prestasjon) presenteres homogene kategorier med en beskrivelse av evidensnivå for progresjon i kognitiv funksjon og akademisk prestasjon ved akutt og kronisk fysisk aktivitetsintervensjoner. De fleste studiene er basert på den nevrobiologiske hypotesen og enkelte studier er i tillegg basert på den psykososiale hypotesen. Et funn kan relateres til atferdshypotesen.

6.2 Hovedhypotese 2

Hovedhypotese 2: Primærstudiene har tilstrekkelig metodologisk kvalitet. Tilstrekkelig metodologisk kvalitet krever minst 60% sannsynlighet for at primærstudiene er fri for metodologiske avvik.

Hypotese 2 er oppfylt for 76.5% av RCT og CRT studiene. Kvalitetskriteriene for studier med tilstrekkelig metodologisk kvalitet har svakheter på vesentlige områder. Eksternvaliditet, utvalgsstørrelse og internvaliditet blir diskutert i diskusjonskapittelet (jmf. 5.1 Vurdering av

validitet). Eksternvaliditet og utvalgsstørrelse har betydning for generaliserbarhet og internvaliditet har betydning for kausalitet.

6.3 Hovedhypotese 3

Hovedhypotese 3: Det systematiske reviewet oppnår en robusthet i evidens som gir implikasjoner for anbefaling av bestemte fysiske aktivitetsintervensjoner.

Hypotese 3 er oppfylt for 4 studier (Hill et al., 2010; Schmidt et al., 2015; Have et al., 2016c; Mullender-Wijnsma et al., 2016). Studiene har funn med sterk evidens for progresjon i kognitiv funksjon og akademisk prestasjon (5.3 Anbefalinger av fysisk aktivitetsintervensjoner i kroppsøving). Hill et al. (2010) målte effekten på kognitiv funksjon med et lærerstyrt klasseromsprogram med løp på stedet ved moderat intensitet, hopping til musikk og strekking. Schmidt et al. (2015) målte effekten av et høyintensitets lagspill (gulvball, basketball) spesielt sammensatt for å stimulere eksekutiv funksjon. Have et al. (2016c) målte effekten på akademisk prestasjon med et klasseromsintegreert fysisk aktivitetsprogram (aktiv matte). Programmet besto av at oppgaver ble løst ved kroppsbevegelse, minst 15 min per time. Mullender- Wijnsma et al. (2016) målte effekten av fysisk aktiv undervisning (F & V lesson). Oppgaver i matematikk og staving ble løst ved bevegelse. Det systematiske reviewet oppnår en robusthet i evidens som gjør det mulig å anbefale intervensjoner hvor fysisk aktivitet er kognitivt beriket.

6.4 Implikasjoner for fremtidige arbeider

Reviewet viser at kun 35%, 6 av 17 primærstudier er gjennomført med tilstrekkelig utvalgsstørrelse ($N \geq 500$). Utvalgsstørrelse utgjør et av flere sentrale kriterier ved vurdering av generalisering. Små utvalg ($N < 100$ per gruppe) benyttes i 47%, 8 av 17 studier. De kroniske studiene har relativt kort varighet. Kun 17.4% (3 av 17 studier) varer lengre enn 1 år. Reviewet har få målinger av akademisk prestasjon. Endelige funn (akademisk prestasjon) er rapportert for 24.3% (9 av 37 funn) av disse har kun 10.8% (4 av 37 funn) tilstrekkelig metodologisk kvalitet.

Det systematiske databasesøket ble innrettet mot RCT design. Et tilsvarende søk innrettet mot CRT design vil kanskje øke andelen studier med en presis utvalgsstørrelse. En innretning mot studier med lengre varighet vil gi et mer omfattende grunnlag for å vurdere effekten på akademisk prestasjon.

Kausalitet mellom fysisk aktivitet og kognisjon eller akademisk prestasjon kan sannsynliggjøres ved å styrke internvaliditet. Internvaliditet kan konstrueres gjennom systematisk kartlegging og kontroll av konfundere. Fysisk aktivitetsintervensjoner har ikke løst problemet med blinding og smitte. Blinding og smitte har betydning for generalisering og kausalitet. Et mulig fremtidig arbeidsområde kan være å utvikle metodologiske kriterier basert på studienes merknader for internvaliditet eller eksternvaliditet (jmf. Tabell 24 og 25 i 5.1.3 Internvaliditet), og utvikle en metodologisk sjekklister som er spesielt tilpasset fysisk aktivitetsintervensjoner.

Syntesen presenterer en samling resultater i form av målt kognitiv funksjon eller akademisk prestasjon. En videreutvikling kan inneholde studiedesign eller komponenter fra delanalyser i primærstudiene (eks. kjønn). Modellen som er utviklet på bakgrunn av syntesen kan også videreutvikles til å inneholde delkomponenter.

Kategorien *forsterket frekvens versus ordinær undervisning* dokumenterer sterk evidens for progresjon i kognitiv funksjon ved akutt kvantitativ og kvalitativ fysisk aktivitetsintervensjon. Chen et al. (2014) og Altenburg et al. (2016) kan repliseres med et CRT design og en robust utvalgsstørrelse. På den måten er en eventuell generalisering av resultatene mulig.

7.0 Referanser

- Aadland, K. N., Ommundsen, Y., Anderssen, S. A., Brønnick, K. S., Moe, V. F., Resaland, G. K., Skrede, T., Stavsmo, M. & Aadland, E. (2017). Effects of the Active Smarter Kids (ASK) Physical Activity School-based Intervention on Executive Functions: A Cluster-Randomized Controlled Trial. *Scandinavian Journal of Educational Research*, 07/11/2017, 1-15. doi: 10.1080/00313831.2017.1336477
- Ahamed, Y., MacDonald, H., Reed, K., Naylor, P. J., Liu-Ambrose, T. & McKay, H. (2007). School-Based Physical Activity Does Not Compromise Children's Academic Performance. *Medicine & Science in Sports & Exercise*, 39(2), 371-376. doi: 10.1249/01.mss.0000241654.45500.8e
- Altenburg, T. M., Chinapaw, M. J. M. & Singh, A. S. (2016). Effects of one versus two bouts of moderate intensity physical activity on selective attention during a school morning in Dutch primary schoolchildren: A randomized controlled trial. *Journal of Science and Medicine in Sport*, 19(10), 820-824.
- Álvarez-Bueno, C., Pesce, C., Cavero-Redondo, I., Sánchez-López, M., Martínez-Hortelano, J. A. & Martínez-Vizcaíno, V. (2017). The Effect of Physical Activity Interventions on Children's Cognition and Metacognition: A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, 56(9), 729-738.
- American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, D.C.: American Psychological Association. Hentet fra <http://kildekompasset.no/referansestiler/apa-6th.aspx>
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, 8(2), 71-82. doi: 10.1076/chin.8.2.71.8724
- Arday, D. N., Fernandez- Rodriguez, J. M., Jimenez-Pavon, D., Castillo, R., Ruiz, J. R. & Ortega, F. B. (2014). A Physical Education trial improves adolescents' cognitive performance and academic achievement: the EDUFIT study. *Scandinavian Journal of Medicine & Science in Sports*, 24(1), 52-61. doi: 10.1111/sms.12093
- Astill, R. G., Van der Heijden, K. B., Van Ijendoorn, M. H. & Van Someren, E. J. W. (2012). Sleep, Cognition, and Behavioral Problems in School-Age Children: A Century of Research Meta-Analyzed. *Psychological Bulletin*, 138(6), 1109-1138. doi:10.1037/a0028204
- Berghmans, L. C. M., Hendriks, H. J. M., Bø, K., Hay-Smith, E. J., de Bie, R. A. & van Waalwijk van Doorn, E. S. C. (1998). Conservative treatment of stress urinary incontinence in women: A systematic review of randomized clinical trials. *British Journal of Urology*, 82(2), 181-191.
- Best, J. R., Miller, P. H. & Jones, L. L. (2009). Executive Functions after Age 5: Changes and Correlates. *Developmental Review*, 29(3), 180-200.
- Best, J. R. (2010). Effects of physical activity on children's executive function: Contributions of experimental research on aerobic exercise. *Development Review*, 30 (4), 331- 351.

- Best, J. R. & Eccles, J. (2012). Exergaming Immediately Enhances Children's Executive Function. *Developmental Psychology*, 48(5), 1501-1510.
- Binder, D. K. & Scharfman, H. E. (2004). Brain-derived Neurotrophic Factor: A Mini Review. *Growth Factors*, 22(3), 123-131.
- Birkeland, P., Venheim, R. & Breiteig, T. (2011). Læring og Undervisning. I Birkeland, P., Venheim, R. & Breiteig, T (Red.). *Matematikk for lærere: 1* (22 - 66). Oslo: Universitetsforlaget.
- Booth, A., Sutton, A. & Papaioannou, D. (2016). *Systematic approaches to a successful literature review* (2nd ed.). Los Angeles, Calif: Sage.
- Bragg, M. A., Tucker, C. M., Kaye, L. B. & Desmond, F. (2009). Motivators of and Barriers to Engaging in Physical Activity: Perspectives of Low-Income Culturally Diverse Adolescents and Adults. *American Journal of Health Education*, 40(3), 146-154. doi: 10.1080/19325037.2009.10599089
- Bryman, A. (2008). Research design. I A. Bryman (Red.), *Social research methods* (s. 29 - 64). Oxford: Oxford University Press.
- Budde, H., Voelcker-Rehage, C., Pietraßyk-Kendziorra, S., Ribeiro, P. & Tidow, G. (2008). Acute coordinative exercise improves attentional performance in adolescents. *Neuroscience Letters*, 441(2), 219-223.
- Bugge, A., Tarp, J., Østergaard, L., Domazet, S. L., Andersen, L. B. & Froberg, K. (2014). LCoMotion - Learning, Cognition and Motion; a multicomponent cluster randomized school-based intervention aimed at increasing learning and cognition - rationale, design and methods. *BMC Public Health*, 14, 1-8.
- Campbell, M. K, Elbourne, D. R. & Altman, D. G. (2004). CONSORT statement: Extension to cluster randomised trials. *British Medical Journal*, 328(7441), 702-708.
- Castelli, D. M., Hillman, C. H., Hirsch, J., Hirsch, A. & Drollette, E. (2011). FIT Kids: Time in target heart zone and cognitive performance. *Preventive Medicine*, 52, 55-59. doi:10.1016/j.ypmed.2011.01.019
- Chaddock, L., Erickson, K. I., Prakash, R. S., VanPatter, M., Voss, M. W., Pontifex, M. B., Raine, L. B., Hillman, C. H. & Kramer, A. F. (2010). Basal ganglia volume is associated with aerobic fitness in preadolescent children. *Developmental Neuroscience*, 32(3), 249-256. doi: 10.1159/000316648
- Chaddock, L., Pontifex, M. B., Hillman, C. H. & Kramer, A. F. (2011). A Review of the Relation of Aerobic Fitness and Physical Activity to Brain Structure and Function in Children. *Journal of the International Neuropsychological Society*, 17(6), 975-985. doi:10.1017/S1355617711000567

- Chang, Y-K., Tsai, Y-J., Chen, T-T. & Hung, T-M. (2013). The impacts of coordinative exercise on executive function in kindergarten children: An ERP study. *Experimental Brain Research*, 225(2), 187-196. doi: 10.1007/s00221-012-3360-9
- Chen, A.-G., Yan, J., Yin, H-C., Pan, C-Y. & Chang, Y-K. (2014). Effects of acute aerobic exercise on multiple aspects of executive function in preadolescent children. *Psychology of Sport & Exercise*, 15(6), 627-636.
- Clark, J. & Paivio, M. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149-210.
- Cliff, D. P., Hesketh, K. D., Vella, S. A., Hinkley, T., Tsiros, M. D., Ridgers, N. D., Carver, A., Veitch, J., Parrish, A-M., Hardy, L. L., Plotnikoff, R. C., Okely, A. D., Salmon, J. & Lubans, D. R. (2016). Objectively measured sedentary behaviour and health and development in children and adolescents: Systematic review and meta-analysis. *Obesity Reviews*, 17(4), 330-344.
- Coon Thompson, J., Boddy, K., Stein, K., Whear, R., Barton, J. & Depledge, M. (2011). Does participating in physical activity in outdoor natural environments have a greater effect on physical and mental wellbeing than physical activity indoors? A systematic review. *Journal Of Epidemiology And Community Health*, 65, 38. doi: 10.1136/jech.2011.143586.85
- Costigan, S. A., Eather, N., Plotnikoff, R. C., Taaffe, D. R., Pollock, E., Kennedy, S. G. & Lubans, D. R. (2015). Preliminary efficacy and feasibility of embedding high intensity interval training into the school day: A pilot randomized controlled trial. *Preventive Medicine Reports*, 2, 973-979.
- Costigan, S. A., Eather, N., Plotnikoff, R. C., Hillman, C. H. & Lubans, D. R. (2016). High-Intensity Interval Training for Cognitive and Mental Health in Adolescents. *Medicine & Science in Sports & Exercise*, 48(10), 1985-1993. doi: 10.1249/MSS.0000000000000993
- Crova, C., Struzzolino, I., Marchetti, R., Masci, I., Vannozzi, G., Forte, R. & Pesce, C. (2013). Cognitively challenging physical activity benefits executive function in overweight children. *Journal of Sports Sciences*, 32(3), 201-211. doi: 10.1080/02640414.2013.828849
- Davis, C., Tomporowski, P. D., McDowell, J. E., Austin, B. P., Miller, P. H., Yanasak, N. E., Allison, J. D. & Naglieri, J. A. (2011). Exercise Improves Executive Function and Achievement and Alters Brain Activation in Overweight Children: A Randomized, Controlled Trial. *Health Psychology*, 30(1), 91-98. doi: 10.1037/a0021766
- De Greeff, J. W., Hartman, E., Mullender-Wijnsma, M. J., Bosker, R. J., Doolaard, S. & Visscher, C. (2016). Long-Term Effects of Physically Active Academic Lessons on Physical Fitness and Executive Functions in Primary School Children. *Health Education Research*, 31(2), 185-194. doi:10.1093/her/cyv102

