



Høgskulen på Vestlandet

Masteroppgave - MASPED3-303

MASPED3-303-O-2021-VÅR-FLOWassign

Predefinert informasjon

Startdato:	30-04-2021 09:00	Termin:	2021 VÅR
Sluttdato:	14-05-2021 14:00	Vurderingsform:	Norsk 6-trinns skala (A-F)
Eksamensform:	Masteroppgave	Studiepoeng:	40
Flowkode:	203 MASPED3-303 1 O 2021 VÅR		
Intern sensor:	(Anonymisert)		

Deltaker

Navn:	Hege Sandal Starnesfet
Kandidatnr.:	228
HVL-id:	575333@hvl.no

Informasjon fra deltaker

Antall ord *:	21495
----------------------	-------

Egenerklæring *: Ja

Jeg bekrefter at jeg har ja registrert oppgavetittelen på norsk og engelsk i StudentWeb og vet at denne vil stå på vitnemålet mitt *:

Jeg godkjenner avtalen om publisering av masteroppgaven min *

Ja

Er masteroppgaven skrevet som del av et større forskningsprosjekt ved HVL? *

Nei

Er masteroppgaven skrevet ved bedrift/virksomhet i næringsliv eller offentlig sektor? *

Nei

MASTEROPPGAVE

En systematisk litteraturgjennomgang av
kvalitet i instrumenter brukt til språkkartlegging
i innføringstilbud i tidligere Hordaland fylke

A Systematic Review of the Quality of Language
Mapping Tests in Introductory Courses in Previously
Hordaland County

Hege Stamnesfet

MASPED3-303 Master i spesialpedagogikk
Fakultet for lærerutdanning, kunst og idrett
Institutt for pedagogikk, religion og samfunnsfag
Veiledere: Arne Ola Lervåg og Hilde Hofslundsengen
Innleveringsdato: 14.05.2021

Jeg bekrefter at arbeidet er selvstendig utarbeidet, og at referanser/kildehenvisninger til alle
kilder som er brukt i arbeidet er oppgitt, jf. Forskrift om studium og eksamen ved Høgskulen
på Vestlandet, § 10.

Innholdsliste

Innholdsliste	2
Forord.....	4
Sammendrag.....	5
Abstract	6
1. Innledning	7
1.1. Formål og hensikt	7
1.2. Problemstilling.....	8
1.3. Begrepsavklaring	8
1.3.1. Test eller prøve?	8
1.3.2. Elevgruppen.....	9
1.3.3. Hva er ‘tilstrekkelige ferdigheter i norsk’?.....	10
1.3.4. Læreren som portvokter?.....	12
2. Teori.....	13
2.1. Krav om kartlegging	13
2.2. Prøvene	13
2.2.1. Kartleggingsmaterieill: Språkkompetanse i grunnleggende norsk	15
2.2.2. Den europeiske språkpermen	15
2.2.3. Vestlandsverktøyet – Vurderingsverktøy for minoritetsspråklige elever i VGS	15
2.2.4. FLORO	16
2.2.5. TOSP	16
2.2.6. Norskprøve 2 og 3 for voksne innvandrere og Norskprøven for voksne innvandrere... ..	17
2.2.7. Migranorsk	17
2.2.8. Trondheimstesten.....	18
2.3. utfordringer med kartleggingsprøvene.....	18
2.4. Språk og ferdigheter.....	19
2.4.1. Lesferdigheter	20
2.4.2. Lytteferdigheter	22
2.4.3. Skriftlige ferdigheter	23
2.4.4. Muntlige ferdigheter.....	24
2.5. Norsk som andrespråk.....	24
2.6. Hva er språktesting?.....	27
2.6.1. Reliabilitet	27
2.6.2. Validitet.....	31

2.7.	Det felles europeiske rammeverket for språk	34
3.	Metode	37
3.1.	Spørreskjema, utvalg og prosedyre for datainnsamling	37
3.2.	Systematisk litteraturgjennomgang av andrespråkstester	39
3.3.	EFPA Test Review Model	39
3.4.	Validitet og reliabilitet	41
3.5.	Analyse	41
4.	Resultater	42
4.1.	Hvilke prøver bruker skolene?	42
4.2.	Vurdering av validitet og reliabilitet i instrumentene	45
5.	Drøfting.....	52
5.1.	Hvilke kartleggingsprøver brukes, hvor ofte og til hvilket formål?.....	52
5.2.	Hvilken validitet og reliabilitet har måleinstrumentene?	53
5.3.	Betydningen av få gode kartleggingsprøver	56
5.4.	Hva er tilstrekkelige norskferdigheter?	58
5.5.	Hva er kriteriene for språkprøvene?.....	59
5.6.	Implikasjoner for praksis og mulige konsekvenser.....	61
6.	Avslutning og veien videre	64
	Kilder.....	66
	Vedlegg 1: Spørreskjema	74
	Vedlegg 2: Informasjonsskriv til skolene.....	75
	Vedlegg 3: Samtykkeskjema	76
	Vedlegg 4: Godkjennelse fra NSD	79
	Vedlegg 5: Veiledende normering av <i>FLORO</i>	81
	Vedlegg 6: Normering av <i>TOSP</i>	83
	Vedlegg 7: Det felles europeiske rammeverket for språk	84

Forord

Denne masteroppgaven har vært en lang prosess som startet i det små allerede før masterstudiet. Innholdet i oppgaven har endret seg drastisk fra de spede førsteutkastene, og veien fra start til slutt har vært en utrolig interessant og lærerik prosess. Jeg vil særlig takke mine veiledere, Arne Lervåg og Hilde Hofslundsengen, som har stilt spørsmål, kommentert og oppmuntret underveis.

Sammendrag

Minoritetsspråklige elever i grunnskole og videregående skole har rett på særskilt språkopplæring, jf. §2-8 og §3-12 i Opplæringslova, dersom elever ikke har ‘tilstrekkelig ferdigheter i norsk’. Fokuset i denne masteroppgaven har vært å se på hvilke kartleggingsmateriell skolene i tidligere Hordaland fylke bruker for å kartlegge norskferdigheter i innføringstilbud, hva disse brukes til og deretter å vurdere reliabilitet og validitet i prøvene.

Undersøkelsen min var todelt. Først sendte jeg ut et kort, nettbasert spørreskjema til alle 25 skoler i tidligere Hordaland fylke som på sine nettsider informerer om at de har innføringstilbud, og jeg fikk inn svar fra lærere på 17 ulike skoler. Deretter ble det gjennomført en systematisk analyse av 14 kartleggingsprøver ved hjelp av *EFPA Test Review Model*. Denne undersøkelsen av kvaliteten av instrumentene ble gjort ved å bruke internasjonale standarder for å undersøke kvaliteten i tester som brukes til utdanningsformål.

Resultatene viser at skoler kartlegger ofte og bruker ofte flere ulike kartleggingsverktøy for å kartlegge elever, og at prøvene brukes til ulike formål, som å tilpasse undervisningen, nivådele grupper og avdekke vansker på første- eller andrespråket. Den viser også at noen av prøvene brukes for å vurdere om elever har rett til særskilt språkopplæring (jf. §2-8 og §3-12 i Opplæringslova), noe som gjør at prøven anses som ‘high stake’ fordi den kan ha store konsekvenser for elevens videre opplæring. I drøftingen problematiseres det også at begrepet ‘tilstrekkelig norskferdigheter’ ikke er avklart i lov eller forskrift, og at det er uklart hva som egentlig avgjør om en elev har eller ikke har rett til særskilt språkopplæring.

Litteraturgjennomgangen viser at de fleste kartleggingsprøvene som brukes i liten eller ingen grad vurderes i forhold til reliabilitet og validitet.

Funnene i denne studien viser at det er et stort behov for at kartleggingsmateriell som brukes i skolen for å kartlegge minoritetsspråkliges språkferdigheter, må vurderes i forhold til reliabilitet og validitet for å sikre at elever vurderes etter prøver som måler det de sier de skal måle og som gir resultater man kan stole på. I tillegg må begrepet ‘tilstrekkelige norskferdigheter’ avklares i lov eller forskrift, slik at alle har den samme forståelsen av når en elev skal ha opplæring etter ordinær læreplan.

Abstract

Minority language-speaking pupils in primary and secondary education are entitled to special language instruction, if they do not have ‘sufficient skills in Norwegian’, cf. §2-8 and §3-12 in the Education Act (Opplæringslova). The focus of this master’s thesis has been to look at which language mapping tests the schools in previously Hordaland county use to map Norwegian skills in introductory courses, what these mapping tests are used for and then to assess reliability and validity in the language mapping tests.

My survey was twofold. First, I sent out a short, online questionnaire to all 25 schools in former Hordaland county, which on their website informs that they have introductory courses in Norwegian, and I received answers from teachers at 17 different schools. Subsequently, a systematic review of 14 language mapping tests was performed using the *EFPA Test Review Model*. This study of the quality of the instruments was done by using international standards to examine the quality of tests used for educational purposes.

The results show that schools often map pupils, that they often use several different language mapping tests, and that the tests are used for different purposes, such as adapting the individual teaching, dividing into groups and to find difficulties in their first or second language. It also shows that some of the tests are used to assess whether pupils are entitled to special language training (cf. §2-8 and §3-12 in the Education Act), which means that the test is considered to be ‘high stake’ because it may have big consequences for the pupil’s education. In the discussion, it is also problematized that the concept of ‘sufficient Norwegian skill’ has not been clarified by law or regulation, and that it is unclear what determines whether a pupil has or has not the right to special language instruction. This literature review shows that most mapping tests that are used to a small or no degree are assessed in relation to reliability and validity.

The findings in this study show that there is a great need for assessing reliability and validity in mapping tests used in schools to map language skills of minority language-speakers. This is to ensure that pupils are assessed on tests that measure what they say they measure and gives results one can trust. In addition, the concept of ‘sufficient Norwegian skills’ must be clarified in law or regulations, so that everyone has the same understanding of when a pupil should have his or her education according to the ordinary curriculum.

1. Innledning

Opplæringslova §2-8 og §3-12 pålegger kommunene og fylkeskommunene å kartlegge alle minoritetsspråklige elever for å avgjøre om de har 'tilstrekkelige ferdigheter i norsk' til å følge ordinær læreplan. Dersom de ikke har det, utløser det rett til særskilt språkopplæring, herunder særskilt norskopplæring, morsmålsundervisning og tospråklig fagopplæring. I praksis tilfaller dette ansvaret på skolene. Mitt fokus vil her være på retten til særskilt norskopplæring, siden det hovedsakelig er det som gis i innføringsklasser i grunnskolen og den videregående skolen. Det er også viktig å ha med seg at et formelt kriterie for å plassere elever i innføringsklasser er vedtak om særskilt språkopplæring jf. §2-8 eller §3-12 i Opplæringslova.

1.1. Formål og hensikt

Det er ingen føringer for *hvordan* kommunene, eller skolene, skal kartlegge elevene, heller ikke *hvilke* kartleggingsmaterieell som skal brukes. I dag er det mange ulike kartleggingsprøver å velge mellom, og skolene kan også velge å lage sine egne kartleggingsprøver. Det er lite informasjon om kvaliteten på de kartleggingsprøvene som brukes, og dette er et felt som det er forsket veldig lite på. Et unntak er Mia C. Heller som undersøkte kartleggingsverktøyet *Norsk som læringsspråk (NSL)* som en del av sitt doktorgradsprosjekt, der hun vurdere de ulike komponentene i NSL ved å analysere reliabilitet, begrepsvaliditet og kriterievaliditet (Heller & Lervåg, 2020). En annen relatert studie, Arnesen, Braeken, Ogden og Melby-Lervåg (2019), undersøkte kvaliteten på kartleggingsverktøy som brukes for å vurdere sosial fungering og leseferdigheter i grunnskolen, der resultatene viser at mange av verktøyene har svak validitet og reliabilitet, og noen av verktøyene manglet slike vurderinger i det hele tatt. Disse resultatene viser at dette er et felt det er viktig å forske mer på, slik at skoler i større grad kan gjøre gode vurderinger av verktøyene før de velger hvilke de ønsker å bruke. Alle lærere ønsker å bruke gode verktøy når de kartlegger, men ofte er det vanskelig å vite hva som er et godt verktøy å bruke til ulike formål.