- Dekkers, O. M., Von Elm, E., Algra, A., Romijn, J. A. & Vandenbroucke, J. P. (2010). How to assess the external validity of therapeutic trials: A conceptual approach. *International Journal of Epidemiology*, 39(1), 89-94. doi:10.1093/ije/dyp174
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64(1), 135-168. 10.1146/annurev-psych-113011-143750
- Diamond, A. & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified ad those that, despite much hype, do not. *Development Cognitive Neuroscience*, 18, 34-48.
- Dishman, R. K., Berthoud, H. R., Booth, F. W., Cotman, C. W., Edgerton, V. R., Fleshner, M. R., Gandevia, S. C., Gomez-Pinilla, F., Greenwood, B. N., Hillman, C. H., Kramer, A. F., Levin, B. E., Moran, T. H., Russo-Neustadt, A. A., Salamone, J. D., Van Hoomissen, J. D., Wade, C. E., York, D. A. & Zigmond, M. J. (2006). Neurobiology of exercise. *Obesity*, 14(3), 345-356.
- Donnelly, J. E., Hillman, C. H., Castelli, D., Etnier, J. N., Lee, S., Tomporowski, P., Lambourne, K. & Szabo-Reed, A. N. (2016). Physical Activity, Fitness, Cognitive Function, and Academic Achievement in Children: A Systematic Review. *Medecine & Science in Sports & Exercise*, 48(6), 1197-1222. doi:10.1249/MSS.0000000000000901
- Downs, S. & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health*, 52(6), 377-384.
- Elleberg, D. & St-Louis-Deschênes, M. (2010). The effect of acute physical exercise on cognitive function during development. *Psychology of Sport & Exercise*, 11(2), 122-126.
- Ericsson, I. & Karlsson, M. K. (2014). Motor skills and school performance in children with daily physical education in school – a 9-year intervention study. *Scandinavian Journal of Medicine & Science in Sports*, 24(2), 273-278.
- Erwin, H., Fedewa, A. & Ahn, S. (2012). Student academic performance outcomes of a classroom physical activity intervention: a pilot study. *International Electronic Journal of Elementary Education*, 4(3), 473-487.
- Etnier, J., Labban, J. D., Piepmeier, A., Davis, M. & Henning, D. A. (2014). Effects of an acute bout of exercise on memory in 6th grade children. *Pediatric Exercise Science*, 26, 250-258. doi: 10.1123/pes.2013-0141
- Fedewa, A. L. & Ahn, S. (2011). The Effects of Physical Activity and Physical Fitness on Children's Achievement and Cognitive Outcomes: A Meta-Analysis. *Pediatric Exercise Science*, 26, 250 - 258. doi: 10.1123/pes.2013-0141
- Fedewa, A. L., Ahn, S., Erwin, H. & Davis, M. C. (2015). A Randomized Controlled Design Investigating the Effects of Classroom-Based Physical Activity on Children's Fluid Intelligence and Achievement. *School Psychology International*, 36(2), 135-153. doi: 10.1177/0143034314565424

- Ferris, L. T., Williams, J. S. & Shen, C-L. (2007). The Effect of Acute Exercise on Serum Brain-Derived Neurotrophic Factor Levels and Cognitive Function. *Medicine & Science in Sports & Exercise*, 39(4), 728-734. doi: 10.1249/mss.0b013e31802f04c7
- Fleischer, A. V. & From, K. (2017). *Eksekutive funksjoner hos barn og unge*. Bryne: Info Vest forlag.
- Fox, K. R. & Wilson, P. M. (2008). Self-perceptual systems and physical activity. I T. S. Horn (Red.), *Advances in Sport Psychology* (s. 49 – 64). Oxford, Ohio: Human Kinetics.
- Gallotta, M. C., Emerenziani, G. P., Franciosi, E., Meucci, M., Guidetti, L. & Baldari, C. (2015). Acute physical activity and delayed attention in primary school students. *Scandinavian Journal of Medicine & Science in Sports*, 25(3), 331-338. doi: 10.1111/sms.12310
- Gligoroska, J. P. & Manchevska, S. (2012). The Effect of Physical Activity on Cognition – Physiological Mechanisms. *Materia Socio-Medica*, 24(3), 198-202. doi: 10.5455/msm.2012.24.198-202
- Grieco, L. A., Jowers, E. M. & Bartholomew, J. B. (2009). Physically Active Academic Lessons and Time on Task: The Moderating Effect of Body Mass Index. *Medicine & Science in Sports & Exercise*, 41(10), 1921-1926. doi: 10.1249/MSS.0b013e3181a61495
- Have, M., Nielsen, J. H., Gejl, A. K., Ernst, M. T., Fredens, K., Støckel, J. T., Wedderkopp, N., Domazet, S. L., Gudex, C., Grøntved, A. & Kristensen, P. L. (2016a). Rationale and design of a randomized controlled trial examining the effect of classroom-based physical activity on math achievement. *BMC Public Health*, 16(1), 1-11. doi: 10.1186/s12889-016-2971-7
- Have, M. & Kristensen, P. L. (2016b). *Count your body: the effect of classroom-based physical activity and academic achievement in math*. (Doktorgradsavhandling, upublisert), University of Southern Denmark, Odense. Hentet fra forfatteren.
- Have, M., Nielsen, J. H., Ernst, M. T., Gejl, A. K., Fredens, K., Grøntved, A. & Kristensen, P. L. (2016c). A classroom-based physical activity intervention improves children`s math achievement – A randomized Controlled Trial. I Have, M. (Red.), *Count your body: the effect of classroom-based physical activity and academic achievement in math* (88 – 107). Upublisert.
- Herbert, R. D. & Gabriel, M. (2002). Effects of stretching before and after exercising on muscle soreness and risk of injury: Systematic review. *British Medical Journal*, 325, 1-5.
- Higgins, J., Green, S. & Cochrane Collaboration. (2009). *Cochrane handbook for systematic reviews of interventions*. Hentet fra: <http://training.cochrane.org/handbook>
- Hill, L., Williams, J. H. G., Aucott, L., Milne, J., Thomson, J., Greig, J., Munro, V. & Mon-Williams, M. (2010). Exercising attention within the classroom. *Developmental Medicine & Child Neurology*, 52(10), 929-934. doi: 10.1111/j.1469-8749.2010.03661.x

- Hillman, C. H., Erickson, K. I. & Kramer, A. F. (2008). Be smart, exercise your heart: Exercise effects on brain and cognition. *Nature Reviews Neuroscience*, 9(1), 58-65.
- Hillman, C. H., Pontifex, M. B., Raine, L. B., Castelli, D. M., Hall, E. E. & Kramer, A. F. (2009). The effect of acute treadmill walking on cognitive control and academic achievement in preadolescent children. *Neuroscience*, 159(3), 1044-1054.
- Hillman, C. H., Kamijo, K. & Scudder, M. (2011). A review of chronic and acute physical activity participation on neuroelectric measures of brain health and cognition during childhood. *Preventive Medicine*, 52, 21- 28. doi:10.1016/j.ypmed.2011.01.024
- Hillman, C. H., Pontifex, M. B., Castelli, D. M., Khan, N. A., Raine, L. B., Schudder, M. R., Drolette, E. S., Moore, R. D., Wu, C-T. & Kamijo, K. (2014). Effects of the FITkids randomized controlled trial on executive control and brain function. *Pediatrics*, 134(4), 1063-1071.
- Hostetter, A. B. & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, 15(3), 495-514. doi: 10.3758/PBR.15.3.495
- Howie, E. K. & Pate, R. R. (2012). Physical activity and academic achievement in children: A historical perspective. *Journal of Sport and Health Science*, 1(3), 160-169.
- Imsen, G. (2005). Veien til elevens verden. I G. Imsen (Red.), *Elevens verden: Innføring i pedagogisk psykologi* (4. utg.) (s. 29 – 50). Oslo: Universitetsforlaget.
- Keeley, T. & Fox, K. (2009). The impact of physical activity and fitness on academic achievement and cognitive performance in children. *International Review of Sport and Exercise Psychology*, 2(2), 198-214. doi: 10.1080/17509840903233822
- Khalsa, S. B., Hickey-Schultz, L., Cohen, D., Steiner, N. & Cope, S. (2012). Evaluation of the mental health benefits of yoga in a secondary school: A preliminary randomized controlled trial. *Journal of Behavioral Health Services & Research*, 39(1), 80-90.
- Kibbe, D. L., Hackett, J., Hurley, M., McFarland, A., Schubert, K. G., Schultz, A. & Harris, S. (2011). Ten Years of TAKE 10!: Integrating physical activity with academic concepts in elementary school classrooms. *Preventive Medicine*, 52, 43- 50.
- Kipp, L. E. & Weiss, M. R. (2013). Physical activity and self-perceptions among children and adolescents. In P. Ekkekakis (Red.), *Routledge handbook of physical activity and mental health* (pp. 187-199). New York: Routledge.
- Konstantina, T., Loyens, S., Mavilidi, M. F. & Paas, F. (2015). Preschool Children's Foreign Language Vocabulary Learning by Embodying Words through Physical Activity and Gesturing. *Educational Psychology Review*, 27(3), 445-456. doi: 10.1007/s10648-015-9316-4
- Kunnskapsdepartementet. (2013). *Lange linjer - kunnskap gir muligheter* (Meld. St. 18. 2012-2013). Hentet fra [https://www.regjeringen.no/no/dokumenter/meld-st-18-20122013/id716040/?q=Lange linjer kunnskap gir muligheter](https://www.regjeringen.no/no/dokumenter/meld-st-18-20122013/id716040/?q=Lange+linjer+kunnskap+gir+muligheter)

- Lakes, K. D. & Hoyt, W. T. (2004). Promoting Self-Regulation through School-Based Martial Arts Training. *Applied Developmental Psychology*, 25(3), 283-302.
- Link, T., Moeller, K., Huber, S., Fischer, U. & Nuerk, H. C. (2013). Walk the number line - An embodied training of numerical concepts. *Trends in Neuroscience and Education*, 2(2), 74-84.
- Lubans, D. R., Morgan, P.J., Cliff, D. P., Barnett, L. M. & Okely, A. D. (2010). Fundamental Movement Skills in Children and Adolescents. *Sports Medicine*, 40(12), 1019-1035.
- Lubans, D. L., Richards, J., Hillman, C., Faulkner, G., Beauchamp, M., Nilsson, M., Kelly, P., Smith, J., Raine, L. & Biddle, S. (2016). Physical activity for cognitive and mental health in youth: A systematic review of mechanisms. *Pediatrics*, 138(3), 1 - 13.
- Mahar, M. T., Murphy, S. K., Rowe, D. A., Golden, J., Shields, A. T. & Raedeke, T. D. (2006). Effects of a Classroom-Based Program on Physical Activity and On-Task Behavior. *Medicine & Science in Sports & Exercise*, 38(12), 2086-2094. doi: 10.1249/01.mss.0000235359.16685.a3
- Maher, C. G., Sheerington, C., Herbert, R. D., Moseley, A. M. & Elkins, M. (2003). Reliability of the PEDro scale for rating quality of randomized controlled trials. *Physical Therapy*, 83(8), 713-21.
- Macintosh, B. J., Crane, D. E., Sage, M. D., Rajab, A. S, Donahue, M. J., McIlroy, W. E. & Middleton, L. E. (2014). Impact of a Single Bout of Aerobic Exercise on Regional Brain Perfusion and Activation Responses in Healthy Young Adults. *PLOS ONE*, 9(1), 1-7. doi:10.1371/journal.pone.0085163.t002
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & the PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of internal medicine*, 151(4), 264-270.
- Mullender-Wijnsma, M. J., Hartman, E., de Greeff, J. W., Doolaard, S., Bosker, R. J. & Visscher, C. (2016). Physically Active Math and Language Lessons Improve Academic Achievement: A Cluster Randomized Controlled Trial. *Pediatrics*, 137(3), 1-9. doi: 10.1542/peds.2015-2743
- Niedenthal, P. (2007). Embodying emotion. *Science*, 316, 1002-1005. doi: 10.1126/science.1137651
- Ogilvie, D., Fayter, D., Petticrew, M., Sowden, A., Thomas, S., Whitehead, M. & Worthy, G. (2008). The harvest plot: A method for synthesising evidence about the differential effects of interventions. *BMC Research Methodology*, 8(8), 1-7. doi:10.1186/1471-2288-8-8
- Ommundsen, Y. (2013). Fysisk-motorisk ferdighet gjennom kroppsøving – et viktig bidrag til elevenes allmenndanning og læring i skolen. *Norsk pedagogisk tidsskrift*, 97(2), 155-166.
- Pesce, C., Crova, C., Cereatti, L., Casella, R. & Bellucci, M. (2009). Physical activity and Mental performance in preadolescents: Effects of acute exercise on free-recall memory. *Mental Health and Physical Activity*, 2(1), 16-22. doi:10.1016/j.mhpa.2009.02.001