Dette er utgangspunktet for temaet for masterprosjektet mitt: Språkkartlegging av minoritetsspråklige elever i innføringstilbud i tidligere Hordaland fylke. Jeg har gjennom en spørreundersøkelse fått vite hvilke kartleggingsverktøy skolene bruker til å vurdere norskferdigheter i innføringsklasser, og deretter har jeg sett på validitet og reliabilitet i disse kartleggingsverktøyene. Utvalget av skoler som er med i prosjektet, er både barneskoler, ungdomsskoler og videregående skoler i hele det tidligere Hordaland fylke. Undersøkelsen

var helt frivillig, og lærere valgte selv om de vil svare på spørreundersøkelsen, som også er helt anonym. Ingen svar kan spores tilbake til den enkelte skole eller lærer.

1.2. Problemstilling

Hvilke kartleggingsprøver brukes for å kartlegge norskspråklige ferdigheter i innføringstilbud i tidligere Hordaland fylke, og hvilken validitet og reliabilitet har måleinstrumentene som brukes?

Forskningsspørsmål:

1. Hvilke kartleggingsprøver brukes for å kartlegge norskspråklige ferdigheter i innføringstilbud i tidligere Hordaland fylke?
2. Hvor ofte og til hvilket formål brukes kartleggingsprøvene?
3. Hvilken validitet og reliabilitet har måleinstrumentene? ‘

1.3. Begrepsavklaring

1.3.1. Test eller prøve?

I litteratur om språktesting og språk-kartlegging, finner vi to begreper: test og prøve. I min oppgave skiller jeg ikke mellom disse to begrepene, og det handler om at det i engelsk litteratur handler om ‘tests’ og på norsk brukes både ‘test’ og ‘prøve’. Når vi snakker om kartlegging, vil det ofte være brukt begrepet ‘prøve’, men samtidig så handler testlitteraturen om mange former for prøver og tester.

Det finnes flere ulike prøveformater, med ulike bruksområder og kriterier. *Standardiserte tester* er tester som ikke endrer seg. Prøvekonstruktet, formatet og antall oppgaver er det samme, og prøven gjennomføres på samme måte hver gang. Det er også klare regler for hva som er bestått eller ikke, og prøven vil bli skåret på samme måte hver gang. I tillegg er prøvene grundig prøvd ut på forhånd gjennom pilotstudier, og resultatene fra disse studiene analyseres før prøvene er klar til gjennomføring. Vanligvis er standardiserte prøver offisielle prøver, og noen ganger får man et prøvebevis, men ikke alltid. Eksempel på dette kan være Utdanningsdirektoratets kartleggingsprøver for grunnskolen. Disse gir ikke prøvebevis fordi det er prøver som tas underveis i opplæringen, mens prøvebevis gis når den markerer noe som er bestått.

Ferdighetsprøver er laget for å måle språkferdigheter. Disse prøvene har ikke pensum, men måler ferdigheter i f.eks. språk ut fra ferdighetsnivåene i *Det felles europeiske rammeverket*

for språk (også kjent som *Rammeverket* eller CEFR¹), eller andre ferdigheter en deltaker må bestå et nivå i (Carlsen & Moe, 2019). *Oppnåelsesprøver* skal finne ut om eleven har lært det som står i kompetansemålene eller på andre måter ligger innenfor pensum (Carlsen & Moe, 2019), mens *diagnostiske prøver* skal gi informasjon om elever som trenger ekstra hjelp (Carlsen & Moe, 2019). Eksempler på slike prøver kan være prøver for å se om elever har lese- og skrivevansker, eller en grammatikkprøve i språkfag for å se om elevene mestrer grammatiske strukturer. *Plasseringsprøver* er prøver som måler elevenes ferdighetsnivå for å plassere dem i grupper basert på ferdighetsnivået fordi man tenker at elever som f.eks. kan lite språk vil kunne oppleve mer læring og progresjon sammen med andre som kan lite språk fordi undervisningen blir lagt på et nivå som passer nivået til gruppen (Carlsen & Moe, 2019). På den måten kan man argumentere for at plasseringsprøver er pedagogisk motiverte prøver fordi målet er best mulig forutsetning for læring. *Fremgangsprøver* skal måle fremgang, og blir vanligvis gitt på ulike tidspunkt i løpet av et læringsløp (Carlsen & Moe, 2019). Det finnes også *egnethetsprøver*, men dette er ikke så vanlige i språktesting, og disse måler talenter for noe, f.eks. å lære et nytt språk (Carlsen & Moe, 2019).

1.3.2. Elevgruppen

Minoritetsspråklige elever er en sammensatt gruppe. Ifølge Arbeids- og inkluderingsdepartementet bør man bruke begrepet 'flerspråklige', men Jortveit (2017) hevder man bør holde seg til begrepet 'minoritetsspråklig' fordi majoritets elever også kan være flerspråklige. Flere institusjoner bruker også begrepene 'minoritetsspråklige' eller 'språklige minoriteter':

Begrepet språklig minoritet brukes vanligvis om alle i Norge som ikke har norsk eller samisk som morsmål. Når begrepet brukes i tiltaksplanen, refererer det til elever eller deltakere som ikke har norsk eller samisk som morsmål, og voksne som ikke har norsk, samisk, dansk eller svensk som morsmål, og som har behov for særskilt tilrettelagt opplæring.

(Kunnskapsdepartementet, 2007, s. 9)

¹ *Common European Framework of Reference for Languages*, som på norsk oversettes til *Det felles europeiske rammeverket for språk*.

Med minoritetsspråklige elever i grunnopplæringen forstår vi barn, unge og voksne som har et annet morsmål enn norsk eller samisk. I denne sammenhengen brukes begrepet om elever i 1.–10. trinn i grunnskolen og videregående opplæring, samt voksne deltagere innenfor grunnopplæringens område.

(Utdanningsdirektoratet, 2016)

Her ser vi at både Kunnskapsdepartementet og Utdanningsdirektoratet bruker begrepene ‘minoritetsspråklige’ eller ‘språklige minoriteter’ når de snakker om denne elevgruppen, også uavhengig av hvor de er i utdanningsløpet.

Når disse elevene kommer til Norge, så møter de ulike sider av utdanningssystemet vårt. Noen blir plassert direkte inn i ordinære klasser, noen plasseres i såkalte innføringstilbud, voksne plasseres i voksenopplæringen, enten grunnskole for voksne, norskkurs eller videregående opplæring for voksne, og noen søker seg direkte til høyere utdanning. Hva som møter de ulike elevene i de ulike tilbudene, er nok veldig ulikt. Mitt fokus for denne oppgaven er innføringstilbudene i grunnskolen og videregående skoler i tidligere Hordaland fylke, som også forutsetter at elevene har fått enkeltvedtak som gir dem rett til særskilt språkopplæring, jf. §2-8 eller §3-12 i Opplæringslova.

Et innføringstilbud er en måte å organisere særskilt norskopplæring på (Rambøll Management, 2016). Det gir skolen rom for å avvike fra den ordinære læreplanen og gi et tilpasset opplæringsløp med fokus på norskopplæring i alle fag. Opplæringen skjer i stor grad på norsk, altså på andrespråket til elevene (Egeberg, 2016). Meld.St.6 peker på viktigheten av at de som driver med særskilt språkopplæring har kunnskap om «språkstimulering, kartlegging av språkferdigheter, organisering av innhold i ulike typer tilbud og kjennskap til god praksis» (Det kongelige barne-, likestillings- og inkluderingsdepartement, 2012-2013, s. 51). I Rambøll Managements *Evaluering av særskilt språkopplæring og innføringstilbud* (2016) konkluderer de med at organiseringen av innføringstilbud er god og at elever ser ut til å ha nytte av en slik organisering, men de peker også på at det ser ut til at en del elever overføres til ordinære klasser for tidlig, uten at de har ‘tilstrekkelige norskferdigheter’. Evalueringen sier ingenting om disse elevene får tilbud om særskilt språkopplæring i, eller som del av, det ordinære løpet.

1.3.3. Hva er ‘tilstrekkelige ferdigheter i norsk’?

I Opplæringsloven §2-8 og §3-12 står det at eleven har rett til særskilt språkopplæring frem til eleven har ‘tilstrekkelige ferdigheter i norsk’ til å følge ordinær læreplan. Det diskuteres ofte

hva som ligger i begrepet 'tilstrekkelig norskferdigheter' for å følge opplæring etter ordinær læreplan. Ifølge Kunnskapsdepartementet er det ikke «fastsatt kriterier i § 2-8 om hva som er "tilstrekkelige kunnskaper i norsk". Det må derfor foretas en skjønnsmessig, lærerbasert vurdering av når eleven har oppnådd tilstrekkelige kunnskaper» (Kunnskapsdepartementet, 2007, s. 16). Likevel kan vi finne noe som kan fortelle oss hva 'tilstrekkelige kunnskaper' er hvis vi ser i læreplanen for grunnleggende norsk: «Når elevene har nådd målene i læreplanens nivå 3, skal de følge ordinær norskopplæring og bli vurdert i samsvar med denne» (Utdanningsdirektoratet, 2007b). I veiledningen til *Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk* forklarer de tolkningen av begrepet 'tilstrekkelige ferdigheter i norsk' som nivået når en elev skal følge ordinær læreplan. Dette begrepet er sentralt i opplæringsloven, som nevnt tidligere. «Nivå 3 i læreplanen tilsvarer nivå B1 og deler av nivå B2 i rammeverket. Opplæring ut over nivå 3 forutsettes å skulle foregå på grunnlag av den ordinære norskplanen» (Utdanningsdirektoratet, 2009, s. 10). Veiledningen til kartleggingsmateriellet trekker også frem bestemmelser gjort for *Læreplan i grunnleggende norsk for språklige minoriteter* der det står at «Det er utarbeidet kartleggingsverktøy i tilknytning til læreplanen. Når elevene har nådd målene i læreplanens nivå 3, skal de følge ordinær norskopplæring og bli vurdert i samsvar med denne» (Utdanningsdirektoratet, 2009, s. 56). Denne formuleringen finnes ikke i den nye læreplanen som innføres fra skoleåret 2020/2021 (Utdanningsdirektoratet, 2020). Ser man disse to i sammenheng, vil det på den ene siden være tydelige kriterier for hva som tilsvarer 'tilstrekkelige norskferdigheter', og på den andre siden at dette er noe som skal vurderes skjønnsmessig av lærerne. Her er altså ikke Kunnskapsdepartementet og Utdanningsdirektoratet enige. Evalueringen til Rambøll Management (2006) konkluderer også med at 'tilstrekkelige norskferdigheter' er et begrep som ikke er avklart. Randen (2013) avla sin doktorgradsavhandling, *Tilstrekkelige ferdigheter i norsk?* i 2013 der hun så på minoritetsspråklige førsteklasingers språkferdigheter. I avslutningen trekker hun frem at det brukes store ressurser på begrepet 'tilstrekkelige ferdigheter i norsk' uten at det finnes begrepsavklaringer eller andre føringer på hvordan dette begrepet skal tolkes i henhold til Opplæringslovas §2-8 og §3-12 (Randen, 2013). Engen (2010, s. 30) ser ut til å ha forståelsen av at 'tilstrekkelige norskferdigheter' tilsvarer læreplanens nivå 3 når han sier at «Kriteriet på for tidlig overføring i denne sammenheng, er det nivået (3) i planen for grunnleggende norsk for språklige minoriteter, som er definert som tilstrekkelig for at overgang skal finne sted». Det blir tydelig at begrepet er uavklart og diffust, og at lærere skal vurdere dette individuelt, noe som kan føre til veldig ulik praksis mellom og på tvers av lærere, skoler og kommuner.