- Pesce, C., Crova, C., Marchetti, M., Struzzolino, I., Masci, I. & Vannozzi, G. (2013). Searching for cognitively optimal challenge point in physical activity for children with typical and atypical motor development. *Mental Health and Physical Activity*, 6, 172 – 180.
- Petersen, S. E. & Posner, M. I. (2012). The Attention System of the Human Brain: 20 Years After. *Annual Review of Neuroscience*, 35, 73-89. doi:10.1146/annurev-neuro-062111-150525
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128-2148. doi:10.1016/j.clinph.2007.04.019
- Polit, D. & Beck, C. T. (2014). *Essentials of nursing research: Appraising evidence for nursing practice*. Philadelphia: Wolters Kluwer Health / Lippincott Williams & Wilkins.
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K. & Duffy, S. (2006). Guidance on narrative synthesis – an overview. I J. Popay, H. Roberts, A. Sowden, M. Petticrew, L. Arai, M. Rodgers, N. Britten, K. Roen & S. Duffy (Red.), *Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC Methods Programme (s. 11-23)*. UK: Lancaster University. doi:10.13140/2.1.1018.4643
- Quinto Romani, A. & Klausen, T. B. (2017). Physical Activity and School Performance: Evidence from a Danish Randomised School-Intervention Study. *Scandinavian Journal of Educational Research*, 61(4), 479-502. doi: 10.1080/00313831.2016.1172498
- Querido, J. & Sheel, S. (2007). Regulation of Cerebral Blood Flow During Exercise. *Sports Medicine*, 37(9), 765-782.
- Randolph, J. J. (2009). A Guide to Writing the Dissertation Literature Review. *Practical Assessment, Research & Evaluation*, 14(13), 1-13.
- Resaland, G. K., Moe, V. F., Aadland, E., Steene-Johannessen, J., Glosvik, Ø., Andersen, J. R., Kvalheim, O. M., McKay, H. A. & Anderssen, S. A. (2015). Active Smarter Kids (ASK): Rationale and design of a cluster-randomized controlled trial investigating the effects of daily physical activity on children's academic performance and risk factors for non-communicable diseases. *BMC Public Health*, 15(1), 709-709. doi: 10.1186/s12889-015-2049-y
- Resaland, G. K., Aadland, E., Moe, V. F., Aadland, K. N., Skrede, T., Stavnsbo, M., Suominen, L., Steene-Johannessen, J., Glosvik, Ø., Andersen, J. R., Kvalheim, O. M., Engelsrud, G., Andersen, L. B., Holme, I. M., Ommundsen, Y., Kriemler, S., van Mechelen, W., McKay, H. A., Ekelund, U. & Andersen, S. A. (2016). Effects of physical activity on schoolchildren's academic performance: The Active Smarter Kids (ASK) cluster-randomized controlled trial. *Preventive Medicine*, 91, 322-328.
- Ryff, C. D. & Keyes, C. L. (1995). The Structure of Psychological Well-Being Revisited. *Journal of Personality and Social Psychology*, 69(4), 719-727.

- Schmidt, M., Jager, K., Egger, F., Roebbers, C. & Conzelmann, A. (2015). Cognitively engaging chronic physical activity, but not aerobic exercise, affects executive functions in primary school children: A group-randomized controlled trial. *Journal of Sport & Exercise Psychology*, 37(6), 575-591.
- Schneider, S., Vogt, T., Frysich, J., Guardiera, P. & Strüder, H. K. (2009). School sport—A neurophysiological approach. *Neuroscience Letters*, 467(2), 131-134. doi:10.1016/j.neulet.2009.10.022
- Singh, A., Uijtdewilligen, L., Twisk, J., Van Mechelen, W. & Chinapaw, M. (2012). Physical Activity and Performance at School: A Systematic Review of the Literature Including a Methodological Quality Assessment. *Archives of Pediatrics & Adolescent Medicine*, 166(1), 49-55.
- Slavin, R. E. (1995). Best evidence synthesis: An intelligent alternative to meta-analysis. *Journal of Clinical Epidemiology*, 48(1), 9-18.
- Sonstroem R. J., Harlow, L. L. & Josephs, L. (1994). Exercise and self-esteem: Validity of Model expansion and Exercise Associations. *Journal of Sports & Exercise Psychology*, 16, 29-42.
- Subramanian, S. K., Sharma, V. K., Arunachalam, V. Radhakrishnan, K. & Ramamurthy, S. (2015). Effect of Structured and Unstructured Physical Activity Training on Cognitive Functions in Adolescents - A Randomized Control Trial. *Journal of Clinical & Diagnostic Research*, 9(11), 4-9. doi: 10.7860/JCDR/2015/14881.6818
- Tarp, J., Domazet, S. L., Froberg, K., Hillman, C. H., Andersen, L. B. & Bugge, A. (2016). Effectiveness of a School-Based Physical Activity Intervention on Cognitive Performance in Danish Adolescents: LCoMotion—Learning, Cognition and Motion – A Cluster Randomized Controlled Trial. *PLoS ONE*, 11(6), 1-19. doi:10.1371/journal.pone.0158087
- Telles, S., Singh, N., Bhardwaj, A. K., Kumar, A. & Balkrishna, A. (2013). Effect of yoga or physical exercise on physical, cognitive and emotional measures in children: a randomized controlled trial. *Child & Adolescent Psychiatry & Mental Health*, 7(37), 1-16.
- Thomas, A. G., Dennis, A., Bandettini, P. A. & Johansen-Berg, H. (2012). The effects of physical activity on brain structure. *Frontiers in Psychology*, 3, 1-9. doi: 10.3389/fpsyg.2012.00086
- Tomprowski, P. D., Davis, C. L., Miller, P. H. & Naglieri, J. A. (2008). Exercise and Children's Intelligence, Cognition, and Academic Achievement. *Educational Psychology Review*, 20(2), 111-131. doi: 10.1007/s10648-007-9057-0
- Tomprowski, P. D., McCullick, B. A. & Horvat, M. (2010). The role of contextual interference and mental engagement on learning. I Edvardsen F., & Kulle, H. (Red.), *Educational Games: Design, Learning and Applications*, 127-155. New York: Nova Science Publishers.

- Tomporowski, P. D., McCullick, B., Pendleton, D. M. & Pesce, C. (2015). Exercise and children's cognition: The role of exercise characteristics and a place for metacognition. *Journal of Sport and Health Science*, 4, 47-55.
- Tooth, L., Ware, R., Bain, C., Purdie, D. M. & Dobson, A. (2005). Quality of reporting of observational longitudinal research. *American Journal of Epidemiology*, 161(3), 280-288. doi: 10.1093/aje/kwi042
- Utdanningsdirektoratet. (2015). *Læreplan i kroppsøving (KRO1-04. 2015)*. Hentet fra <https://www.udir.no/kl06/KRO1-04>
- Van Praag, H. (2008). Neurogenesis and exercise: Past and future directions. *Neuromolecular Medicine*, 10, 128-140. doi: 10.1007/s12017-008-8028-z
- Vazou, S., Pesce, C., Lakes, K. & Smiley-Oyen, A. (2016). More than one road leads to Rome: A narrative review and meta-analysis of physical activity intervention effects on cognition in youth. *International Journal of Sport and Exercise Psychology*, 1-26. doi: 10.1080/1612197X.2016.1223423
- Voss, M. W., Vivar, C., Kramer, A. F. & Van Praag, H. (2013). Bridging animal and human models of exercise-induced brain plasticity. *Trends in Cognitive Sciences*, 17(10), 525-544.
- Wasserman, T. & Wasserman, L. D. (2013). Toward an Integrated Model of Executive Functioning in Children. *Applied Neuropsychology, Child*, 2, 88 – 96. doi: 10.1080/21622965.2013.748394
- Williamson, P. & Gamble, C. (2005). Identification and impact of outcome selection bias in meta-analysis. *Statistics in Medicine*, 24, 1547-1561. doi: 10.1002/sim.2025
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625-636.
- Ylvisaker, M. & Feeney, T. (2008). Helping children without making them helpless: Facilitating development of executive self-regulation in children and adolescents. I V. Anderson, R. Jacobs & P. J. Anderson (Red.), *Executive Functions and the Frontal Lobes: A Lifespan Perspective* (409 - 438). Hove, East Sussex: Taylor and Francis.

8.0 Vedlegg

8.1 Vedlegg 1: Glossarium – begreper, forkortelser og administrering av testinstrumenter

8.1.1 Begreper

Under presenteres sentrale begreper som brukes i rapporten.

Aerob fitness

Aerob fitness er det kardiorespiratoriske systems maksimale kapasitet til opptak og bruk av oksygen (Hillman et al., 2008).

Akademisk prestasjon

Akademisk prestasjon er et mål på i hvilken grad en elev eller en institusjon har oppnådd deres akademiske målsetning. Akademisk prestasjon blir vanligvis målt ved eksaminering eller kontinuerlig vurdering (karakterer, standardisert testscore eller annen formell vurdering i akademiske fag som lesing, matematikk og språk) (Donnelly et al., 2016).

Akutt fysisk aktivitet

Akutt fysisk aktivitet er enkeltøkter fysisk aktivitet med kortere varighet (Best, 2010).

BDNF «brain derived neurotrophin factor»

BDNF er i en familie av nevrotropiske faktorer som inngår i nevrologisk signalgang, modulering og plastisitet (tilpasning til atferd, miljø og opplevelse). Studier har vist at akutt og kronisk fysisk aktivitet fører til økt nivå av BDNF i ulike områder av hjernen. Økningen i BDNF er intensitetsavhengig (Ferris, Williams & Shen, 2007).

Dipolmodellering

Dipolmodellering er en metode for å bestemme lokalisering av kilden som gir en målt respons i et elektroencefalografisk eksperiment. Modelleringen gir et estimat på lokalisering, retning og styrke på kilden som funksjon av tid etter at stimuli ble presentert (Hillman et al., 2008).

Effektstørrelse (effect size - ES)

Effektstørrelse er et kvantitativt mål på styrken til et fenomen. Eksempler på effektstørrelse: korrelasjon mellom to variabler, regresjonskoeffisienten i en regresjonslinje, en gjennomsnittlig differanse. En større absoluttverdi indikerer sterkere effekt. Effektstørrelse kompletterer statistisk hypotesetesting og har en viktig rolle i power analyser, beregning av utvalgsstørrelse og i meta-analyser (Polit & Beck, 2014).

Eksekutiv kontroll

Eksekutiv kontroll er en beregningsprosess som involverer seleksjon, planlegging og koordinering av komplekse kognitive funksjoner (Hillman et al., 2008).

Fitness

Fitness er en fysiologisk tilstand av velvære som er grunnlag for god helse og som muliggjør dagliglivets funksjonelle oppgaver. Fitness inneholder komponenter som kardiorespiratorisk kapasitet, muskelstyrke, utholdenhet, fleksibilitet og kroppssammensetning (Donnelly et al., 2016).

Fysisk aktivitet

Fysisk aktivitet er kroppslig bevegelse produsert av skjelettmuskulatur ved forbruk av energi (Hillman et al., 2008)

Hendelses-relatert nevrologisk potensial, "Event-related brain potential (eRP)

Hendelses-relatert nevrologisk potensial (eRP) er en tidsbestemt indeks for nevrologisk aktivering som assosieres med spesifikke kognitive prosesser (Hillman et al., 2008).