Engen (2010) trekker også frem at det ser ut til å råde en ideologi om at minoritetsspråklige elever skal overføres til 'ordinær opplæring' så snart som mulig, ofte sosialt begrunnet, og at dette fører til at mange minoritetsspråklige elever ender opp med individuelle tiltak seinere for å bøte på manglende forståelse i det ordinære opplæringsløpet. Han begrunner dette med at faget «Norsk som andrespråk» er fjernet fra skolen og viser til Rambøll Managements rapport, *Evaluering av implementering av læreplaner i grunnleggende norsk og morsmål for språklige minoriteter*, fra 2009. Denne rapporten sier ingenting om at elever skal overføres til ordinær undervisning raskt, men den trekker frem at beslutningen om *når* eleven skal overføres til ordinær undervisning i grunnskolen ofte skjer i et samarbeid mellom lærer og skoleledelse, mens i videregående skole blir denne oppgaven oftere lagt på den enkelte læreren (Rambøll Management, 2009). *Evaluering av særskilt språkopplæring og innføringstilbud* (Rambøll Management, 2016) trekker derimot frem at skoleledere avgrensner innføringstilbud til ett år, som de begrunner med at det er viktig at elevene kommer tilbake til ordinær undervisning. Denne rapporten kommer likevel seks år etter Engens kommentar.

1.3.4. Læreren som portvokter?

Utsagnet fra Kunnskapsdepartementet om at det er læreren som skal vurdere begrepet 'tilstrekkelige norskerferdigheter' står i sterk kontrast til studien til Antoniazzi, Snow og Dickson-Swift (2010), som undersøker i hvilken grad lærere klarer å vurdere om elever har språkvansker i de første skoleårene, og metastudien til Südkamp, Kaiser og Möller (2012), som ser på samsvar mellom lærers vurderinger og elevers akademiske ferdigheter.

Antoniazzi, Snow og Dickson-Swift (2019) konkluderer med at lærere i liten grad klarer å vurdere hvilke elever som har språkvansker, både ved at de i liten grad klarer å identifisere de som har språkvansker og at de finner språkvansker hos elever som ikke har dette. Ifølge studien ble 21-50 % av elevene med språkvansker ikke identifisert av lærerne. 59 % av elevene lærerne mente ikke hadde språkvansker, ble gjennom screeningtesten identifisert som elever med mulige språkvansker. Südkamp, Kaiser og Möller (2012) konkluderer med at lærere i noen grad klarer å vurdere elevers akademiske ferdigheter, og at disse vurderingene er uavhengige av fag. Det betyr at de ikke finner forskjeller i lærers vurderinger f.eks. av elevers matematikkferdigheter og språkerferdigheter. Selv om de peker på at lærere i noen grad klarer å vurdere elevers ferdigheter, så trekker de frem at mer enn to tredjedeler av variansen i lærervurderingene ikke kan forklares av elevenes ferdigheter. Südkamp, Kaiser og Möller (2012) hevder også at lærere i stor grad overvurderer elevenes ferdigheter, og i særlig grad gjelder dette for elever på eller under aldersadekvat nivå i matematikk og lesing.

2. Teori

2.1. Krav om kartlegging

For elever i grunnskolen og den videregående skolen er det krav om at alle minoritetsspråklige elever skal kartlegges for å vurdere hvilken læreplan de skal følge. I tillegg kan kartleggingen vise at de har rett på særskilt språkopplæring, jf. §2-8 (grunnskolen) og §3-12 (videregående skole) i Opplæringslova. Disse paragrafene er like, men de omtaler henholdsvis grunnskolen og den videregående skolen, så noen av begrepene endres derfor.

Elevar i grunnskolen med anna morsmål enn norsk og samisk har rett til særskild norskopplæring til dei har tilstrekkeleg dugleik i norsk til å følgje den vanlege opplæringa i skolen [...] Kommunen skal kartleggje kva dugleik elevane har i norsk, før det vert gjort vedtak om særskild norskopplæring. Slik kartlegging skal også utførast undervegs i opplæringa for elevar som får særskild språkopplæring etter føresegna, som grunnlag for å vurdere om elevane har tilstrekkeleg dugleik i norsk til å følgje den vanlege opplæringa i skolen.

(Opplæringslova, u.d., §2-8)

Dette betyr at skolen er pliktig å kartlegge de minoritetsspråklige elevene ved skolestart, men også flere ganger i løpet av året. Resultatet av kartleggingen vil være grunnlaget for et enkeltvedtak som avgjør hvilken læreplan elevene skal følge og hvilke rettigheter de har i skoleløpet. Når eleven får vedtak om å følge ordinær læreplan, vil normalt de andre rettighetene falle bort. Det betyr at kartleggingen kan ha store konsekvenser for den enkelte eleven. Til tross for dette, er det opp til kommunene og skolene å velge *hvordan* de utfører kartleggingen, og *hvilke(t)* kartleggingsmaterieill de ønsker å bruke. De står dermed helt fritt til å velge hva de vil bruke av det som er tilgjengelig, f.eks. veletablerte prøver eller egenproduserte prøver.

2.2. Prøvene

I Norge har det vært språkprøver som har vært basert på ferdighetsnivåene i *Det felles europeiske rammeverket for språk* (heretter *Rammeverket*) siden 2005 (Carlsen & Moe, 2019). Ferdighetene er knyttet til hvilke ferdigheter man har i et språk innenfor områdene lytte, lese, skrive og snakke. Innenfor disse kategoriene avgjøres det så hvilket språknivå man forventer ulike ferdigheter innenfor i disse kategoriene. Figur 1 viser et eksempel på hvordan ferdighetsbeskrivelsene er for lytte- og leseforståelse på de laveste tre nivåene.

		A1	A2	B1
F O R S T Å E L S E	L Y T T I N G	Jeg kan kjenne igjen vanlige ord og svært enkle uttrykk som har med meg selv, min familie og mine nære omgivelser å gjøre, når man snakker langsomt og tydelig.	Jeg kan forstå vanlige ord og ytringer knyttet til områder som er viktige for meg personlig (f.eks. svært enkel informasjon som gjelder meg selv, min familie, nærmiljø, innkjøp og arbeidsliv). Jeg kan oppfatte hovedinnholdet i enkle, korte og tydelige beskjeder og meldinger.	Jeg kan forstå hovedinnholdet i klar, standardtale om kjente emner knyttet til arbeid, skole og fritid osv. Jeg kan forstå hovedinnholdet i mange radio- og TV-program om aktuelle saker eller om emner som er av personlig eller faglig interesse for meg, når man snakker relativt langsomt og tydelig.
	L E S I N G	Jeg kan oppfatte kjente navn og forstå ord og svært enkle setninger, for eksempel i annonser, på plakater eller i kataloger.	Jeg kan lese svært korte, enkle tekster. Jeg kan finne konkrete, forutsigbare opplysninger i enkel, dagligdags informasjon som reklame, brosjyrer, menyer og rutetabeller. Jeg kan forstå korte, enkle, personlige brev.	Jeg kan forstå tekster der språket i stor grad er preget av frekvente, dagligdags ord og uttrykk eller er knyttet til arbeidet mitt. Jeg kan forstå personlige brev med beskrivelser av hendelser, følelser og ønsker.

Figur 1. Utdrag fra nivåbeskrivelser i Rammeverket (Utdanningsdirektoratet, 2011, s. 32)

Norsk språktest, et samarbeid mellom Universitetet i Bergen og Folkeuniversitetet, utviklet prøvene *Norskprøve 2 for voksne* og *Norskprøve 3 for voksne* (heretter *Norskprøve 2 og 3*), som måler henholdsvis språknivå A2 og B1. Kompetanse Norge har fra 2014 utviklet *Norskprøven for voksne* (ny, digital versjon, heretter *Norskprøven*), som måler språknivå A1-B2. Fagbokforlagets digitale kartleggingsprøve, *Migranorsk*, og *Den europeiske språkpermen*, som er laget på oppdrag fra Utdanningsdirektoratet, måler språkkompetanse ut fra de ferdighetene vi finner i *Rammeverket*: lytting, lesing, muntlig kommunikasjon og tekstproduksjon. *Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk* fra Utdanningsdirektoratet, måler ferdigheter definert etter nivå 1, 2 og 3, og disse sammenfaller i stor grad med nivå A1, A2 og B1-B2 i *Rammeverket*.

Vestlandsverktøyet er et relativt nytt kartleggingsverktøy, og her finnes bare muntlig og skriftlig prøve. Instruksjonene spesifiserer at eleven skal svare på alle spørsmål, på alle nivåer (A1-B1+), uavhengig av elevens antatte nivå eller forutsetning. I denne prøven måles eleven ut fra nivåbeskrivelser som Carlsen og Moe (2019) peker på er hentet fra *Rammeverket* og tilpasset en yngre brukergruppe.

I tillegg finnes det prøver som kartlegger ferdigheter uten at disse kobles til språknivåene i *Rammeverket*, men gjerne til norskfaglige ferdigheter man tenker elevene trenger på ulike trinn i skolen, som *Trondheimstesten*. Testene *TOSP* og *FLORO* måler kognitive forutsetninger for læring, både på norsk og på morsmål.

2.2.1. Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk

Kartleggingsmaterieellet er bygd opp av tre deler, en språkbiografi, kartleggingsverktøyet og en språkmappe. Kartleggingsmaterieellet har vært mye omtalt, og også kritisert. Det er 97 'eleven kan'-utsagn. Disse skal vurderes etter om eleven kan gjøre dette 'med mye hjelp', 'med litt hjelp' eller 'uten hjelp'. Carlsen og Moe (2019) peker på at det er lite samsvar mellom nivåbeskrivelsene og de 97 'eleven kan'-utsagnene, og det gjør det veldig krevende for lærere å bruke verktøyet. I tillegg er det flere punkter som krever tolkning, og uten en felles forståelse for hvordan det skal tolkes og vurderes, så vil elever bli vurdert veldig ulikt. I rapporten fra Rambøll Management kommer det frem at kartleggingsmaterieellet inneholder «for generelle ferdighetsmål, som krever tolkninger og skjønnsmessige vurderinger fra lærerne. Lærere vurderer måloppnåelse forskjellig, og følgelig blir lærernes kompetanse avgjørende» (Carlsen & Moe, 2019, s. 61). Når man også ser at det er dette som ligger til grunn for enkeltvedtak, med føringer for hvilken læreplan eleven skal bruke, og om eleven har rett på særskilt språkopplæring, morsmålsopplæring eller tospråklig fagopplæring, så er det viktig å evaluere bruken av verktøyet. Noe av det som er positivt med denne kartleggingsprøven, er at den flerspråklig kompetanse til elevene kommer godt frem og at elevene selv kan være med å vurdere sine egne ferdigheter (Carlsen & Moe, 2019).

2.2.2. Den europeiske språkpermen

Språkpermen er bygd opp på samme måte som kartleggingsprøven til Utdanningsdirektoratet. Hovedforskjellen er at språkpermen er laget for å rette seg mot deltakeren eller eleven som bruker den. Det er eleven selv som fyller ut hva han/hun mestrer eller ikke, og på hvilket nivå. I veiledningsmaterieellet spesifiseres det at dette bør gjøres i samråd med lærer, og de anbefaler også at ikke eleven får hele språkpermen med en gang, men at lærer leverer ut ark fra den som eleven kan fylle i etter hvert i undervisningen (Simonsen & Speitz, 2006).

2.2.3. Vestlandsverktøyet – Vurderingsverktøy for minoritetsspråklige elever i VGS

Vestlandsverktøyet er et nytt kartleggingsverktøy som er laget i et samarbeid mellom Hordaland Fylkeskommune og Høgskulen på Vestlandet. Våren 2020 inneholder verktøyet en skriftlig og en muntlig prøve, samt vurderingskriterier, veiledning og eksempelbesvarelser med nivåbeskrivelser. Ifølge Carlsen og Moe (2019) skal det arbeides med å lage en lesetest skoleåret 2019/2020, men denne er ikke klar pr. april 2021.

Vestlandsverktøyet er ikke en standardisert prøve og prøven er tenkt som et verktøy for å sikre at de som har rett til særskilt språkopplæring, får dette (Carlsen, 2021, e-postutveksling). Tanken er altså at dette skal være et pedagogisk verktøy, både som del i å fatte enkeltvedtak,

men også underveis i undervisningen. Hun forklarer videre at de elevene som bruker disse prøvene for å komme inn på særskilt tilrettelagte tilbud for minoritetsspråklige på VG1 i Vestland fylkeskommune, blir vurdert av de som har vært med å utvikle prøvene, og at dette kvalitetssikres gjennom erfarne sensorer ved voksenopplæringen på Nygård skole. Carlsen (2021, e-postutveksling) trekker også frem at det vurderes å gjennomføre en reliabilitetsanalyse av sensorvurderingene, men dette er ikke gjennomført når dette prosjektet leveres.

Vestlandsverktøyet bygger på ferdighetsbeskrivelsene i *Rammeverket*, og Carlsen og Moe (2019) peker på at de opprinnelige nivåbeskrivelene er beholdt: nivå A1, A2, B1 og B1+ (tilsvarende nivå 3 i *Læreplan for grunnleggende norsk for språklige minoriteter*). Nivåbeskrivelsene er i utgangspunktet hentet fra *Rammeverket*, men er tilpasset en yngre brukergruppe (Carlsen & Moe, 2019).

2.2.4. FLORO

FLORO består av tre kognitive prøver som måler ordminne, RAN og ordrepetisjon. Ordminneprøven er laget etter et samarbeid med Charles Hulmes i 1997 og ble den gang pilotert på norsk, tyrkisk, albansk og vietnamesisk på begynnertrinnet i barneskolen (Utdanningsdirektoratet & NAFO, 2013b). RAN (Rapid Automated Naming) er omarbeidet etter en mal fra Stefan Samuelson ved Lindköpings universitet og Lesesenteret i Stavanger. Ordrepetisjon er en prøve som tester fonologiske ferdigheter, og også denne er prøvd ut av Stefan Samuelson og teamet hans.

Prøvene skal bidra til å finne elever med språk-, lese- eller skrivevansker. Språk og lærevansker er kompliserte og sammensatte fenomen og det er derfor viktig å få belyst ulike sider ved vanskene og dermed få et helhetlig bilde (Utdanningsdirektoratet & NAFO, 2013b).

2.2.5. TOSP

TOSP er ment som et hjelpemiddel for å vurdere elevens ferdigheter på tre områder innen språklig og kognitiv utvikling: Begrepsforståelse, ordforråd og leseferdighet. Ferdigheter på alle disse områdene har en innvirkning på språk- og fagopplæring (Loona, 2008a). *TOSP* er også en tospråklig prøve, som gjør at elever kan testes også på morsmål dersom man ønsker å se om eleven har de samme vanskene, eller utfordringene, på morsmålet som eleven viser på norsk. 'Vansker' i denne sammenhengen er ikke nødvendigvis språk- eller lærevansker, men at eleven har svake andrespråksferdigheter og at man bruker kartleggingsprøver for å finne ut mer om hvorfor eleven strever med dette.

2.2.6. Norskprøve 2 og 3 for voksne innvandrere og Norskprøven for voksne innvandrere
Norskprøve 2 og 3 for voksne innvandrere (Norskprøve 2 og 3) er en tidligere versjon av *Norskprøven for voksne innvandrere (Norskprøven)*. Både *Norskprøve 2 og 3* og *Norskprøven* er bygd opp med fire deltester: lytte, lese, skrive, snakke. *Norskprøve 2 og 3* er en papirbasert prøve, unntatt den muntlige delprøven. Lyttedelen kommer med CD som spilles av i rommet for deltakerne mens de svarer på papir. Lesedel og skrivedel følger etterpå. Deltakerne må velge selv hvilket nivå de skal testes i i denne versjonen. *Norskprøve 2* måler A2, mens *Norskprøve 3* måler B1.

Norskprøven er digital, med unntak av den muntlige delprøven, og lytte- og leseprøven er delvis adaptiv. Det betyr at prøven, gjennom innledede oppgaver, ser hvilket nivå deltakeren er på, og velger deretter hvilken prøve som blir gitt deltakeren. Lytte- og leseprøven rettes automatisk av datamaskinen. For den skriftlige prøven må deltakerne selv velge hvilket nivå de testes på: A1-A2, A2-B1 eller B1-B2. Disse prøvene rettes av to sensorer som fører sin vurdering inn i datasystemet. Dersom begge registrerer det samme nivået, vil deltakeren automatisk få tilsendt resultatene, men dersom de to sensorene er uenige vil prøven sendes til en tredje sensor (Carlsen & Moe 2019). Nivået som da blir registrert på deltakeren vil bli det som to av de tre sensorene er enige om.

I 2018 ble *Norskprøven* tildelt kvalitetsstempelet Q-mark fra ALTE, som betyr at prøven er vurdert av Association of Language Testers in Europe (ALTE) til å tilfredsstillere minstekravene til en god språkprøve.

2.2.7. Migranorsk

Migranorsk er en digital, adaptiv språkprøve som måler lytting, lesing, grammatikk og skriving. De tre første delene rettes av datasystemet, mens skrivdelen må rettes av læreren selv. En svakhet ved prøven er dersom eleven får en skriveoppgave knyttet til bilde. Lærer får ikke se hvilket bilde eleven har fått i oppgave å beskrive, som gjør det vanskelig å vurdere skriveoppgaver på A1. På øvrige nivåer vil oppgaveteksten følge med elevens besvarelse. Prøven er adaptiv i den forstand at eleven må svare rett på ca. 60% av oppgavene på A1 for å få oppgaver på A2, osv. Fagbokforlaget hevder at *Migranorsk* tester er utviklet av erfarne tetsutviklere og testleddskrivere, men de vil ikke gi ut informasjon om validitet og reliabilitet når dette blir etterspurt i en e-post. De forklarer at de ikke kan gi ut slik informasjon pga. kommersielle hensyn.

2.2.8. Trondheimstesten

Trondheimstesten ble laget for å synliggjøre fremmedspråklige elevers norskkompetanse i overgangen fra mottaksklasse til ordinær klasse. Den inneholder de mest grunnleggende temaene med ord og begreper man må kunne på norsk, og er ikke en normert språkprøve. Denne prøven kan gi en *indikasjon* på om eleven har rett på særskilt språkopplæring, men den er bare del av en helhetsvurdering når elever får tilbud som særskilt språkopplæring etter Opplæringslova §2-8 (Amundsen, 2021, e-postutveksling).

2.3. utfordringer med kartleggingsprøvene

Kjelsaas, Eide, Vidarsdottir og Hauge (2020) har gjennom en kvalitativ studie undersøkt språkkartlegging ved fem skoler i to fylker. De har 31 forskningsdeltakere der 22 er elever og 9 ansatte i ulike funksjoner på skolen og hos pedagogisk-psykologisk tjeneste (PPT). Det er uklart om disse forskningsdeltakerne kommer fra alle fem skolene eller bare den ene. De trakk frem eksempler på elever ved én videregående skole som har nektet å la seg kartlegge fordi de opplevde at denne kartleggingen hadde svært negative konsekvenser for dem i skoleløpet og var redd for at nye kartlegginger ville gi flere negative konsekvenser for dem. «'De syns bare vi er dumme', sa flere, og en annen beskrev hvordan kartlegginga fikk henne til å føle seg som 'en idiot'» (Kjelsaas, et al., 2020, s. 2). I dette eksemplet har elevene etter kartlegging i VG1 blitt plassert i toårig VG1-løp basert på kartleggingsresultatene, og disse reaksjonene kommer når elevene skal starte sitt andre år i dette VG1-løpet. Kjelsaas, et al. (2020) trekker videre fram at både verktøyene og praksisen man bruker for å kartlegge elever ikke er nøytrale, men er ideologiske fenomener (Kjelsaas, et al., 2020). Når man har sett på konsekvenser av kartlegging, så handler det om hvilke konsekvenser det får for opplæring og i samfunnet, ikke hvilke konsekvenser det har, eller oppleves å ha, for elevene (Kjelsaas, et al., 2020). I undersøkelsen viste de til at elevene og de ansatte ved skolen sitter på to ulike oppfatninger av kartleggingen. Elevene opplevde kartleggingen som et mindreverdighetsstempel, mens de ansatte hevdet elevene ikke forsto at kartleggingen skulle gi elevene best mulige forutsetninger i den videre opplæringen (Kjelsaas, et al., 2020). Og dette er kanskje noe av det viktigste arbeidet i møte med minoritetsspråklige elever, å få dem til å forstå at særskilt språkopplæring er en rett og et gode, ikke noe som ødelegger for dem. Det er likevel viktig å ha med seg at denne retten også kan føre til en følelse av ekskludering fra fellesskapet dersom elever tas ut av ordinære klasser hele eller deler av tiden uten at eleven selv ønsker og ser behovet av dette.

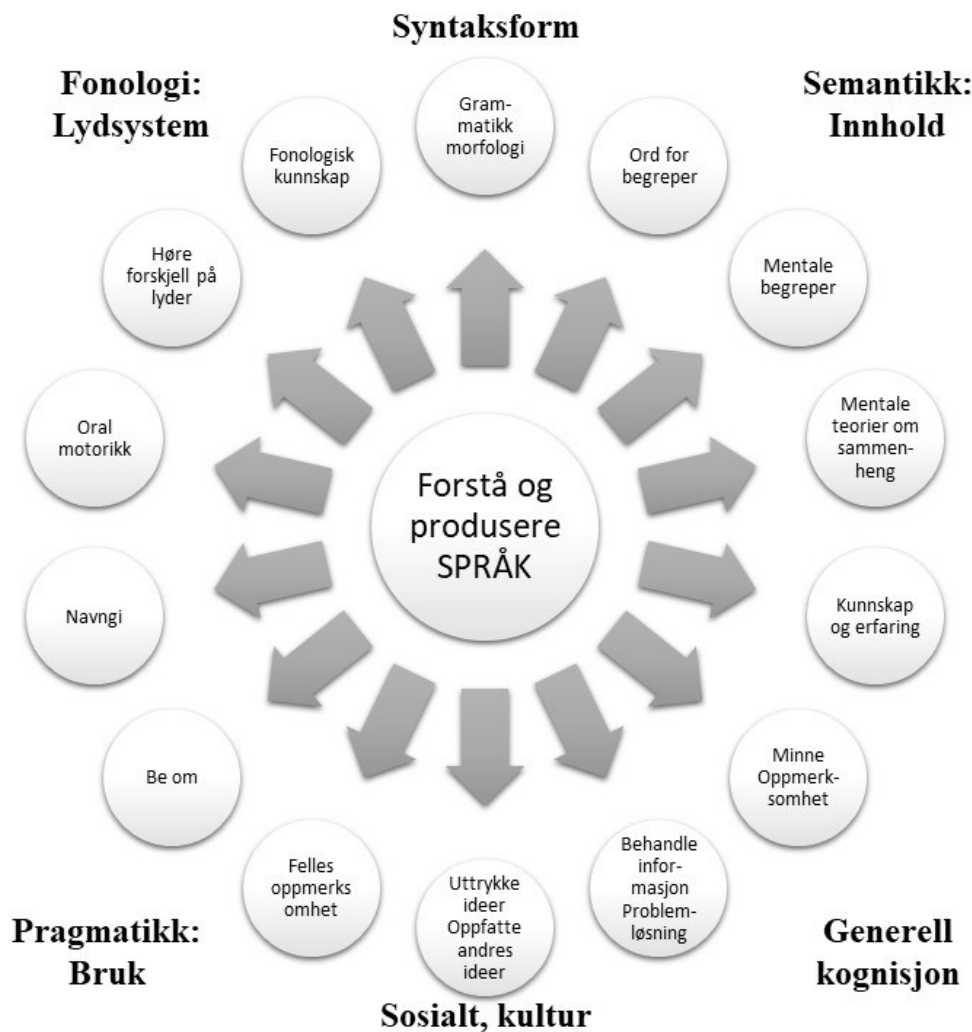
En annen utfordring med prøvene er normeringen. Hva skal prøvene normeres mot? Skal de normeres mot den generelle populasjonen i norsk skole på ulike trinn? Skal prøvene på morsmål normeres mot den generelle populasjonen i det aktuelle landet, eller den språklig tilgjengelige populasjonen i Norge? Og trenger vi det? Ifølge Opplæringslova §2-8 og §3-12 så er *rett til særskilt språkopplæring* delt inn i tre rettigheter: Særskilt norskopplæring, morsmålsopplæring og tospråklig fagopplæring, og elever kan ha rett til en eller flere av disse. Hvis skolen gir eleven særskilt norskopplæring, så kan vi tenke oss at å teste på morsmål ikke er så relevant, men dersom eleven også får morsmålsopplæring eller tospråklig fagopplæring, så er det gjerne også viktig å vite noe om morsmålet. Spørsmålet er om dette skal kartlegges, i hvilket omfang og om det er behov for å normere dette.