Heterogenitet (statistical heterogeneity)

Heterogenitet er et mål på variasjonen mellom studier for en rekke utvalgte variabler (Booth, Sutton & Papaioannou, 2016). Eksempel: måleresultater er ikke identiske for alle deltagere i en studie, derfor oppstår en uunngåelig variasjon i effekt over flere studier. Dersom måleresultater befinner seg innenfor en rimelig variasjon da kan et viktig grep i forskningsdesignet være å undersøke kilden til variasjon. Eksempel: effekten av en intervensjon kan være systematisk forskjellig for kvinner og menn. Delanalyser planlegges vanligvis for å undersøke slike systematiske forskjeller nærmere og for å redusere heterogenitet (Polit & Beck, 2014).

IGF-1 «insulin-like growth factor»

IGF-1 er en vekstfaktor, peptider som ligner insulin, og som stimulerer motoriske enheter og andre vev perifert. Fysisk aktivitet stimulerer vekstfaktorer (som IGF-1) og forsterker proteinsyntese, skjellettvekst og celleproliferasjon (Van Praag, 2008).

Katekolaminer

Katekolaminer er hormoner med viktigste oppgave å forberede kroppen på belastende og kritiske situasjoner. Katekolaminer øker nedbrytning av glykogen i lever og muskel, øker nedbrytning av fett lagret i fettvev, øker hjertets kontraksjonsstyrke og frekvens, endrer regional blodgjennomstrømning slik at mer blod pumpes gjennom arbeidende skjelett-muskulatur, øker blodtrykk og ikke minst stimulerer retikulærsubstansen i hjernestammen slik at individets aktpågivenhet og beredskap øker (Gligoroska et al., 2012).

Kognisjon

Kognisjon er et sett mentale prosesser som bidrar til persepsjon, minne, intellekt og handling (Donnely et al., 2016).

Kognitivt beriket

Kognitivt beriket er en kodeenhet som essensensielt handler om fysisk aktivitet som er tilført et kognitivt stimulerende innhold. Intervensjonene som beskrives som kognitivt beriket favner et bredt spekter av aktiviteter. I enkelttilfeller innrettes aktiviteten slik at man stimulerer spesifikke kognitive komponenter som f.eks. i bevegelsesbasert problemløsning, beslutningsoppgaver og reaksjonstid (Galotta et al., 2015). Andre eksempler er hopping til musikk (Hill et al., 2010), kognitivt krevende lagspill (Schmidt et al., 2015) og Yoga (Telles et al., 2013). Akademisk prestasjon blir stimulert ved klasseromsintegrert aktiv matematikk (Mullender-Winsma et al., 2016; Have et al., 2016c).

Konfundere (seleksjonsavvik)

Konfundere er variabler som innvirker på måleresultater, som et tillegg til fysisk aktivitetsintervensjon, og dermed reduserer internvaliditeten. Årsakssammenhengen mellom uavhengig og avhengig variabel svekkes. Eksempel: manglende randomisering ved allokering av deltagere, intervensjon og kontrollgruppe med ulikheter i baselineverdier slik at differansen i resultater endres, intervensjon og kontrollgruppe med treningsprogram som

avviker fra det planlagte programmet (fysisk aktivitets- type og dosering, kognitivt eller akademisk stimuli) (Downs & Black, 1998).

Kronisk fysisk aktivitet

Kronisk fysisk aktivitet er repeterte økter aktivitet over lengre tid (Best, 2010).

Mediator

Mediator er en medierende variabel som fungerer som et mellomledd i kausale sammenhenger, og som kobler 2 variabler sammen (Polit & Beck, 2014).

Moderator

Moderator er en variabel som påvirker (modererer) sammenhengen mellom uavhengig og avhengig variabel (Polit & Beck, 2014).

P3

P3 er et positivt utslag som følge av stimuli-bestemt eRP som avspeiler en nevrologisk representasjon av stimuli fra omgivelsen og er proporsjonal med oppmerksomheten som kreves for å tolke et gitt stimuli (amplituden) og gir i tillegg hurtigheten i evaluering av stimuli (latenstid) (Hillman et al., 2008).

Trening

Trening er en type fysisk aktivitet som er planlagt, strukturert og repetert og som har som målsetning å bedre eller vedlikeholde fysisk form (Donnelly et al., 2016).

VEGF “vascular endothelial growth factor” (vascular permeability factor – VGF)

VEGF er en vekstfaktor som stimulerer dannelsen av blodkar, og er et viktig signalprotein for angiogenese (Van Praag, 2008).

8.1.2 Forkortelser

Under presenteres sentrale forkortelser som brukes i rapporten (= leses, «betyr» eller «står for»).

2-back = modifisert visual «N-back» (2 - back) task.

AP = akademisk prestasjon.

BPM = beats per min.te,

BRIEF = Behavior rating inventory of executive function.

BSEI = Battles self-esteem inventory.

C-a FT = Child-adapted Flanker task.

CAT-3 = Canadian Achievement test.

D2- a = D2-test of attention.

DSB= Digit-span backwards.

DSE = Digit-symbol encoding.

EF = Eksekutiv funksjon.

EFT = Erickson Flanker task.

EFT = Modifisert Eriksen flanker test.

EPB = Eurofit physical battery.

F & V lessons = Physically active mathematics and language lessons.

FA = Fysisk aktivitet.

fMRI = Funksjonell magnetic resonance imaging.

FS = Flourishing Scale.

GPA= Grade point average.

K10 = The Kessler Psychological Distress Scale.

LA = Lærervurdert atferd - lydighet, akademisk prestasjon, oppmerksomhet, punktlighet, atferd med venner og lærere.

LCT = Two target letter cancellation test.

LS = Listening span.

MAP = Measures of Academic Progress.

MB FT = Mixed block Flanker task.

MG = 45-minute standardized math test.

M-OT = More-odd task (for skifte av setting).

MVPA = Moderate-to-vigorous physical activity.

NSn-b = Nonspatial n-back task.

PE = Physical education, kroppsøving.

PSA = Paced serial addition.

PS-DQ = Physical Self-Description Questionnaire.

RAVLT = Rey Auditory Verbal Learning Test.

RFFT = Ruff figural Fluency test.

RT = Reaksjonstid.

SCWT = Stroop color-word task for children.

SD = Standard avvik.

SOFI = Spanish Overall and Fractional Intelligence Test.

SOT = Size ordering test.

SPM = Raven`s Standard Progressive Matrics.

SS = Sky Search.

TEA-Ch = Test of selective attention in children.

TMT = The Trail making test.

TTA og TTB = Trail Making test A and B.

TTCT = Torrance tests of creative Thinking.

VF = Verbal fluency.

WISC-IV = Wechsler Intelligence Scale for Children, 4th ed.

8.1.3 Administrering av testinstrumenter

Under beskrives administrering av tester for inkluderte studier. Testene måler kognitiv funksjon og akademisk måloppnåelse (St1 – St17).

St1. (Etnier et al., 2014) bruker AB versjonen av *Rey Auditory Verbal Learning test (RAVLT)* som måler en rekke funksjoner. Blant disse funksjonene er verbalt (muntlig) minne, læringshurtighet og læringsstrategier. En forhåndsbestemt liste med ord ble spilt av likt for alle forsøkspersoner. Listen besto av 15 ord. Deltagerne fikk høre at dette var liste A. Et blankt papir med identitetsnummer ble utdelt etter avspilling og deltagerne fikk i oppgave å notere ned alle ord de kunne huske. Det ble gjort klart at hverken stavemåte eller rekkefølge av ordene spilte noen rolle. Etter at deltagerne signaliserte at de ikke husket flere ord, ble listen samlet inn (varighet maksimalt 2 min.). Denne umiddelbare rekapituleringen ble gjentatt 5 ganger (test 1 – 5). Testen ble betraktet som umiddelbar rekapitulering (STM). Dersom rekapituleringen bedret seg over de neste 4 testene, ble dette tolket som en indikasjon på verbal læring. Etter test 5 ble deltagerne fortalt at de skulle høre en ny liste med ord – Liste B. Etter at liste B var avspilt, ble deltagerne bedt om å rekapitulere så mange ord fra liste B som mulig (Test 6). Test 6 ble gjennomført på samme måte som de foregående 5 testene. Test 6 ble brukt som en slutningstest, hvor deltagerne ble forhindrede fra å repetere (øve) STM som strategi i rekapituleringen. I test 7 ble deltagerne bedt om å rekapitulere liste A nok en gang, uten å høre listen på forhånd og skrive ned ordene (LTM). Testen ble også gjentatt etter 24 timer, der oppgaven gikk ut på å gjenkjenne ord fra liste A og B i en større liste med 50 ord (20 nye ord i tillegg til liste A og B). De 50 ordene ble gitt hver deltager skriftlig på et stykke papir. Instruksjonen var å markere ord fra liste A med «A» ved siden av ordet, og markere ord

fra liste B med «B» ved siden av ordet. Denne målingen ble tolket som langtids-gjenkjennings minne (Etnier et al., 2014).

St2. (Chen et al., 2014) bruker en *modifisert Eriksen flanker test (EFT)* for å måle inhibering. En pc- basert nevropsykologisk applikasjon måler respons inhibering (evne til å motstå respons), arbeidsminne og skifte av setting. Inhibering er påvirket av akutt fysisk aktivitet. Flanker-testen er todelt med kongruente (likeformede) og ikke-kongruente oppgaver. Den kongruente testen har horisontalt organisert 5 bokstaver (eks. LLLLL eller FFFFF), mens den ikke kongruente testen har horisontalt organisert 5 bokstaver der midterste bokstav er ulik (eks. LLFLL eller FFLFF). Deltagerne ble bedt om å trykke «F» eller «L» med venstre eller høyre pekefinger, avhengig av midterste bokstav i rekken. Et fokus kryss (+) ble først presentert i 500 ms i senter av skjermen for å tiltrekke deltagers oppmerksomhet. Mellomtid 2000 ms deretter ble et kongruent eller inkongruent sett med bokstaver presentert i 1000 ms. All stimulus ble presentert på en svart bakgrunn. Deltagerne ble instruert om å trykke riktig svar raskest mulig. Uriktig svar (ikke lest eller ikke reagert) ble registrert dersom deltageren ikke responderte mellom 150 ms – 1000 ms. Deltagerne gjennomførte 12 øvelsestester, deretter 2 enheter med 48 tester i hver. Deltagerne fikk 1 minutt pause mellom hver enhet. Kongruente og inkongruente tester ble presentert i randomisert rekkefølge med lik sannsynlighet i hver enhet. Total tid for testen var ca. 6 minutter. Responstiden for kongruente og inkongruente tester ble registrert og brukt for å lage en indeks for inhibering definert som differansen på responstid mellom kongruente og inkongruente tester. En mindre differanse i responstid ble tolket som bedre prestasjon. En *modifisert visual «N-back» (2 – back) task* ble brukt for å måle oppdatering av arbeidsminne og arbeidsminnets kapasitet. Oppgaven består i en serie bokstaver som raskt endres i senter av skjermen (eks. B, D, L, Y, O). Alle bokstaver ble presentert på en svart bakgrunn i 2000 ms, med mellomtid 3000 ms. I testen ble deltagerne instruert om å respondere «Ja» dersom bokstaven var identisk med bokstaven som ble vist 2 bokstaver tilbake i rekken (2 – back). For å indikere «Ja». Trykket deltagerne «F» ved samsvarende 2-back, med venstre pekefinger. For å respondere «Nei» trykket deltagerne «L» med høyre pekefinger. Testene inneholdt 50% «Ja» og 50% «Nei» oppgaver. Feil svar ble registrert dersom feil knapp ble trykket eller reaksjonstiden ikke lå mellom 300 ms – 1500 ms. Deltagerne utførte 12 øvelsestester, deretter 2 enheter med 25 tester i hver. Deltagerne fikk 1 minutt pause mellom hver enhet. Total tid for testen var ca. 7 minutter. Responstid for korrekte tester ble registrert og beregnet med gjennomsnitt og ga en atferdsindeks. Kortere responstid ga bedre prestasjon. En *more-odd task* ble brukt for å måle skifte av setting. En

serie siffer fra 1 til 4 eller 6 til 9 ble vist i senter av skjermen. Sifrene ble presentert i 2000 ms, med 3000 ms mellomtid. Testen hadde 3 typer oppgaver. En oppgave (A) inneholdt 16 homogene tester der tall ble presentert i svart. Deltagerne ble instruert om å trykke «F» eller «L» med venstre og høyre pekefinger for å indikere tallet henholdsvis var større eller mindre enn 5. Neste oppgave (B) hadde 16 homogene tester der tall ble presentert med fargen grønn. Deltagerne ble instruert om å trykke «F» og «L» med venstre og høyre pekefinger for å indikere om tallet henholdsvis, var oddetall eller partall. Tredje oppgavesett inneholdt 32 heterogene tester (C) som inkluderte både A og B oppgaver (16 av hver). Oppgavene vekslet i annenhver test. Deltagerne ble instruert om å trykke «F» eller «L» for å indikere om tallet i svart farge var under eller over 5, eller om tallet i grønn farge var odde- eller partall. Feil svar ble registrert dersom deltager trykket feil tast eller ikke hadde en responstid mellom 150 ms og 1000 ms for homogene tester og mellom 300 ms og 1500 ms for heterogene tester. Deltagerne utført 8 A og B øvelsesoppgaver (totalt 16 tester), deretter ble 6 oppgaver gjennomført i følgende rekkefølge: ABCCBA. Deltagerne fikk 1 minutt pause mellom hver oppgave. Totaltid for testene var 12 minutter. Indeksen for skifte av setting, ble beregnet som differanse i responstid mellom heterogene (gjennomsnitt for C oppgaver) og homogene enheter (gjennomsnitt for A og B oppgaver) (Chen et al., 2014).