Normering i forhold til morsmålenes land kan være problematisk fordi disse kanskje ser på opplæring og kunnskaper på andre måter enn vi gjør i Norge. Skal normeringen da gå på den språklige populasjonen som er tilgjengelig i Norge? Det kan bety at vi for ulike språk kan få ulike nivåforskjeller fordi f.eks. sosiokulturell bakgrunn i ulike språkgrupper i Norge ikke nødvendigvis sammenfaller med normal-populasjon eller med hverandre. Både *TOSP* og *FLORO* har brukt gjennomsnitt og standardavvik i denne sammenhengen selv om bare *TOSP* er normert, og da bare normert ut fra en generell norsk populasjon på ulike trinn i grunnskolen. Skal vi se på normeringen gjort i en norsk populasjon og overføre denne normeringen til andre språk? Vil det være den 'beste' løsningen av disse? Men vil en norsk normering egentlig være gyldig for andre språk?

2.4. Språk og ferdigheter

Egeberg (2016) hevder at språket er det viktigste verktøyet mennesker har for å kunne forstå og samhandle med verden. Gjennom språket kan vi organisere, motta, behandle og formidle informasjon (Egeberg, 2016). I figur 2 ser vi at det er veldig mange ferdigheter som er med å påvirke språket når vi snakker om å forstå og produsere språk gjennom lesing og skriving.

Det er viktig å huske at språk – både forståelse og produksjon – er avanserte og sammensatte kognitive ferdigheter. Under er ferdighetene delt inn i leseferdigheter, lytteferdigheter, skriftlige ferdigheter og muntlige ferdigheter. Dette betyr ikke at dette er avgrensede ferdigheter, og mange av ferdighetene overlapper hverandre. Hovedgrunnen til at jeg har valgt å dele de inn på denne måten, er at kartleggingsverktøyene i stor grad ser ut til å følge denne inndelingen når de kartlegger norskferdighetene. Noen har også med elementer om språklæring, kultur og grammatikk som egne deler som kartlegges.



Figur 2. Ulike sider ved språk (Egeberg 2016, s.57)

2.4.1. Leseferdigheter

Lesing er en kombinasjon av avkoding og forståelse. 'The simple view of reading' (SVR) er en anerkjent lesemodell som ble utviklet av Gough og Tunmer i 1986. Denne modellen presenterer lesing som to komponenter: Avkoding og språkforståelse. Altså, leseforståelse = avkoding + språkforståelse (Gough, 1996). Elevene må først lære den tekniske ferdigheten, avkoding, før de kan forstå innholdet i det de leser, men de må også ha språkforståelse slik at de forstår innholdet i det de leser og da opplever leseforståelse. Avkoding er den tekniske koblingen fra grafisk tegn til tale, først på grafem-nivå (bokstaver) og fonologisk nivå (tilhørende språklyd), og seinere kunne trekke dette sammen til større helheter som stavelser og ord (Shultz, Haugen & Støre, 2008).

SVR legger til grunn at dersom avkodingsferdighetene eller språkforståelsen mangler, så vil leseforståelsen mangle. Følger man SVR, så vil en elev som mestrer avkoding ikke forstå

innholdet dersom ikke eleven har språkforståelse, eller et ordforråd, som gjør at de avkodede ordene gir mening. På samme måte vil en elev som mestrer avkodingen, men bare har noe språkforståelse, bare få en delvis leseforståelse. Hoover og Gough (1990) trekker også frem at ordforrådet er sterkt knyttet til språkforståelsen. I denne modellen blir det hevdet at elevens ordforråd har en betydning for leseforståelse, men ikke for avkodingsferdigheter (Hasle, 2013). Modellen har også blitt mye kritisert, hovedsakelig for å være for enkel. Flere forskere har pekt på at språklæring og lesing er så kompliserte kognitive prosesser, at det ikke kan forklares så 'enkelt' som SVR fremstiller det (Hasle, 2013).

Schultz, Haugen og Støre (2008) peker også på at det er viktig at elevene mestrer talespråket godt for at lesing ikke bare blir den tekniske avkodingen, men at det også blir en ferdighet for å få nye kunnskaper, og at eleven dermed forstår det som leses. De hevder at dersom talespråket ikke er godt utviklet, vil lesingen forbli en teknisk ferdighet fordi teksten ikke gir leseren mening, og dermed ingen nye kunnskaper eller forståelse. Det er en klar sammenheng mellom lesing og talespråket, og det kan ofte være en fordel for elevenes lese- og skriveopplæring at de først har hatt en lengre periode med muntlig opplæring (Schultz, Haugen & Støre, 2008). Premissene for lesing gjelder også for lytting, men her er avkodingen koblet til lyd i stedet for skrift. I *Rammeverket* er leseferdigheter knyttet til å kunne lese og forstå ord, fraser, setninger og tekst i ulike sjangere (Utdanningsdirektoratet, 2011).

I 2014 ble det gjort en metaanalyse av studier som sammenlignet leseferdigheter hos førstespråksinnlærere (L1) og andrespråksinnlærere (L2) (Melby-Lervåg & Lervåg, 2014). Her konkluderer de med at det ikke er forskjell på avkodingsferdighetene hos L1 og L2, men at det er et tydelig gap i leseforståelsen som kan knyttes til at L2 har svakere språkforståelse enn L1, og dette gapet holder seg ganske stabilt oppover i trinnene (Melby-Lervåg & Lervåg, 2014). Hasle (2013) peker på at forskning viser en sammenheng mellom ordforråd og leseforståelse, men hun viser også til forskning som hevder at det kan være en sammenheng mellom ordforråd og avkoding. Denne koblingen finner ikke Melby-Lervåg og Lervåg (2014) i sin metastudie.

Shultz, Haugen og Støre (2008) trekker frem at kartlegging av norskferdigheter hos minoritetsspråklige i skolen ikke nødvendigvis forteller noe om elevens faktiske nivå og evner, og ofte kan det se ut som om eleven har språkvansker fordi de ikke kan formidle sin forståelse, eller man kan bli lurt til å tro at eleven er flinkere enn de er fordi de gjetter seg frem til hva som er riktige svar. At elever kan forstå innholdet i en tekst uten at de nødvendigvis klarer å formidle denne forståelsen, eksemplifiserer de med en samtale om en

tekst elever får høre og så skal gjenfortelle. På norsk klarer ikke den ene eleven å gjenfortelle noe særlig fra teksten, men på morsmålet kan eleven gjenfortelle store deler av handlingen (Schultz, Haugen & Støre, 2008).

2.4.2. Lytteferdigheter

Husby og Kløve definerer lytteferdigheter som det å kunne forstå det talespråket som snakkes i samfunnet på de arenaene man ferdes i (Husby & Kløve, 2007). Husby og Kløve (2007) peker på at det er forskjell på resepsjon og persepsjon når man snakker om lytteferdigheter, og sier at når man snakker om lytteferdigheter i et andrespråksperspektiv, så snakker vi om persepsjon. Resepsjon er den lyden øret oppfatter, mens persepsjon er hjernens fortolkning av lyden (Husby & Kløve, 2007). Selv om øret vil motta den samme lyden uavhengig av morsmål, så vil hjernens fortolkning av lyden endres med morsmålet (Husby & Kløve, 2007).

Det å lytte er en kompleks ferdighet, og for å forstå det man hører så må man oppfatte elementer på flere språklige nivåer: fonologi, leksikon, syntaks og semantikk, tekst og pragmatikk (Husby & Kløve, 2007). Husby og Kløve (2007) hevder også at i en samtale vil man ha semantiske og pragmatiske forventninger til innhold som skal komme seinere i samtalen, og dette er viktig for å forstå den kommunikasjonen som skjer. Hvis man er i kassen på butikken så forventer man at man får informasjon om hva man skal betale, og gjerne spørsmål om man ønsker pose. Dersom denne personen plutselig snakker om været eller at det er forferdelig med rushtidavgift, så kan man oppleve kommunikative brudd hvis man ikke forstår skiftet i samtalen. Undersøkelser har vist at innfødte stiller andre krav når man vurderer andrespråkkompetansen enn det som fokuseres på i klasserommet (Husby & Kløve, 2007). En elevs lytteferdigheter må derfor vurderes ut fra i hvilken grad de fungerer i en kommunikativ sammenheng, altså det *Rammeverket* kaller samhandling, ikke hva eleven kan oppnå i 'laboratoriesammenheng' (Husby & Kløve, 2007). Denne tanken om at testing i andrespråksperspektivet skal være virkelighetsnært, er noe av det Bachman (2002) tar opp som viktig i vurdering av et andrespråk. Man må altså vurdere om eleven kan bruke språket sammen med andre. I eksempelet over vil det da være viktig at innlæreren klarer å mestre samtalen, selv med digresjoner. Denne måten å tenke på lytting som del av en kommunikativ kompetanse som skal skje i samhandling med andre har fått gjennomslag på viktige arenaer, slik som *Rammeverket*, og har dermed vært med å forme en endring i hvordan andrespråkselever skal vurderes i bl.a. lytteferdigheter.

Ifølge Husby og Kløve (2007) har undervisning i lytting hovedsakelig rettet seg mot kommunikative ferdigheter samtidig som testing av lytteferdigheter har hatt et delt fokus på

komponentene i språket (lyder, vokabular, struktur) og kommunikative ferdigheter. De forklarer at tester i lytting i dag vanligvis er utformet på samme måte og har oppgavetyper med «flervalg, ja/nei-spørsmål, rett/galt, utfylling av skjema/diagram/bilder/kart, diktat eller lukediktat (ortografi eller transkripsjon), notat- eller referatskriving» (Husby & Kløve, 2007, ss. 152-153). Noen tester gir også rom for å svare med egne ord. I *Rammeverket* er lytteferdigheter knyttet til å gjenkjenne ord og uttrykk og forstå innhold i talt språk, først med et veldig begrenset ordforråd som seinere utvides (Utdanningsdirektoratet, 2011).

2.4.3. Skriftlige ferdigheter

Korsgaard, Hannibal og Vitger (2011) hevder at leseprosessen er vanskeligere enn skriveprosessen fordi lesing handler om å gjøre noe ukjent til noe kjent. Når man skriver, vil man ha en formening om hva man ønsker å formidle, og må da finne ord og setninger som formidler dette som allerede er kjent. Korsgaard, Hannibal og Vitger (2011) forklarer at skriving er et komplekst fenomen fordi det krever ulike tilnæringsmåter avhengig av innhold, form og formål. Skriving er viktig i det vestlige samfunnet for det gir oss mulighet til å kommunisere med mennesker som er langt borte eller vi kan organisere tankene og hvordan vi ønsker å formidle noe (Korsgaard, Hannibal & Vitger, 2011). Dette skiller skriftlig og muntlig kommunikasjon fordi muntlig kommunikasjon er en 'her-og-nå'-aktivitet, mens skriftlig kommunikasjon ikke er avhengig av at mottakeren er tilstede, og man kan dermed planlegge og endre fremstillingen til man er fornøyd og velger å sende eller levere.

Rammeverket bruker begrepet kommunikativ kompetanse. I dette begrepet snakker de om lingvistisk, sosiolingvistisk og pragmatisk kompetanse. Altså, språket er alt fra de leksikalske, fonologiske og syntaktiske strukturene i språket, til sosiokulturelle betingelser for språkbruk og til den praktiske bruken av språket (Utdanningsdirektoratet, 2011). Det at noen bruker denne kommunikative kompetansen, kaller de språkhandlinger, og dette kan inkludere resepsjon, produksjon, samhandling eller videreformidling.

Flere av kartleggingsprøvene legger opp til at eleven bl.a. skal mestre å skrive i ulike sjangere, noe vi også ser konturene av i *Rammeverket*:

Tekstutformingskompetanse omfatter kunnskaper om tradisjoner for utforming av tekster i det aktuelle samfunnet, for eksempel

- hvordan det er vanlig å bygge opp informasjon ved hjelp av makrofunksjoner (beskrivelse, fortelling, redegjørelse osv.)
- hvordan historier, anekdoter, vitser osv. fortelles

- hvordan en sak legges fram (i retten, diskusjoner osv.)
- hvordan man setter opp, binder sammen og ordner skriftlige tekster (artikler, formelle brev osv.)

(Utdanningsdirektoratet, 2011, s. 148)

Dette ser vi også igjen i nivåbeskrivelsene i *Rammeverket* (vedlegg 7), der man f.eks. på nivå A1 skal kunne skrive enkle postkort og fylle ut skjema med personopplysninger, mens man på A2 avanserer til å skrive korte beskjeder og enkle, personlige brev. Fra B1 og oppover forventes det at man skal skrive detaljrikt og om mange ulike emner, samt at argumentasjon kommer inn som en viktig del.

2.4.4. Muntlige ferdigheter

Muntlig kommunikasjon handler både om å kunne uttrykke seg selv muntlig, og å kunne delta i samtaler med andre. I *Rammeverket* deles muntlige ferdigheter inn i to kategorier:

‘Samhandling’ og ‘Muntlig produksjon’. Samhandling fokuserer på å kunne delta i en samtale med andre, mens muntlig produksjon handler om å kunne fortelle om noe

(Utdanningsdirektoratet, 2011). Mens man på A1 i muntlig samhandling skal kunne delta i en enkel samtale med mye hjelp, skal man på A2 kunne delta mer selvstendig, men fremdeles på et enkelt nivå. Når man kommer opp på nivå B1 skal man kunne kommunisere i «de fleste situasjoner som kan oppstå når man reiser i et område der språk snakkes»

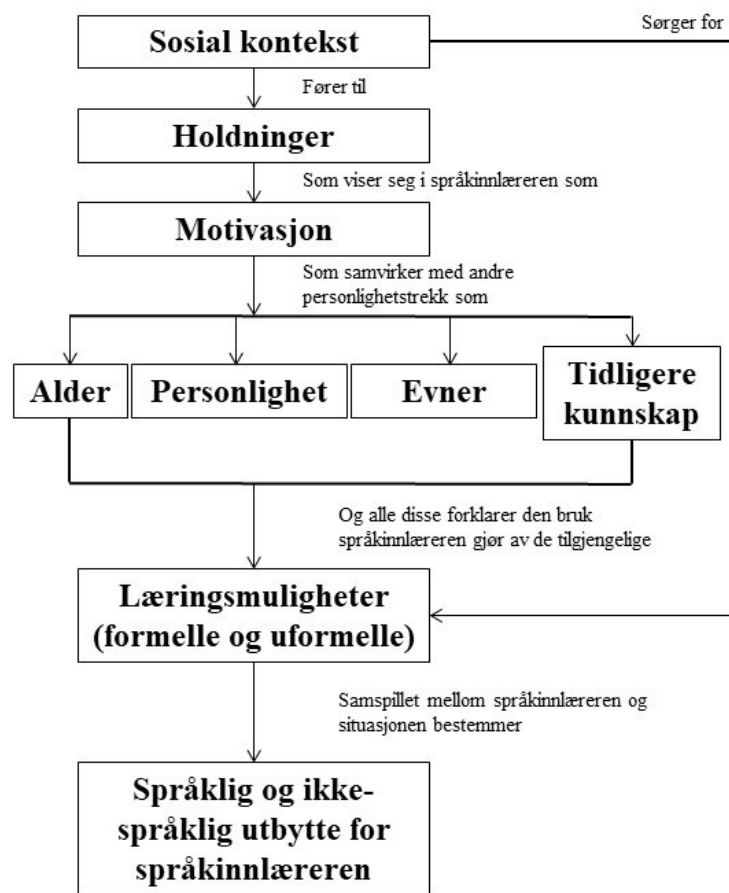
(Utdanningsdirektoratet, 2011, s. 32), så lenge det er snakk om hverdagslige emner. I muntlig produksjon går man fra å kunne beskrive bosted og kjente mennesker på A1 til å skrive om familie, andre mennesker, utdanning og arbeid på A2 og til å kunne gi «en enkel, sammenhengende framstilling av hendelser, erfaringer, drømmer, håp og planer»

(Utdanningsdirektoratet, 2011, s. 32), på B1, i tillegg til at man på B1 skal kunne begrunne meninger, planer, osv.

2.5. Norsk som andrespråk

Andrespråk er alle språk man lærer etter førstespråket (Berggreen & Tenfjord, 2007).

Berggreen og Tenfjord (2007) forklarer at når vi snakker om opplæring i fremmedspråk, så gis denne opplæringen hovedsakelig på førstespråket og vanligvis i et annet land enn fremmedspråkets, mens når vi snakker om opplæring i andrespråk, så skjer opplæringen et sted der dette andrespråket er dagligspråk i samfunnet rundt eleven. Det er veldig mange elementer som spiller inn når vi snakker om andrespråklæring. Det kan være sosiale faktorer, alder, førstespråkskompetanse og rammer for bruk og øvelse i språket (Berggreen & Tenfjord, 2007), slik vi kan se i figur 3.



Figur 3. Spolskys modell for andrespråklæring (Berggreen & Tenfjord 2007, s.28)

Som vi ser, er det å lære et nytt språk er en kompleks ferdighet med mange variabler som er med å påvirke språklæringen. Språkinnlæring skjer over tid, og er en gradvis prosess (Carlsen, 2012). Noen innlærer bruker lang tid på å lære et språk mens andre lærer fort, noen lærer fort til et visst nivå før det så stagnerer (Carlsen, 2012). Denne språkinnlæringsforståelsen kan sammenlignes med fjellklatring. Det er mange veier til toppen av fjellet, og ulike innlærere velger ulike veier med ulik vanskegrad (Carlsen, 2012; Carlsen & Moe, 2019). Noe av det som gjør språklæring ekstra komplekst er at man må jobbe med, eller forholde seg til, to ulike språk samtidig, også kalt å bytte kode (Tornberg, 2000; Koda, 2012). Polio (2012) peker på noe av det samme i forhold til skrive, at forskning har vist at det er vanskeligere å skrive på et andrespråk enn på et førstespråk fordi disse møter flere hindre i forhold til innhold og struktur i skriftlige tekster enn i språket i seg selv.

Innen andrespråksforskning er det kjent at når en innlærer holder på å lære et nytt språk, så 'lager' de er språk som ligger mellom førstespråket og målspråket, altså det språket de skal

lære. Dette kaller vi for *mellomspråk*. Mellomspråk er en beskrivelse for de grammatiske reglene og systemene som en innlærer ‘lager’ mens den holder på å lære et nytt språk, som et slags kodesystem (Berggreen & Tenfjord, 2007). I starten av innlæringen er dette mellomspråket veldig enkelt, med få regler, enkel struktur og ofte i stor grad påvirket av regler fra morsmål. Pickering (2012) har sett på hvordan morsmål påvirker innlæringen i et andrespråksperspektiv, og hennes undersøkelser viser at morsmål påvirker en del i begynnerinnlæringen, men jo mer man lærer om andrespråket, jo mindre vil morsmålet påvirke det. Berggreen og Tenfjord (2007) viser til en læringsrekkefølge for mellomspråk, og i figur 4 kan vi se hvordan denne ser ut for norsk.

Fase		Norske eksempler
1	Formularer: Utenat lærte konstituent	<i>Bare bra!</i>
2	Kanonisk rekkefølge: Fast forhold mellom konstituent	SV(OA) <i>Vi spiser middag seint.</i> <i>Vi spiser nå?</i>
3	Flytting av kant-konstituent	Framflytting av adverbiale ledd, ev. objekter om de står til slutt i setningen Spørreordspørsmål uten inversjon: <i>*Da vi spiste middag.</i> <i>*Middag vi spiste.</i> <i>*Hva du spiller?</i>
4	Flytting av interne konstituent	Framflytting av objekter fra setningsintern posisjon: <i>*Middag vi spiste seint.</i> Inversjon i spørsmål uten spørreord: <i>Spiste du middag?</i> Inversjon i helsetninger der subjektet havner i kantposisjon: <i>Da spiste vi ...</i>
5	Rokking av plassene til to interne konstituent	Inversjon i spørreordssetninger: <i>Når spiste du? Hva spiste du?</i> Inversjon i helsetninger med initialt ikke-subjekt: <i>Da spiste vi middag.</i>
6	Mestring av ulike konstituentrekkefølger i helsetning og leddsetning Løfting av Konstituent ut av leddsetninger	Ulik plassering av setningsadverbial i leddsetninger kontra helsetninger: <i>Som vi ikke spiste / vi spiste ikke</i> Inversjon i helsetning, ikke-inversjon i leddsetninger: <i>Når spiste du?</i> <i>Han spurte når jeg spiste.</i> Objektsinfinitiv: <i>Vi så dem spise.</i>

Figur 4. Utviklingsstadier ved læring av leddstilling i norsk/dansk (Berggreen & Tenfjord, 2007, s.272)

Her ser vi at språkinnlæringen starter med å bruke helt enkle ord og fraser i fastlærte fraser (1), og at man derfra avanserer med et utvidet ordforråd, men med en ganske fast

setningsoppbygging (2). Seinere vil man flytte på ulike setningsledd (3-5), både i enden og innad i setningene, før man til slutt mestrer setningsoppbyggingen i norsk (6).

2.6. Hva er språktesting?

Cyril J. Weir (2005) sier at noe av det viktigste man jobber med når man skal lage en språktest, er å definere hva man skal måle i testen og hvilke ferdigheter som måles i den aktuelle testen. Skal man f.eks. lese en tekst for å forstå innholdet, skal man skrive en tekst, måler man ordforråd, kunnskaper eller samhandling med andre? Dette er det viktig å vite på forhånd. Weir (2005) forklarer at man må ha kontroll på *både* hvilken metode man bruker *og* hvilke språkferdigheter man måler med den aktuelle språktesten. Når man lager en språktest kan man ikke utelukke f.eks. metodedelen fordi den er en del av det som gir svar på testen. Et eksempel kan være at dersom man skal måle leseferdigheter, så vil det være uheldig at den som testes må skrive svar, fordi skriving måler andre ferdigheter. Den som testes kan kanskje ikke skrive, men er god til å lese, og får da ikke vise ferdighetene sine pga. manglende skriveferdigheter.

To elementer er essensielle å ta opp i forbindelse med språktester er reliabilitet og validitet. Dette er to begreper som henger nært sammen, og under vil jeg først gjøre rede for reliabilitet for deretter gå inn på validitet.

2.6.1. Reliabilitet

Reliabilitet er når man måler i hvor stor grad resultatene er stabile eller konsistente. Med andre ord, kan vi stole på resultatene som gis deltakerne etter endt prøve? Bachman og Palmer (1996) definerer reliabilitet som konsistente målinger. Det betyr at resultatene fra en test skal være konsistente i ulike testsituasjoner uavhengig av hvilke deltakere som testes. Deltakerne skal altså få samme resultat på prøven uavhengig av hvor og når de testes, så lenge de har den samme kompetansen. Ifølge Bachman og Palmer er det ikke mulig å lage en test som er helt reliabel fordi resultatet alltid i noen grad vil være påvirket av andre elementer enn den ferdigheten prøven skal måle (Carlsen & Moe, 2019).