St3. (Altenburg et al., 2016) bruker *Sky search (SS)* en deltest av *Test of selective Attention in Children (TEA-Ch)* som måler kontroll av fokus og seleksjon. TEA-Ch har 9 ulike deltester som måler ulike aspekter ved oppmerksomhet. Deltagerne ble instruert til å finne 20 par identiske romskip blant 108 figurer (distraksjoner) så raskt som mulig. For å kontrollere for forskjeller i motorisk hurtighet ble først testen gjennomført uten distraksjoner. Deltageren registrerte selv tidsbruken for hver test ved hjelp av stoppeklokke. Forskere dobbeltsjekket tidsbruk og at testen ble gjennomført korrekt. Antall korrekte par identiske romskip ble talt. 2 ulike versjoner av SS testen ble tilfeldig tildelt deltagerne ved ulike tidspunkt. Tidsbruk per par ble beregnet etter at total tid var registrert. Selektiv oppmerksomhet ble beregnet ved tidsdifferansen mellom SS versjonen med 108 figurer og testen uten distraksjon. Med dette ble motorisk treghet og klossethet eliminert fra testen. Kortere tidsbruk med lav score indikerer bedre prestasjon for oppmerksomhet. Test-retest (etter 20 - 110 min) ble gjennomført for å undersøke korttids reliabilitet (Altenburg et al., 2016).

St4. (Galotta et al., 2015) bruker en *d2-test of attention (2 letter cancellation test)* som måler oppmerksomhet og spesielt visuell oppmerksomhet i løpet av en gitt tid. Deltagerne instrueres

i bokstavkansellering med papir og blyant. Testen består av 14 ulike linjer, hver linje med 47 bokstaver (p, d) med tilfeldig rekkefølge, med 1 – 4 enkle og/eller doble anførselstegn, enten under eller over hver bokstav. Deltagerne ble instruert om å avmerke bokstaven «d» hvis den hadde doble anførselstegn over eller under tegnet. Total tid for testen var 4 min 40 sek. d2-testen gir poeng avhengig av totalt antall prosesserte enheter (TN), minus antallet feil avmerkinger (CP) og ved prosentvis feil (E%) av totalt antall besvarte enheter. TN ble tolket som et mål på prosesseringshurtighet, og mengden fullført arbeid. CP ble tolket som et mål på prestasjon i konsentrasjon. E% ble tolket som et mål på kvaliteten i prestasjon. Måleinstrumentet var på forhånd reliabilitetstestet med rekkevidde 0.95 - 0.98, og validitetstestet med koeffisient 0.47. d2-testen måler kapasiteten til å fokusere et stimuli/fakta, mens man undertrykker forstyrrelser. Prestasjonen tallfester visuell perseptuell hurtighet og konsentrasjonskapasitet (Galotta et al., 2015).

St5. (Hill et al., 2010) bruker *Paced serial addition (PSA)* som måler kapasitet for informasjonsprosessering. Deltagerne blir instruert til å legge sammen tallpar. Fra en verbal avspilt rekke en-sifrede tall (summeres tall 1 og tall 2, deretter tall 2 og tall 3, deretter tall 3 og tall 4, også videre). Hver korrekte sum gir 1 poeng. Det ble gitt 3 oppgavesett henholdsvis med 3 sekunders pause, 2 sekunders pause og 1 sekunds pause, mellom hvert tall i rekken. I *size-ordering testen (SOT)* som måler arbeidsminne, lyttet deltagerne til en sekvens objekter (eks. ball - størst og blyant – minst) med ulik angitt størrelse. Deltagerne ble instruert til å rekapitulere objektene ordnet etter størrelse (i virkeligheten). Det ble gitt 2 poeng for hvert objekt som ble avgitt i riktig rekkefølge. *Sentence-span testen (SST)* måler arbeidsminne. Deltagerne ble instruert om å vurdere påstand som sann eller falsk. Vurderingen av utsagn var en avledning. Deltagerne fikk avspilt 9 trivielle utsagn med omtrentlig lik lengde. Siste ord i hver setning måtte rekapituleres ved avslutning av testen. 2 poeng ble gitt for hvert riktige endeord. *Digit-span backward testen (DSB)* måler arbeidsminnets kapasitet. Deltagerne ble instruert om å rekapitulere en sekvens (1 - 9 tall) i reversert rekkefølge. 2 poeng ble gitt for hver riktige besvarelse. *Digit-symbol encoding testen (DSE)* måler korttidsminne. Deltagerne fikk en tabell med tall og tilhørende symboler. Deltageren fikk en ny tabell med tall i tilfeldig rekkefølge og ble instruert om å fylle inn symboler under tilhørende tall. Testen ble gjennomført på tid - maksimalt 2 minutter. 1 poeng ble gitt for hvert riktig plasserte symbol. To ulike oppgavesett α (uke 1) og β (uke 2) ble brukt for alle tester for å unngå øvelseselementet ved retesting (Hill et al., 2010).

St6. (Schmidt et al., 2015) bruker *nonspatial n-back task* som måler eksekutiv funksjon, oppdatering av arbeidsminnet og arbeidsminnets kapasitet. Flere bilder av frukt ble presentert på skjermen en etter en. Deltagerne ble instruert om å trykke høyre tast når frukten var identisk med frukten to bilder tilbake (2- back). Venstre tast for alle andre bilder. Total tid for testen var 10 minutter. Testen besto av 2 deler med 24 oppgaver. 1 poeng ble gitt for hvert korrekte svar. *Child- adapted Flanker task* ble brukt for å måle respons inhibering (beskrevet i St2. over). Skifte av setting ble målt med en *kombinasjonsoppgave i Flanker testen*. En tilleggsregel ble introdusert ved at ulike farger ble brukt i oppgavene. Testen krevde skifte mellom 2 regler ved fargeendring. Differansen mellom kombinasjonsoppgaver og enkle oppgaver ga et mål på inhibering (Schmidt et al., 2015).

St7. (Costigan et al., 2016) bruker *Trail making test (TMT)* som måler visuell oppmerksomhet og skifte av setting. Testen var todelt. Først ble deltagerne instruert om å tegne linjer (trail A) mellom tall i stigende rekkefølge (eks. 1 – 2 – 3 – 4) så hurtig og presist som mulig. Den andre delen besto i å tegne linjer (trail B) mellom tall og bokstaver i stigende rekkefølge (eks. 1 – a – 2 – b – 3 – c – 4 – d) så raskt som mulig. Tidtaking ga poengscore. Lav score indikerte bedre kognitiv prestasjon. Kognitiv fleksibilitet ble etter litteraturen, beregnet på 3 måter: tid på trail B minus tid på trail A, tid på trail B dividert med tid på trail A eller kun tid på trail B. *Flourishing Scale* måler (spørreskjema) deltagerens persepsjon av suksess i relasjoner, global selvpoppfatning, hensikt og optimisme. Testen hadde 8 enheter og deltagerne selvrapporterte på Likert skala (1 = svært uenig til 7 = svært enig) i hvilken grad de var enige i utsagn som handler om sosialt velvære (eks. Jeg har et hensiktsmessig og meningsfullt liv). Måleinstrumentet ga en score på psykologisk velvære. Instrumentet ble reliabilitetstestet med Chronbachs α . *Kessler Psychological Distress Scale (K10)* måler globalt distress (psykologisk lidelse - negativt stress basert på spørreskjema med 10 spørsmål om emosjonell tilstand). Spørsmålene handler om engstelse og depressive symptomer siste 4 uker. Poengskalaen gikk fra 0 (aldri) til 5 (hele tiden). Poengtelling lå fra 0 – 50. Lav poeng score indikerte en lite uttalt psykologisk lidelse. *Physical Self - Description Questionnaire* måler global selvpoppfatning med fokus på det ikke-akademiske domenet «selvpoppfattet fysisk fremtoning» (hierarkisk modell (Fox & Wilson, 2008)). Deltagerne selvrapporterte på Likert skala (1 = falsk til 6 = sann) som tilsvar til påstander (eks. Jeg er attraktiv for alderen og Jeg er tilfreds med hvem jeg er og hva jeg kan oppnå fysisk). Gjennomsnittsverdien ble beregnet fra total score dividert med antall enheter og ga en tallfesting av globalt fysisk selvbilde og selvpoppfattet fremtoning. Instrumentet ble reliabilitetsvurdert med Chronbachs α . *Feeling*

state questionnaire måler hva deltager føler før og etter hver høyintensitetsøkt (formen).

Deltagerne besvarte 1 spørsmål (Hvordan føler du deg akkurat nå?) på en 11 punktskala (-5 = svært dårlig til 5 = svært bra). Gjennomsnitt, pre- og post- intervensjon ble beregnet og målt for hver økt (Costigan et al., 2016).

St8. (Telles et al., 2013) bruker *Eurofit physical fitness test* som måler antropometri (lengde, vekt). *Flamingo balance testen* måler balanse (på ett bein). *Plate tapping test* måler øye-hånd koordinasjon ved å holde en hånd i sentrum og flytte den andre frem og tilbake (tapping) raskest mulig, innen en viss tid. *Standing broad jump* måler stille lengde. *Handgrip test* måler styrke i håndgrep med dynamometer. *Trunk strength* ble målt ved å telle antall situps i løpet av 30 sekunder. *Bent arm hang* ble målt ved å statisk henge i bom så lenge som mulig. *Shuttle run* (10*5 meter) måler hurtighet ved løp mellom to punkter med 10 meters avstand. *Children`s version of the Stroop color and word test (SCWT)* måler evne til å motstå kognitiv forstyrrelse. Testen tallfester viljeskontroll over nevropsykologiske funksjoner som inngår i respons både på ord og farge. Testen var tredelt. Først en måling av hvor fort deltagerne kunne lese ord. Den neste en test på hvor hurtig deltagerne kan navngi fargene ordene var skrevet i på siden. Den tredje testen instruerte deltagerne om å navngi fargen ordene var skrevet med (og ignorere bokstavbetydningen). Ved feil ble deltagerne stoppet og måtte selv korrigere feil før fortsettelse. Hver deltager fikk bruke 45 sekunder per test. *Indian adaptation of Battle`s self - esteem questinnaire (BSEI)* måler global selvpoppfatning («self-esteem») (Sonstroem, Harlow & Josephs, 1994) ved 50 lukkede spørsmål. Deltagerne ble instruert om å besvare spørsmål innen 4 domener: global selvpoppfatning, selvpoppfatning i sosial setting, akademisk selvpoppfatning og selvpoppfatning som forelder. Total tid for testen var 15 minutter. Lærerens rangering, med en seksdelt analog skala målte lydighet, akademisk prestasjon, oppmerksomhet, punktlighet, atferd med venner og atferd med lærere. Skalaen var en linjal fra 0 til 10 cm. Læreren markerte et merke på hver individuelle skala for alle deltagerne etter egen vurdering (Telles et at., 2013).

St9. (Ardoy et al., 2014) bruker *Spanish Overall and Factorial Intelligence Test, medium version (IGF-M)* som måler spesifikke kognitive elementer som: verbal og ikke-verbal dyktighet, abstrakt resonnement, spatial evne, verbalt resonnement og numerisk dyktighet. Hvert kognitive element hadde poenguttelling fra 0 til 100. Høyere prestasjon ga høyere score. En overordnet kognitiv prestasjon ble beregnet inneholdt gjennomsnitt fra alle målte kognitive variabler. Måleinstrumentet var på forhånd reliabilitets- og validitetstestet.

Deltagernes karakterer i kjernefag (matematikk, språk) og andre fag (naturfag, engelsk, osv.) var grunnlag for akademisk prestasjon. Datainnsamling foregikk i 2 omganger: første semester (desember 2006, pre intervensjon), og ved slutten av året (juni, 2007, post-intervensjon). Numeriske karakterer hadde rekkevidde 1 (svakest) til 10 (beste). Karakteren settes basert på elevens: holdninger, atferd, hjemmelekse, ferdigheter og kunnskap innen faget (etter spansk læreplan). Gjennomsnittskarakterer i matematikk, spansk (morsmål), engelsk (fremmedspråk), naturfag, samfunnsfag, teknologi, kunst og håndverk og musikk ble inkludert. Total karakterscore ble beregnet av gjennomsnittet for alle fag (Arday et al., 2014).

St10. (Tarp et al., 2016) bruker *Modifisert Eriksen Flanker task* som måler inhibering. Deltagerne ble instruert til å respondere på midterste pil ved å trykke med høyre hånd dersom pilen pekte mot høyre og venstre hånd når piler pekte mot venstre. Oppgavene ble presentert på skjermen både kongruente (>>>>) eller <<<< <) og inkongruente (<<<< eller >>>>). Forsøk har vist at inkongruente oppgaver gir en forsinket responstid og redusert nøyaktighet sammenlignet med kongruente oppgaver. Årsaken er at oppgaven krever en høyere respons inhibering (se St2. over). Hver oppgave ble presentert i 120 ms. Testomgivelsen var mest mulig lik for alle deltagerne. Testen ble gjennomført på samme måte ved baseline og ved oppfølging. Testen besto av $2 * 75$ oppgaver med kongruente og inkongruente oppgaver tilfeldig fordelt med lik sannsynlighet. Reaksjonstid (ms) og nøyaktighet ble registrert som hovedresultat. Nøyaktighet ble definert som prosentandelen korrekte svar. Reaksjonstid ble definert som responstiden på korrekte oppgaver. Målingene ble sammenfattet separat for kongruente og inkongruente oppgaver. Kun svar som ble avgitt innen 200 - 1470 ms ble akseptert. Interferens score ble beregnet som differansen mellom inkongruent og kongruent reaksjonstid og nøyaktighet. Deltagere som hadde færre enn 50% korrekte svar ble ekskludert fra analysen på bakgrunn av at deltageren ikke hadde forstått oppgaven. En *Custom made grade specific mathematics test* ble brukt for å måle akademisk prestasjon. Testen inkluderte aritmetikk, algebra, problemløsning og geometri, basert på pensum det aktuelle skoleåret. Totaltid for testen var 45 min. Testen hadde 50 oppgaver med poenguttelling fra 0 (svakest) – 50 (best). Ingen hjelpemidler var tillatt. Testen etter intervensjon ble gitt med samme oppgaver, men med ulike tall. Custom testen ble validert mot en standard prøve i matematikk (tilpasset pensum) med høy korrelasjonskoeffisient (Tarp et al., 2016).

St11. (Subramanian et al., 2015) bruker *Two target letter cancellation test (LCT)* som måler oppmerksomhet, spesielt visuell oppmerksomhet og opprettholdelse av oppmerksomhet.

Deltagerne ble instruert om å kansellere «E» og «C» blant 6 linjer med bokstaver, i det engelske alfabetet. Responstid på fullført oppgave bestemte poengscore (se St4. over). *Trail making Test A og B (TTA og TTB)* måler visuell oppmerksomhet og skifte av setting. Deltagerne ble instruert om å tegne linjer (trail A) mellom 25 nummererte sirkler i økende rekkefølge så hurtig og presist som mulig. Deltagerne ble instruert om å tegne linjer (trail B) mellom 25 tall og bokstaver med økende rekkefølge. Responstid på fullført oppgave bestemte poengscore (se St7. over). *Ruff figural fluency test (RFFT)* måler ikke-verbal kapasitet for initiativ, planlegging og resonnement. 40 «terninger» var angitt på et ark. Deltagerne ble instruert om å tegne stadig nye mønster mellom terningprikkene. Total tid for testen var 5 minutter. Antallet unike mønster ga en totalscore, rotasjonen av disse mønstrene og antallet mønster som ble fastholdt (perseverans) ble registrert (Subramanian et al., 2015).

St12. (Resaland et al., 2016) bruker *Standardized Norwegian National test (NDET)* som måler akademisk prestasjon i lesing, regning og engelsk (fremmedspråk). Testene ble gjennomført på 3 ulike dager ved baseline og post-intervensjon. Testene var på forhånd reliabilitets- og validitetstestet og tilpasset pensum for aktuelt klassetrinn. Resultatene ble analysert samlet og separat, for kognitiv funksjon og inhibering, kognitiv fleksibilitet og arbeidsminne (Aadland et al., 2017). Tre ulike tester *Goldens version of the Stroop test (StroopCW)* måler inhibering (se St8. over), *Verbal fluency (VF)* og *(nonverbal) trail making test (TMT)* (se St7. over) målte visuell oppmerksomhet og skifte av setting. *Wechsler Intelligence Scale for Children (WISC-IV)* og en *Digit span test* med rekapitulering av siffer i motsatt og samme rekkefølge måler intelligens og arbeidsminne (se St5. over). De testene som er mest utslagsgivende for eksekutiv funksjon ble brukt i analysen (StroopCW, total score i VF, WISC-IV backward og TMT del B). Analysen (per protokoll) inneholdt også sammensatt score for eksekutiv funksjon (Resaland et al., 2016).

St13. (Fedewa et al., 2015) bruker *Raven`s Standard Progressive Matrics (SPM)* som måler evne til abstrakt resonnement (ikke-verbal fluid intelligens). Testen gir et mål på generell situasjonsuavhengig prestasjon. Raven`s SPM består av 5 sett med figurer (A, B, C, D, og E). Hvert sett har 60 oppgaver som blir stadig vanskeligere. Hver figur står i forhold til hverandre. Deltagerne blir instruert om å finne den manglende figuren blant en liste med alternativer. Testen er på forhånd reliabilitetstestet og testet for intern konsistens. Testen ble gjennomført i klasserom med inntil 25 elever av gangen. Total tid for testen var 45 minutter.

Measures of Academic Progress (MAP) måler akademisk prestasjon i matematikk og lesing. Testscore på akademisk prestasjon ble friggitt fra kun fra 2 av 4 skoler (Fedewa et al., 2015).

St14. (Have et al., 2016c) bruker *Standardized mathematics test (MG)* som måler matematikkferdighet med forståelsen av tall størrelse, relasjon, addisjon, subtraksjon og geometri. 24 oppgaver med multiple choice var tilpasset aldersgruppen. Tid per oppgave var 1 minutt. Eksekutiv funksjon ble målt med *Flanker test* for å tallfeste inhibering (se St2. over). Testfigurer var 5 fisk, standard flanker (blå fisk, trykk i retning av senterfisk), en reversert flanker (rosa fisk, trykk i samme retning som fisk utenom senter) og blandet oppgave (blå fisk og rosa fisk). Testen målte kognitiv fleksibilitet. Stimulus ble presentert i 1500 ms. Responstiden ble registrert og følgende variabler ble beregnet: % nøyaktighet for kongruente oppgaver, % nøyaktighet for inkongruente oppgaver, reaksjonstid for korrekte kongruente svar, og reaksjonstid for korrekte inkongruente svar. *The Behavior Rating Inventory of Executive Function (BRIEF)* er spørreskjema som ble fylt ut av foreldre eller lærer og målte 8 variabler: initiativ, arbeidsminne, planlegging og organisering, organisering av materiell, årvåkenhet, inhibering, skifte av setting, emosjonell kontroll og metakognitiv indeks. Kartleggingen var todelt: *atferdsregulering (BRI)* og *eksekutive komponenter (GEC)*. Spørreskjema ble normert etter en kontrollgruppe. Cronbachs α viste høy intern konsistens. 30-minutters *standardized Figural Torrance Tests of creative Thinking (TTCT)* måler kreativ tankegang (fem deldomener: uttrykksevne, elaborasjon, originalitet, abstraksjon av titler). Deltagerne reflekterte over livserfaring med utgangspunkt i enkle figurer tilpasset alder og språk. Testen målte antall tolkbare, meningsfulle og relevante ideer som ble generert som respons til stimuli, videre ble fleksibilitet gitt av antall ulike kategorier relevant respons. Derneft, ble originalitet målt basert på statistisk sjeldenhet i respons, og til sist, elaborering basert på utdyping av antall ideer. Tid per aktivitet var 10 minutter. I aktivitet 1 ble deltagerne instruert om å tegne et bilde med utgangspunkt i en pæreform eller en bønneform. I aktivitet 2 ble 10 ufullstendige figurer brukt til å lage et helt bilde. I aktivitet 3 ble 3 sider med linjer eller sirkler brukt til å sette sammen et bilde. Ingen hjelpemidler var tillatt (Have et al., 2016c).

St15. (Ahamed et al., 2007) bruker *Canadian Achievement Test (CAT-3)* som måler akademisk prestasjon. CAT-3 er en standardisert test som måler deltagerens kunnskap og ferdighet i spesifikke fag, som matematikk, lesing, og språk (morsmål). Testen ble kontrollert for alder, kjønn og etnisitet. Innhold og tema ble hentet fra geografi, kultur og etnisk

mangfold i det canadiske samfunnet. Testene ble gjennomført ved et testsenter (Markham, Ontario) med unntak av en kontrollskole som inkluderte testen lokalt. Totalscore (Totscore) ble brukt for å summere resultatene fra matematikk, lesing og språk og utgjorde deltagerens akademiske prestasjon (Ahamed et al., 2007).

St16. (Mullender – Wijnsma et al., 2016) bruker *Speed Test-Arithmetic* som måler matematikkhurtighet. Oppgavene skulle løses raskest mulig. Testen besto av 5 rekker med summasjoner. Deltageren ble instruert om å løse flest mulig regneoppgaver. Maksimal tidsbruk ble satt til 1 minutt per rekke. Poenguttelling ble beregnet fra antall korrekte oppgaver (maksimalt 200). *Child academic monitoring system (CAM)* er et standardisert og normbasert testbatteri som måler staveprestasjon og matematikkferdighet. Stavetesten var todelt: I første del ble deltagerne instruert om å stave ordet etter at læreren leste en setning og gjentok testordet. I den andre delen ble deltagerne instruert om å identifisere feilstavede ord. Testene ble på forhånd reliabilitets- og validitetstestet. Matematikktesten inneholdt tallforståelse, aritmetikk, algebra, geometri, tid og pengeverdi, kunnskap om desimaltall og brøk. Testen var tilpasset klassetrinn (Mullender- Wijnsma et al., 2016).

St17. (Quinto Romani & Klausen, 2017) bruker *Grade point average (GPA) karakterer* som gir et mål på skoleprestasjon. Datainnsamling av karakterer ble foretatt i 2011/2012 da deltagerne var på trinn 9. Intervensjonen foregikk fra trinn 6 (2008) til trinn 8 (2010). GPA ble samlet inn en stund etter intervensjonsperioden. I 2008 ble GPA samlet inn for hver skole generelt fra 9 trinn elever (baseline). Disse resultatene ga en differanse i GPA mellom intervensjon- og kontrollgruppe og baseline (Quinto Romani & Klausen, 2017). Differansen ble brukt for å dokumentere bedre resultater etter intervensjon sammenlignet med kontroll (Quinto Romani & Klausen, 2017).

8.2 Vedlegg 2: Tverrfaglig databasesøk.

Tabell 26: Tverrfaglig databasesøk. Tabellen leses nedenfra og opp med S1 som første søk. Query er søkestrengen. S1, S2, S3 og S4 er enkeltsøk for utfall, populasjon, intervensjon, og forskningsdesign. Enkeltsøk gir bredde (med logisk ELLER). S5, S6, S7, S8 gir presisjon i søket (med logisk OG). Limiters brukes for S6, S7 og S8. Last run via navngir databasene som er brukt (SPORTDiscus, CINAHL Plus with Full Text, MEDLINE, Academic Search Complete, Health and Psychosocial Instruments, ERIC under EBSCOhost). Results tallfester antall treff for hvert søk (og reduseres til 1807 i S8). Disse artiklene blir undersøkt nærmere (jmf. 3.3 Utvelgelse av litteraturreferanser ved PRISMAS flytdiagram).

#	Query	Limiters/Expanders	Last Run Via	Results
S8	S1 and S2 AND S3 AND S4	Limiters - Peer Reviewed; Published Date: 20070101-20171231 Narrow by Language: - english Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - SPORTDiscus; CINAHL Plus with Full Text; MEDLINE; Academic Search Complete; Health and Psychosocial Instruments; ERIC	1,807
S7	S1 and S2 AND S3 AND S4	Limiters - Peer Reviewed; Published Date: 20070101-20171231 Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - SPORTDiscus; CINAHL Plus with Full Text; MEDLINE; Academic Search Complete; Health and Psychosocial Instruments; ERIC	1,813
S6	S1 and S2 AND S3 AND S4	Limiters - Published Date: 20070101-20171231 Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - SPORTDiscus; CINAHL Plus with Full Text; MEDLINE; Academic Search Complete; Health and Psychosocial Instruments; ERIC	2,124
S5	S1 and S2 AND S3 AND S4	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - SPORTDiscus; CINAHL Plus with Full Text; MEDLINE; Academic Search Complete; Health and Psychosocial Instruments; ERIC	2,280
S4	randomized controlled trials or rct or randomized control trials or random control trial	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - SPORTDiscus; CINAHL Plus with Full Text; MEDLINE; Academic Search Complete; Health and Psychosocial Instruments; ERIC	372,968
S3	Activity or aerobic exercise or physical activity or physical fitness or exercise or movement	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - SPORTDiscus;	8,320,582

	skills or physical education or motor skills training or motor skills or physical		CINAHL Plus with Full Text; MEDLINE; Academic Search Complete; Health and Psychosocial Instruments; ERIC	
S2	infancy or childhood or children or adolescents or pupil or young people	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - SPORTDiscus; CINAHL Plus with Full Text; MEDLINE; Academic Search Complete; Health and Psychosocial Instruments; ERIC	4,016,309
S1	learning or school health or executive function or executive functioning or cognition or cognitive function or cognitive performance or cognitive ability or cognitive engagement or intelligence or academic achievement or academic performance or scholastic achievement or grading	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - SPORTDiscus; CINAHL Plus with Full Text; MEDLINE; Academic Search Complete; Health and Psychosocial Instruments; ERIC	2,919,828

8.3 Vedlegg 3: Ekskluderte artikler med begrunnelse

Fulltekst artikler ble ekskludert med begrunnelse (n=15) (jmf. 3.3 Utvelgelse av litteraturreferanser ved PRISMA's flytdiagram).