Bachman og Palmer (1996) peker på at f.eks. flervalgsprøver i grammatikk ikke er nok for å kunne si at dette er en god innplasseringstest siden den bare måler én snever ferdighet innenfor språkbruk. Noe av det viktigste Bachman og Palmer (1996) trekker frem, er at reliabiliteten i en test blir påvirket av formålet for testen, måten begrepene har blitt definert og hvordan oppgavene har blitt utformet. I såkalte 'high-stakes tests' er det viktigere at utviklerne setter høye krav til reliabilitet, mens i såkalte 'low-stakes tests' er det akseptabelt at

reliabiliteten er lavere. Carlsen og Moe (2019) forklarer at en 'high stake test' er et begrep som fagmiljøet bruker om prøver der resultatene innvirker i stor grad på avgjørelser i en persons liv. Dette kan f.eks. være prøver som avgjør om man får permanent opphold, statsborgerskap, eller eksamener som kan avgjøre hvilke videre studier man kan komme inn på eller hvilken jobb man kan få. På den andre siden har man 'low stakes tests', som ikke har direkte konsekvenser i personenes liv, men f.eks. måler fremgang og gir lærer informasjon som de kan bruke i den videre undervisningen (Carlsen & Moe, 2019).

En reliabel test er når vi kan forvente at resultatene på testen vil være veldig like ved gjentatt bruk (Weir, 2005; Hargreaves, 2010). Dette oppnår man bl.a. gjennom å kontrollere alle utenforliggende variabler: miljøet man testes i, rommet, instruksjoner, tidsbegrensninger, osv. Weir peker på at når de utenforliggende faktorene i liten grad varierer, så vil det redusere feil i skåring av testen, og dermed vil testen bli mer reliabel (Weir, 2005). Deretter ser man på om resultatene er reliable ved å se om de er stabile over tid og uten fordommer. På denne måten handler det om i hvilken grad prøven er uten målingsfeil og at man på den måten kan stole på resultatene av prøven (Weir, 2005). Hargreaves (2010) viser til at dersom det er 3 eller færre svaralternativ på en flervalgsprøve, så vil sjansen for å gjette riktig øke til over 30 %, som vil gjøre testen mindre reliabel. Hargreaves (2010) forklarer bl.a. at spørsmål med fire eller fem svaralternativer er gode og veldig gode sett reliabelt, mens spørsmål med bare to eller tre svaralternativer er lite reliable.

Weir (2005) sammenligner også språktester med fly, og sier at på samme måte som fly sjekkes før, under og etter flyvning, så må også språktester sjekkes. Han mener at reliabiliteten til testens resultater (besvarelser) skal forklares og begrunnes, både i henhold til innhold og metode og hvilke forutsetninger testen ble gitt under. Dette mener han kan gi testutviklerne uvurderlig informasjon i videre arbeid med utvikling og videreutvikling av tester.

Under er noen måter å undersøke reliabilitet på: *test-retest*, *alternate form*, *split-half*, *mellomtestledd* og *sensorreliabilitet*, i tillegg til et avsnitt om standardfeil og standardavvik.

2.6.1.1. *Test-retest reliabilitet*

Hvis vi bruker test-retest reliabilitet så betyr det å gi en elevgruppe en prøve og etter en periode å teste dem på nytt med den samme prøven. Da kan vi måle hvor stabile eller konsistente svarene som gis er, gjennom å undersøke korrelasjonskoeffisienten (Domino & Domino, 2006). Test-retest har utfordringer knyttet til tid mellom test og retest. Er tiden for

kort kan deltakerne huske spørsmålene fra forrige gang, og dette kan påvirke svarene, og om det går for lang tid mellom prøvene så kan deltakerne ha endret utgangspunkt, altså at de ikke er så 'like hverandre' som de var da de tok prøven første gang.

2.6.1.2. Alternat form reliabilitet

Ved å bruke alternate form reliabilitet, så er det to prøver som skal besvares der den ene gruppen starter med prøve A mens den andre gruppen starter med prøve B. Disse to prøvene skal være ekvivalente, eller likeverdige, på alle områder, som f.eks. instruksjoner, antall testledd, testleddinnhold, osv. (Domino & Domino, 2006). Domino og Domino (2006) peker på at hvis disse to prøvene deles ut etter hverandre, altså at man får prøve B ut etter å ha levert prøve A eller motsatt, så vil forskjeller i svarene hos den enkelte være pga. testleddinnholdet og på den måten kan reliabiliteten bli svakere fordi man faktisk måler forskjellige ting i de to prøvene selv om testleddene i utgangspunktet skal være ekvivalente (Domino & Domino, 2006). Om det derimot er lengre tid mellom de to prøvene, så vil reliabilitetskoeffisienten reflektere variasjon pga. både testleddinnholdet og tidsaspektet.

2.6.1.3. Split-half reliabilitet

Dersom man gir en gruppe elever en prøve med 100 testledd og deretter gir hver deltaker to skårer, en for partalltestledd og en for oddetalltestledd, så kan vi korrelere disse to svarsettene (Domino & Domino, 2006). Deltakerne tror de har svart på en prøve med 100 testledd, mens det egentlig er to prøver med 50 testledd. Utfordringen vil her bli å passe på at testleddinnholdet i partall- og oddetalltestledd er ekvivalente. Altså, at ikke de f.eks. er påvirket av at de første spørsmålene er enkle og de andre mer avanserte, men dette er noe man må ta hensyn til i andre reliabilitetsmodeller også. Split-half reliabilitet fokuserer på den indre konsistensen i prøven. Dette kan man regne ut ved å f.eks. bruke Rulon-formelen (Brown & Hudson, 2002; Domino & Domino, 2006). Da kalkulerer man først fire skårer for hver deltaker: partalltestledd, oddetalltestledd, differansen mellom oddetall- og partalltestledd og totalskåren (Domino & Domino, 2006). Deretter regner man ut variansen i differansen og totalskåren og setter inn i formelen. Dersom oddetall- og partalltestledd er helt konsistente så vil differansen være 0 og r vil være 1.

2.6.1.4. Mellomtestledd reliabilitet

Mellomtestledd reliabilitet er når hvert testledd i en test måler den samme variabelen (Domino & Domino, 2006). Dette kan man måle ved Kuder-Richardson 20 (K-R20) analyse, Kuder-Richardson 21 (K-R21) analyse eller Cronbachs alfa. K-R20 og K-R21 brukes til å teste testledd som kan skåres som rett/feil, sant/usant, ja/nei, mens Cronbachs alfa brukes for å teste

testledd der svarene har ulike verdier, som f.eks. likert-skala der man svarer i hvor stor grad man er enig/uenig eller frekvens.

2.6.1.5. *Sensorreliabilitet*

Det finnes ulike måter å måle sensorreliabilitet på. En måte er å bruke Many-Facet Rasch Måling, eller MFRM. Denne modellen måler i hvilken grad vi kan stole på resultatene sensorer gir deltakere på prøver. MFRM estimerer både kandidatens ferdigheter og oppgavens vanskelighet og om sensor er mild eller streng i sine vurderinger (Midtbø, Rossow & Sagbakken, 2018). På denne måten kan man analysere sensors vurdering av resultatene og deres retteatferd i tillegg til analysen av kandidatens ferdigheter og oppgavens vanskelighetsgrad.

2.6.1.6. *Standardfeil og standardavvik*

Ved å vite reliabilitetskoeffisienten av en prøve så har vi en indikasjon på stabiliteten i den prøven. Standardfeil er estimerer på hvor langt fra gjennomsnittet resultater vil være, mens standardavvik er spredningen i resultatene. Å regne ut standardfeil (SEM) må gjøres i flere omganger, og det er flere måter å gjøre det på. En måte er å begynne med å regne ut aritmetisk gjennomsnitt (\bar{X}), som er gjennomsnittsverdien.

$$\bar{X} = \frac{\text{summen av alle skårverdiene}}{\text{frekvensen av observasjonene}} = \frac{\sum X}{N}$$

Deretter kan man måle standardavviket (SD), som ser på spredningen i resultatene på en test og setter punkter i et diagram med gjennomsnittlige avstander til det aritmetiske gjennomsnittet.

$$SD = \sqrt{\frac{(\text{summen av alle skårverdiene} - \text{aritmetisk gjennomsnitt})^2}{\text{størrelsen på utvalget}}} = \sqrt{\frac{\sum(X - \bar{X})^2}{N}}$$

I formelen er X summen av alle skårverdiene, \bar{X} er aritmetisk gjennomsnitt, og N er størrelsen på utvalget. For å kunne regne ut standardfeil i måleinstrumentet, må man også regne ut reliabilitetskoeffisienten.

$$SEM = SD\sqrt{1 - r_{xx}}$$

Dermed kan vi regne ut standardfeilen i måleinstrumentet (SEM). I formelen har man standardavviket for måleresultatet (SD), og r_{xx} er den utregnede reliabilitetskoeffisienten (Befring, 2002; Brown & Hudson, 2002; Domino & Domino, 2006).

2.6.2. Validitet

Ifølge Huges handler validitet om i hvor stor grad en prøve måler det den sier den skal måle (Carlsen & Moe, 2019). Derfor er det viktig å sikre validitet i testutvikling, og Carlsen og Moe (2019) peker på at det er testutviklernes oppgave å passe på at testen er valid og rettferdig.

Det finnes ulike typer validitet som kan måles, og ulike forskere og fagfolk bruker ulike begreper på validitet. Under går jeg kort gjennom først Chronbach og Meels validitetsbegreper, deretter kritikken Messick kom med mot dette synet på validitet, så kort om Messicks validitetsbegreper, som American Psychology Association (APA) og Association of Language Testers in Europe (ALTE) i stor grad ser ut til å ha tatt en dreining mot.

2.6.2.1. *Begrepsvaliditet, kriterievaliditet og innholdsvaliditet*

Denne inndelingen av validitet i begrepsvaliditet, kriterievaliditet og innholdsvaliditet ble presentert av Cronbach og Meel allerede på 1950-tallet (Colliver, Conlee & Verhulst, 2012). Begrepsvaliditet brukes som begrep for å tolke testskårer som en indikator på de ferdighetene man måler (Bachman & Palmer, 1996). Begrepsvaliditeten er konstruerte teoretiske begreper som innehar et sett spesifikke trekk som er laget for et bestemt formål. Det betyr at begrepsvaliditet er en prosess som aldri tar slutt og som er nødvendig for å få et innblikk i om det er en refleksjon av den teoretiske rammen (Domino & Domino, 2006). Weir (2005) peker på at vi statistisk kan gå inn og undersøke hvilke språkferdigheter som måles i ulike tester, og sier at vi da vurderer begrepsvaliditeten. Bachman definerer begrepsvaliditet som en undersøkelse som tar for seg at f.eks. lesing og skriving vurderes som separate ferdigheter, og dette henger tett sammen med Weirs syn på at de teoretiske begrepene til testen må avklares før testen lages (Weir, 2005). Bachman (2002) peker på at når man måler ferdigheter i språktester så kreves det vanligvis språk for å fortolke og forstå oppgaven, eller det man skal gjøre, men man skal også bruke språk til å kunne gi svar. Dette mener han må være tydelig fra start fordi det inkluderer flere ferdigheter som måles sammen. Begrepsvaliditet handler om hvor meningsfulle og formålstjenlige fortolkninger vi gjør av testresultater. Vi må altså ha beviser for at resultatene på testen måler de områdene vi ønsker å måle, og i liten grad måler andre elementer eller ferdigheter (Bachman & Palmer, 1996). Hvis vi har en lyttetest, så er det viktig at det er lytteferdigheten som måles, og ikke f.eks. lesing eller skriving. Dersom man må lese for å krysse av for et svar eller skrive svaret, så vil det svekke begrepsvaliditeten i testen.