Asfour, L., Koussa, M., Perrino, T., Stoutenberg, M. & Prado, G. (2016). The association of organized and unorganized physical activity and sedentary behavior with internalizing and externalizing symptoms in Hispanic adolescents. *Child & Adolescent Mental Health*, 21(2), 109-114. doi:10.1111/camh.12139

Eksklusjon: Problemstillingen var ikke relevant.

Blair, C. & Raver, C. C. (2014). Closing the Achievement Gap through Modification of Neurocognitive and Neuroendocrine Function: Results from a Cluster Randomized Controlled Trial of an Innovative Approach to the Education of Children in Kindergarten. *PLoS ONE*, 9(11), 1-13.

Eksklusjon: Problemstilling var ikke relevant.

Britton, W. B., Lepp, N. E., Niles, H. F., Rocha, T., Fisher, N. E. & Gold, J. S. (2014). A randomized controlled pilot trial of classroom-based mindfulness meditation compared to an active control condition in sixth-grade children. *Journal of School Psychology, 52*(3), 263-278.

Eksklusjon: Intervensjonen inneholdt ikke fysisk aktivitet.

Dodick, D., Starling, A. J., Wethe, J., Pang, Y., Messner, L. V., Smith, C., Master, C. L., Halker-Singh, R. B., Vargas, B. B., Bogle, J. M., Mandrekar, J., Talaber, A. & Leong, D. (2017). The Effect of In-School Saccadic Training on Reading Fluency and Comprehension in First and Second Grade Students. *Journal of Child Neurology, 32*(1), 104-111. doi: 10.1177/0883073816668704

Eksklusjon: Problemstilling var ikke relevant.

Gallotta, M. C., Emerenziani, G. P., Iazzoni, S., Iasevoli, L., Guidetti, L. & Baldari, C. (2017). Effects of different physical education programmes on children's skill- and health-related outcomes: a pilot randomised controlled trial. *Journal of Sports Sciences, 35*(15), 1547-1555.

Eksklusjon: Problemstillingen var ikke relevant.

Hardy, J. L., Nelson, R. A., Thomason, M. E., Sternberg, D. A. Katovich, K., Farzin, F. & Scanlon, M. (2015). Enhancing Cognitive Abilities with Comprehensive Training: A Large, Online, Randomized, Active-Controlled Trial. *PLOS ONE, 10*(9), 1-17. doi:10.1371/journal.pone.0134467

Eksklusjon: Intervensjonen inneholdt ikke fysisk aktivitet.

James-Burdumy, S., Bleeker, M., Beyler, N., London, R. A., Westrich, L., Stokes- Guinan, K. & Castrechini, S. (2013). *Does Playworks Work? Findings from a Randomized Controlled Trial*. Paper presentert på *Society for Research on Educational Effectiveness, SREE*.

Eksklusjon: Problemstillingen var ikke relevant.

Kamijo, K. (2016). Physical activity, aerobic fitness, and cognitive function in children. *Advances in Exercise & Sports Physiology, 22*(4), 77-77.

Eksklusjon: forskningsdesign var ikke RCT eller CRT.

Keeley, T. J. H. and K. R. Fox (2009). The impact of physical activity and fitness on academic achievement and cognitive performance in children. *International Review of Sport & Exercise Psychology*, 2(2), 198-214.

Eksklusjon: Review (ikke en primærstudie).

Matthews, M. J., Yusuf, M., Doyle, C. & Thompson, C. (2016). Quadrupedal movement training improves markers of cognition and joint repositioning. *Human Movement Science*, 47, 70-80.

Eksklusjon: Utvalget i alder over 18 år.

Prakash, R. S., Voss, M. W., Erickson, K. I. & Kramer, A. F. (2015). Physical Activity and Cognitive Vitality. *Annual Review of Psychology*, 66(1), 769-797.

Eksklusjon: Review (ikke en primærstudie).

Riley, N., Lubans, D. R. Holmes, K. & Morgan, P. J. (2014). Rationale and study protocol of the EASY Minds (Encouraging Activity to Stimulate Young Minds) program: cluster randomized controlled trial of a primary school-based physical activity integration program for mathematics. *BMC Public Health*, 14(1), 816-816.

Eksklusjon: Populærvitenskapelig tidsskrift?

Sutherland, R., Campbell, E., Lubans, D., Morgan, P., Nathan, N., Okley, A., Gillham, K., Davies, L. & Wiggers, J. (2017). Physical Activity 4 Everyone cluster RCT: 24-month physical activity outcomes of a school-based physical activity intervention targeting adolescents. Overall and school day physical activity outcomes. *Journal of Science & Medicine in Sport*, 20, 106-128.

Eksklusjon: Problemstilling ikke relevant.

Tse, A. C. Y., Fong, S. S. M., Wong, T. W. L. & Masters, R. (2017). Analogy motor learning by young children: a study of rope skipping. *European Journal of Sport Science*, 17(2), 152-159. doi: 10.1080/17461391.2016.1214184

Eksklusjon: Ikke relevant problemstilling.

Westendorp, M., Houwen, S., Hartman, E., Mombarg, R., Smith, J. & Visscher, C. (2014).

Effect of a Ball Skill Intervention on Children's Ball Skills and Cognitive Functions.

Medicine & Science in Sports & Exercise, 46(2), 414-422. doi:

10.1249/MSS.0b013e3182a532b3

Eksklusjon: Barn utover normal utvikling.

8.4 Vedlegg 4: Prosjektbeskrivelse for Masteroppgave i Idrett

Litteraturstudie med arbeidstitel: Effekten av fysisk aktivitet på kognitiv funksjon

Student: Erik Borgersen

Veileder: Yngvar Ommundsen

Prosjektplan

Prosjektplanen har en ramme på ca 3000 ord. Punktene nedenfor beskriver prosjektplanen i henhold til retningslinjene i Masterhåndbok avsnitt 4.0 Prosjektplan.

1. Forskningsområde: Effekten av fysisk aktivitet på kognitiv funksjon

Bakgrunn for valg av tema:

Ifølge Ommundsen (2013) kan fysisk-motorisk ferdighet ha betydning for kognitiv funksjon og skoleprestasjoner. I lys av nyere forskning som viser at økt fysisk aktivitet kan være viktig for elevers kognitive funksjon og læring i skolen, nevnes en rekke måter å nærme seg problemstillingen. Nyere perspektiver innen nevro-psykologisk forskning (Hillman, Kamijo & Scudder, 2011) og empiriske studier, en revidert av Tomporowski, Davis, Miller & Naglieri (2008), støtter at systematiske treningsprogrammer kan fremme mentale prosesser som har betydning for akademisk prestasjon.

I en longitudinell studie (Ericsson & Karlsson, 2014), fulgte man svenske elever fra begynnelsen til slutten av grunnskoletrinnet. Intervensjonsgruppen fikk kroppsøving 5 ganger pr. uke, og motorisk svakt fungerende elever fikk 1 time ekstra pr. uke med vekt på motorisk trening fra 1. til 9. skoleår. Kontrollgruppen fikk ordinær kroppsøvingundervisning etter læreplanen to ganger pr. uke. I niende klasse scoret intervensjonsgruppen signifikant bedre på motorisk ferdighet. Guttene i intervensjonsgruppen oppnådde bedre karakterer i teoretiske fag, og man fant sammenheng mellom bedre skoleprestasjoner og bedre motoriske ferdigheter.

Resultater fra en studie av Fedewa og Ahn (2011) viser også verdien av fysisk aktivitet for kognitiv funksjon. Artikkelen poengterer at denne type målinger er kompliserte. Studien konkluderer med gevinst for barn og unges kognitive funksjoner og skoleprestasjoner (særlig lesing og matematikk) hvis man i skolesammenheng fokuserer på økt fysisk aktivitet.

I en annen oversiktsartikkel, et langtidsdesign (Singh, Uijtdewilligen, Twisk, Van Mechelen & Chinapaw, 2012), er forfatterne noe mer avventende, grunnet mangel på studier med tilstrekkelig høy kvalitet; det vil si kontrollerte og longitudinelle studier.

2. Litteraturgjennomgang og teoretiske rammeverk

Litteraturstudien innebærer å søke relevant litteratur. Disse studiene blir gjort tilgjengelig underveis i prosjektet.

I en litteraturstudie av Best (2010) undersøkes effekten av akutt og kronisk aerob trening på eksekutiv funksjon. Forskningsdesign er konstruert ved enten umiddelbart å måle kognitiv effekt av en aerob økt (10 min til 1 time) eller måle kognitiv effekt av differansen før og etter noen ukers aerob trening. Ved akutt aerob trening har Best undersøkt effekten av ulike former for trening (individuell vs. gruppetrening). Ved aerob trening over noen uker har Best undersøkt aerob spill (løpsspill, modifisert basketball eller fotball) med ulik øktvarighet (20 minutter og 40 minutter) sammenlignet med kontrollgruppe i hvile. Resultatene pekte i retning av en effekt på eksekutiv funksjon etter enkeltøkter. Effekten var bedre hvis oppgavene under trening stimulerte eksekutiv funksjon. Ved aerob trening over tid fant Best en varig forbedring av eksekutiv funksjon. Noen underliggende mekanismer forklarer hvorfor aerob trening har en effekt på eksekutiv funksjon.

1. Måltrettet trening og det å forplikte seg til trening krever en kognitiv kapasitet. Barn trenger komplekse kognitive egenskaper for å samarbeide med jevnaldrende, for å kunne forutse spillbevegelser hos lagkamerater og konkurrenter, anvende strategier, og tilpasse seg stadige endrede oppgavekrav (Best, 2010).

2. Kompleks motorisk bevegelse krever en kognitiv kapasitet. Barn aktiverer nevrologiske nervestrukturer som forbindes med eksekutiv funksjon. Cerebellum er viktig for kompleks og koordinert bevegelse. Dorsolateral prefrontal korteks (DL- PFC) er viktig både for kompleks motorfunksjon og for kognitive funksjoner. Hjernen stimuleres ulikt ved å operere ved ulik modus. Et eksempel er at hjernen skiller mellom ensidige repeterte oppgaver som å gå og komplekse oppgaver med kognitive krav som strategiske spillalternativer på fotballbanen (Best, 2010).

3. Aerob trening fører til fysiologiske tilpasninger i hjernen. Den fysiologiske endringen som følger av trening er ulik for akutt aerob trening og kronisk trening over tid. Trening over tid påvirker ulike vekstfaktorer (insulin-like growth factor IGF-1, vaskular endothelial growth factor VEGF og brain-derived neurotrophic factor BDNF). BDNF er en viktig mediator for synaptisk transmisjon og synaptisk plastisitet. BDNF bidrar til utviklingen av nye nevroner (er mediator for treningsindusert nevrogenese), og legger grunnlag for bedret læring og hukommelse ved en økt nevrogenese i Hippocampus. Eks. dyreforsøk – rotter, løp på tredemølle kan bedre den visuell-spatial hukommelsen og øke konsentrasjonen av nerveceller

i Hippocampus. Akutt trening innvirker med umiddelbare nevrokjemiske endringer som forbereder sentralnervesystemet (CNS) på øyeblikkelig eller påfølgende ferdighetstilegnelse som kreves i den fysiske aktiviteten. Den kjemiske endringen kan vekke individet til økt kognitiv prestasjon (Best, 2010).

En litteraturstudie (Tomporowski et al., 2008) undersøker effekten av fysisk aktivitet på barns intelligens, kognisjon og akademisk prestasjon. Teorigrunnlaget i studien er kognitiv læringsteori som retter seg mot fysisk aktivitet. Problemstillingen er belyst for effekten av trening på kognitiv funksjon hos voksne, og studier gjennomført på dyr har gitt resultater med sammenheng til nevrologisk utvikling. Trening fasiliterer barns eksekutiv funksjon. Det vil si prosesser som er avgjørende for valg, organisering, og initiering av målrettet handling. I følge Tomporowski et al. (2008) har eksekutiv funksjon en sterk sammenheng med fysisk aktivitet. I tillegg blir prosessering av kontroll, visuell-spatial prosessering og reaksjonstid nevnt å ha sammenheng med aerob trening. Resultatene ble tolket som evidens på en kausal kobling mellom fitnessnivå og vitalitet. Eksekutiv funksjon inngår i planlegging, valg av strategier som organiserer målrettet handling og skilles fra grunnleggende prosesser for informasjonsbehandling (koding, vurdering av stimuli, valg av respons og utførelse av respons). Eksekutiv funksjon har en rekke underliggende mekanismer. Testing av voksne utøvere har resultert i tre variabler. *Skifte av setting*, som innebærer at individet kutter ut irrelevante oppgaver og engasjerer seg i en relevant oppgave. *Oppdatering*, som koples til arbeidsminnet og innebærer en egenvurdering av mental fremtoning. Og *inhibering*, som innebærer en undertrykking av impulsiv respons. Hypotesen om påvirkning av eksekutiv funksjon har blitt testet på eldre utøvere ved å undersøke effekten av aerob trening både på ikke eksekutiv og eksekutiv funksjon. Det ble avdekket en differanse for utøvere før og etter treningsintervensjon, på en rekke kognitive tester.

Motsatsen til en god eksekutiv funksjon er manglende atferdskontroll, manglende oppmerksomhet og vurderingsevne. Barn med redusert eksekutiv-funksjon vil ha problemer med å følge med og utvikle seg akademisk i klasserommet. Videre vil evnen til å kontrollere eller hemme respons være redusert. Responsen har betydning for kapasitet til utvikling av fantasi, erfare empati, kreativitet, og selvevaluering av egne tanker og handlinger (Tomporowski et al., 2008). Eksekutiv funksjon vil være en forutsetning for akademisk måloppnåelse. Fysisk aktive barn har bedret *kognitiv atferdskontroll* og øker sin *konsentrasjonsevne* og *oppmerksomhet* i innlærings situasjoner (Singh et al., 2012).

3. Problemstilling formuleres og drøftes

På samme måte som valg av problemstilling har konsekvens for valg av design, har også valg av evidens i form av randomiserte eksperiment (Randomized controlled trials, RCT design) konsekvens for problemstillingen. I litteraturen (Vazou, Pesce, Lakes & Smiley-Oyen, 2016) har forskere brukt RCT studier for å undersøke sammenhenger mellom fysisk aktivitet og kognisjon med et ønske om å kunne forklare årsak og virkning. Vazou et al. (2016) nøyer seg likevel med å predikere en sammenheng og mangler en kausal forklaring i konklusjonen (Vazou et al., 2016). Keeley & Fox (2009) etterlyser design som har en kvalitet og dybde i evidens ut over korrelasjonsstudier med svake sammenhenger. På bakgrunn av evidens bør operasjonaliserte mål på fysisk aktivitet kunne forklare operasjonaliserte mål på kognisjon. Donnely et al., (2016), hevder at avansert teknologi (Magnetic resonance imaging MRI og Elektroencefalogram EEG) er nødvendig for å danne anatomiske og biologiske modeller for å etablere en biologisk basis for observert effekt på kognisjon og akademisk prestasjon.

Denne litteraturstudien tar sikte på å besvare følgende problemstilling: *Kan fysisk, aerob, motorisk aktivitet, motorisk ferdighet og tradisjonell kroppsøving være årsak til kognitiv psykologisk respons som eksekutiv funksjon sammen med akademiske tester?*

Litteraturstudien er en kritisk oppsummering av enkeltstudier siste 10 år, og gir en oppdatering av kunnskap om effekten av fysisk aktivitet på kognisjon.

4. Metode velges og begrunnes med utgangspunkt i problemstillingen

Protokoll beskriver en systematisk utvelgelse og kritisk gjennomgang av studier. De ulike studiene bør ha en lik eller lignende problemstilling. Avhengig og uavhengig variabel bør være like eller lignende. Variablene som brukes for fysisk aktivitet blir konkretisert og beskrevet som aerob, motorisk ferdighet, eller motorisk aktivitet med kognitivt innhold. Variablene som brukes for kognisjon er kognitiv psykologisk respons og innebærer en presisering av eksekutiv funksjon som har basis i etablerte modeller. Akademisk prestasjon måles i etablerte ferdighetstester.

Studiepopulasjonen bør være like eller lignende. En sterk design bør kunne legge til rette for høyt evidensnivå. Ulike studier bør resultatmessig peke i samme retning. Tiden satt av til undersøkelsen bør være lik eller lignende. Begreper bør defineres og skilles på ulike nivå (fysisk aktivitet, motorisk aktivitet). Søkestreng utarbeides. Medical subject headings (MeSH) brukes som indeksert begrep for presise søk. PRISMA 2009 Flow Diagram er et standardisert flytdiagram og gir en oversikt over eksklusjon og inklusjon av studier.

- **Søkestrategier**

- Litteratursøk i elektroniske databaser, sensitive og spesifikke søk.
- Kildehenvisninger i tidligere reviews og meta analyser.
- Kildehenvisning i relevante artikler.
- Søk i tidsskrift.

- **Inklusjonskriterier,**

Fysisk aktivitet:

- Aerob aktivitet, løping, hoppetau, spill med intensitet,
- motorisk aktivitet, ferdighetsutvikling,
- tradisjonell kroppsøving, fritt spill, basketball, volleyball, badminton, bordtennis,
- motorisk ferdighet, grunnleggende ferdighet, koordinasjon, balanse, lateralitet, objekt kontroll

Kognitiv psykologisk respons:

- eksekutiv funksjon, akademisk måloppnåelse, planlegging, kreativitet, arbeidsminne, romlig oppfatning og tankegang, matematikkompetanse, lesekompetanse, staving, reaksjonstid og nøyaktighet, oppmerksomhet, rekapitulering, kognitive skifter, språk, perseveranstendens (evne til opprettholdelse av aktivitet).

- **Eksklusjonskriterier,**

- Dersom intervensjonen ikke inneholder fysisk aktivitet.
- Deltagere med alder > 18 år eller barn utover normal utvikling.
- Tverrsnittsstudier eller korrelasjonsstudier
- Eksperimenter med enkeltøkter
- Studier uten kontroll- eller sammenligningsgruppe
- Dersom avhengig variabel kun er akademisk prestasjon og ikke inneholder en direkte måling av kognitiv funksjon.
- Studiekvalitet, vurdering av eksklusjon basert på antallet mangler eller spesifikke mangler. Vurderingen baseres på sjekklister fra Kunnskapscenteret oversatt fra Critical Appraisal Skills Program (CASP).
- Vurdering av risiko for systematiske avvik i resultater (Higgins, Green & Cochrane Collaboration, 2009).

5. Dataekstrahering – koding

Resultatene i denne litteraturstudien vil inneholde datasett som er produsert for statistisk analyse.

- Deskriptive data for hver studie vil inneholde årstall, opprinnelsesland, metodeinformasjon, utvalg (størrelse), randomisering, blinding, frafall, oppfølgingsperiode osv.
- Resultatene fra enkeltstudier blir presentert med hensiktsmessige statistiske begreper.
- Statistisk analyse for hver studie vil inneholde vektet gjennomsnitt basert på studiens størrelse, studiekvalitet (?), heterogenitet osv.
- Definerings av effekt. Frequency effect size er styrken på funnene. Intensity effect size er konsentrasjonen av funn innen den enkelte rapport.
- Beregning av effektstørrelse og konfidensintervall

6. Utvalget

Utvalget er gitt av inkluderte studier. Populasjonen bør kartlegges med tanke på alder, utdanning, inntekt, kontekst, landsbygd, by, helsestatus osv.

7. Detaljert tids- og arbeidsfordelingsplan

Tabell 27: Tids- arbeidsfordelingsplan for mai 2017 til mai 2018. Planen er utarbeidet som et dynamisk dokument. Poster som følger planlagt fremdrift blir fortløpende avmerket med font «bold».

Plan og fremdrift Bold – planen fulgt	Mai 2017	Juni	Aug	Sept	Okt	Nov	Des	Jan 2018	Feb	Mars	April	Mai
Veileder-møter	X	X	X	X	X	X	X	X	X	X	X	X
Prosjektplan	Oppstart 20.04	Inn-levering 15.06										
Litteratursøk			X	X					X			
Lese artikler			X	X	X				X			
Presentasjon av resultater				X	X	X				X		
Diskutere resultater					X	X	X				X	
Revidering							X	X	X			
Skrive-prosess	X	X	X	X	X	X	X	X	X	X	X	X
Avslutning												X

8. Etisk overveielse

Litteratursøk legger grunnlag for datainnsamlingen. Studiene som inkluderes i denne studien er tidligere behandlet i respektive etiske komiteer. St. meld. Nr. 18 (2012 – 2013) Lange linjer – kunnskap gir muligheter, behandler en rekke forskningsetiske tema og områder.

Undertegnede rolle i håndtering av forskningsartikler er vurdert opp mot stortingsmeldingen. Forskningsmetoden vil bli beskrevet transparent og etterprøvbart. Det forskningsetiske perspektivet er derfor ivarettatt (Kunnskapsdepartementet, 2013, s. 32).

9. Budsjett for prosjektet

Ingen utgifter påløper prosjektet.

10. Innleveringsfrister

- i. **Forskningsskisse**, torsdag **15.06.2017** eller etter avtale med veileder. Kan mailes eller legges i posthylla til Yngvar Ommundsen.
- ii. **Masteroppgaven innleveres** 15. mai 2018.
- iii. **Muntlig eksamen** avholdes i løpet av mai 2018.

11. Referanser

Best, R. J. (2010). Effects of physical activity on children`s executive function: Contributions of experimental research on aerobic exercise. *Development Review*, 30 (4), 331- 351.

Higgins, J., Green, S. & Cochrane Collaboration. (2009). *Cochrane handbook for systematic reviews of interventions*. Hentet fra: <http://training.cochrane.org/handbook>

Donnelly, E. J., Hillman, H. C., Castelli, L., D., Etnier, N. J., Lee, N. S., Tomporowski, N. P., Lambourne, N. K. & Szabo-Reed, N. A. (2016). Physical Activity, Fitness, Cognitive Function, and Academic Achievement in Children: A Systematic Review. *Medicine & Science in Sports & Exercise*, 48(6), 1223-1224.

Ericsson, I. & Karlsson, M. (2014). Motor skills and school performance in children with daily physical education in school – a 9-year intervention study. *Scandinavian Journal of Medicine & Science in Sports*, 24(2), 273-278.

Fedewa, A. L. & Ahn, S. (2011). The Effects of Physical Activity and Physical Fitness on Children's Achievement and Cognitive Outcomes: A Meta-Analysis. *Research Quarterly for Exercise and Sport*, 82(3), 521-535.

- Hillman, C. H., Kamijo, K. & Scudder, M. (2011). A review of chronic and acute physical activity participation on neuroelectric measures of brain health and cognition during childhood. *Preventive Medicine*, 52, 21- 28.
- Keeley, T. & Fox, K. (2009). The impact of physical activity and fitness on academic achievement and cognitive performance in children. *International Review of Sport and Exercise Psychology*, 2(2), 198-214.
- Kunnskapsdepartementet. (2013). *Lange linjer - kunnskap gir muligheter*. (Meld. St. 18 2012-2013), Hentet fra [https://www.regjeringen.no/no/dokumenter/meld-st-18-20122013/id716040/?q=Lange linjer kunnskap gir muligheter](https://www.regjeringen.no/no/dokumenter/meld-st-18-20122013/id716040/?q=Lange+linjer+kunnskap+gir+muligheter).
- Ommundsen, Y. (2013). Fysisk-motorisk ferdighet gjennom kroppsøving – et viktig bidrag til elevenes allmenndanning og læring i skolen. *Norsk pedagogisk tidsskrift*, 97(02), 155-166.
- Singh, A., Uijtdewilligen, L., Twisk, J., Van Mechelen, W. & Chinapaw, M. (2012). Physical Activity and Performance at School: A Systematic Review of the Literature Including a Methodological Quality Assessment. *Archives of Pediatrics & Adolescent Medicine*, 166(1), 49-55.
- Tomporowski, P., Davis, D., Miller, C. & Naglieri, L. (2008). Exercise and Children's Intelligence, Cognition, and Academic Achievement. *Educational Psychology Review*, 20(2), 111-131.
- Vazou, S., Pesce, C., Lakes, K. & Smiley-Oyen, A. (2016). More than one road leads to Rome: A narrative review and meta-analysis of physical activity intervention effects on cognition in youth. *International Journal of Sport and Exercise Psychology*, 1-26.