Mens begrepsvaliditeten handler om i hvilken grad testen måler det den sier den måler, så handler kriterievaliditet om hvilke studier som er gjort for å se om kriteriene for testen kan anses som gyldige (Evers, Hagemester, Høstmælingen, Lindley, Muñoz & Sjöberg, 2013). Kleven og Hjordemaal (2018) forklarer dette som i hvilken grad man kan overføre konklusjonen til andre sammenhenger og situasjoner, eller i hvilken kontekst resultatene kan anses å være gyldige i. Kriterievaliditet er når hensikten med prøven er å måle noe viktig som ligger utenfor selve måleinstrumentet (Carmines & Zeller, 2011), og disse kriteriene kan det være vanskelig å fastslå. Dette betyr at man gjennom måleinstrumentet blir målt mot kriterier utenfor testen. Hvis man gjennomfører en skrivetekst, så vil man lese oppgaver og skrive en eller flere tekster basert på oppgaveteksten. Deretter måles man gjerne etter kriteriene som ligger i språknivåene i *Rammeverket*, som handler om hvilke skriftlige ferdigheter man mestrer ifølge disse kriteriene. Her er det viktig at det er klart definert hva det skal legges vekt på i vurderingen: Formuleringssevne, ordforråd, argumentasjon, at man svarer på oppgaven, sjanger, rettskriving, syntaks, osv. Det er også viktig at man bruker kriterier som hører til i det domenet som man skal måle noe i (Domino & Domino, 2006; Carmines & Zeller, 2011). Det betyr at dersom man skal måle muntlig kommunikasjon i en språktest, så er det viktig at kriteriene måler de elementene som er nødvendig for muntlig produksjon, f.eks. ordforråd, uttale, konstruksjon av setninger, osv., og at den ikke også måler f.eks. lesing. Det er også viktig å merke seg at prøven ikke kan bli bedre enn kriteriene den blir målt mot (Domino & Domino, 2006).

Innholdsvaliditet er om prøven dekker alle områdene som skal måles, og dette er gjerne særlig viktig i oppnåelsesprøver, prøver som «[...] er knyttet til pensum eller viktige kompetansemål i læreplanen» (Carlsen & Moe, 2019, s. 41), der målet er å se om eleven har lært det de skal. Innholdsvaliditet handler om tolkningen av relasjonene mellom ulike variabler innad i prøven (Kleven & Hjordemaal, 2018). Denne typen validitet viser viktigheten av at prøven er representativ for de områdene eller domenene man skal måle.

2.6.2.2. *Messicks kritikk*

De tredelte validitetsbegrepene til Cronbach og Meel anså Messick som problematisk fordi disse tre kategoriene ikke var likevektige, og det førte til en kategoriseringstenkning av validitetsbegrepet (Goodwin, 2002). Messick (1994) peker på at når man f.eks. har få beviser på kriterievaliditet, så kan man likevel gjerne finne beviser for andre typer validitet i testen, men da må man kunne undersøke andre typer validitet. Diskusjonen om validitet har dreid i retning av hvilke typer valide beviser man kan finne i en validitetsundersøkelse, fremfor

hvilken type validitet noe kan kategoriseres som (Goodwin, 2002). APA og ALTE har i stor grad har dreid fra tredelingen av validitetsbegrepet og over til Messicks definisjon i *The Standards for Educational and Psychological Testing* (Goodwin, 2002; Goodwin & Leech, 2017). Her har Messicks begrepsvaliditet blitt stående som en paraplybetegnelse, og under denne paraplyen finner vi innholdsvaliditet, responsvaliditet, validitet i indre strukturer, validitet i relasjon mellom variabler, og konsekvensvaliditet (Downing, 2003).

Innholdsvaliditet ser på testspesifikasjoner, i hvilken grad testleddene inneholder det testspesifikasjonene beskriver og i hvilken grad testleddene er representative for det domenet, eller området, som testes. Responsvaliditet handler om at den som tar prøven skal være kjent med prøveformatet, at det er nøyaktighet i skåring av resultatene og at det gjennomføres analyser på de ferdighetene som testes, at pilotresultater valideres, og at prøveresultatene er forståelige for deltakerne. Validitet i indre struktur handler om testleddanalyser, stabilitet i resultatene og standardfeil, mens validitet i relasjon mellom variabler handler om korrelasjoner med andre relevante variabler, og generaliserbarhet. Konsekvensvaliditet ser på konsekvenser prøveresultatet kan ha for deltakerne og samfunnet, konsekvenser for læring, om positive konsekvenser utligner negative konsekvenser, konsekvenser for bestått/ikke bestått og metoder for å sette bestått-skårer. Som vi kan se under, overlapper dette i stor grad med Messicks begrepsvaliditet.

2.6.2.3. *Messicks begrepsvaliditet*

Messick (1994; 1996) ser på begrepsvaliditet som et overordnet validitetsbegrep, som favner om seks aspekter: innhold, substans, struktur, generaliserbarhet, ekstern og konsekvens. Innholdsaspektet handler om at innholdet i prøven skal gjenspeile det prøven hevder å måle, og ofte er dette knyttet til en eller annen form for pensum. Substansaspektet handler om at de ferdighetene kandidaten bruker for å løse oppgaven, er de ferdighetene som måles i prøven. Strukturaspectet ser på hvordan prøven er bygd opp, både oppgaver og vurderingskriterier, og om prøven i tilstrekkelig grad gjenspeiler det teoretiske fundamentet prøven er bygget på. Det betyr at dersom oppgaver eller vurderingskriterier ikke omfatter alt det man tenker prøven skal måle, så vil validiteten i strukturaspectet svekkes. Generaliserbarhetsaspektet undersøker om det er mulig å generalisere ut fra prøveresultater, og i så fall i hvilken grad. Dette aspektet knyttes tradisjonelt til reliabilitet fordi det ser på stabilitet i målinger (Carlsen & Moe, 2019). Det eksterne aspektet ser på prøveresultatet i forhold til andre eksterne mål, f.eks. at man sammenligner testresultater til en gruppe kandidater, eller at man ser på samsvar mellom prøver som skal måle den/de samme ferdighet(ene). Konsekvensaspektet handler om hvilke

konsekvenser tolkninger av resultater kan få for kandidatene og for samfunnet, særlig i forhold til fordommer og rettferdighet.

Ifølge Messicks definisjon, handler begrepsvaliditet om at det empiriske materialet og det teoretiske fundamentet til testen i tilstrekkelig grad må støtte det som er passende tolkninger og handlinger basert på prøveresultatet (Messick, 1994; Messick, 1996). Bachman og Palmer tar i bruk Messicks definisjon av validitet når de sier at en 'god prøve' skal fungere godt i en definert sammenheng, for et bestemt formål og en bestemt gruppe (Carlsen & Moe, 2019). Det er også viktig å ta høyde for hvordan prøveresultater blir tolket og brukt i etterkant. Dette gjelder særlig store og viktige ferdighetsprøver fordi disse prøvene kan påvirke læring som skjer før prøven kommer. Hvis en muntlig prøve fokuserer på uttale og egenproduksjon, så vil fokuset i undervisningen være på dette og ikke på f.eks. kommunikative ferdigheter. Carlsen og Moe (2019) henviser også til Kunnan, som sier at selv om en prøve er god for én gruppe, så betyr ikke det at den er god eller like god for en annen gruppe. En viktig oppgave testutviklere har, er dermed ikke bare å utvikle gode språktester, men også å kvalitetssikre språktester som allerede er på markedet (Carlsen & Moe, 2019).

Ifølge Weir (2005) og Carlsen og Moe (2019) er det hovedsakelig to store trusler mot validiteten til en test: konstruktunderrepresentasjon og konstrukt irrelevans (Weir, 2005). Konstruktunderrepresentasjon betyr at dersom elementer ikke legges vekt på i testen, så kan det føre til at disse elementene ikke undervises på språkkurs eller i fag som er relevante før man skal ta testen (Weir, 2005; Carlsen & Moe, 2019). Weir hevder at dersom det f.eks. ikke brukes lytteprøver i testen, så vil ikke lærere undervise i lytteøvelser. Konstrukt irrelevans handler om viktigheten av at oppgavene i testen lages på en slik måte at de er relevante for det virkelige livet, og at det ikke er abstrakte oppgaver uten relevans. Carlsen og Moe (2019) tar med et eksempel om at dersom prøven skal måle grammatikk og ordforråd, så vil det være irrelevant for prøvekonstruktet om den samtidig tester leseforståelse. I tillegg mener Weir (2005) at oppgaver i språktester bør ligge så nært autentiske situasjoner som mulig.

2.7. Det felles europeiske rammeverket for språk

Det er helt klart at *Det felles europeiske rammeverket for språk (Rammeverket)* har vært nyttig for å skape en felles forståelse av språknivåer. Nivåbeskrivelsene i *Rammeverket* er bygd opp med hele 56 ulike nivåer som beskriver språklige ferdigheter innenfor områdene å lese, lytte, skrive og snakke (Carlsen & Moe, 2019). I tillegg beskrives de ulike delkomponenter, som grammatikk, ortografi, vokabular, koherens og kohesjon (Carlsen, 2012; Carlsen & Moe,

2019). Carlsen trekker frem at noe av det mest positive med nivåbeskrivelsene, er at det er *positive* beskrivelser av hva eleven *kan* eller *mestrer*, i motsetning til det vi ellers ofte ser i skolen, at elevene ved å få en lavere karakter enn 6 blir fortalt hva de *ikke kan*. *Rammeverket* bygger på Bachmans kommunikative språksyn, der språk er en ferdighet i kommunikasjon, ikke noe man har kunnskap om (Carlsen, 2012).

Nivåbeskrivelsene i *Rammeverket* er dogmatisk, og de er teoretiske konstruksjoner (Carlsen, 2012; Carlsen & Moe, 2019), men resten av materialet i *Rammeverket* er ikke dogmatisk, og kan dermed tilpasses situasjonen og ulike innlærergrupper. Det har vært en del kritikk mot nivåbeskrivelsene i *Rammeverket* fordi de er teoretiske konstruksjoner fremfor å være basert på empirisk fundament (Alderson, 2007; Carlsen, 2012). Carlsen og Moe (2019) forklarer at det er viktig at nivåbeskrivelsene er dogmatiske, fordi om man endrer innholdet i beskrivelsene for å passe innlærergrupper eller land, så vil man snakke om nivåer med ulikt innhold, og ikke lenger ha den samme forståelsen av språknivåene. På samme måte er det viktig at resten av materialet i *Rammeverket* ikke er dogmatisk, slik at man kan tilpasse materialet best mulig til undervisningssituasjonene. Alderson (2007) går så langt i sin kritikk at han spør hvordan man kan identifisere styrker og svakheter når det gjelder språk, og om disse i det hele tatt kan hevdes å være ekvivalente, eller likeverdige, mellom ulike språk. Han viser til en undersøkelse som ble gjort med språktesten DIALANG², som viser at innlærere på A2 også kan lese mellom linjene og forstå meninger i tekster på samme måte som innlærere på C1, men han peker på at forskjellen ligger i hvor avansert teksten innlærere møter er, ikke hvilke evner eller ferdigheter innlærerne klarer å bruke (Alderson, 2007). Dette er bakgrunnen hans for å hevde at man må se på hvorfor noen *ikke* mestrer ferdigheter fremfor å ensidig fokusere på hva innlæreren får til.

Når man skal måle ferdigheter, så er det viktig at de som bruker måleenheten har samme forståelse for hva som ligger i vurderingene i måleenheten (Carlsen & Moe, 2019). Carlsen og Moe (2019) peker også på at selv om man bruker like benevnelser i vurderinger, så betyr ikke dette at innholdet i nivåene er de samme. Her er man avhengig av en felles forståelse for hva som ligger i ulike nivåer, og at man forholder seg til denne og ikke lager egne kriterier. Dette

² DIALANG er et prosjekt der det er utviklet en digital språktest på 14 europeiske språk som måler ferdigheter i lese, lytte, skrive, grammatikk og ordforråd. Testen er bygd opp rundt *Rammeverket* og har som formål å gi en diagnostisk tilbakemelding på prøven uten å gi en skår slik at innlærere kan forstå hvor de er innenfor *Rammeverket*, den skal diagnostisere evner innenfor hvert ferdighetsområde, og innlæreren kan se hva han gjorde rett og feil underveis og dermed også vurdere egne ferdigheter. (Alderson, 2007)

er en av grunnene til at man skal se på ferdighetsnivåene i *Rammeverket* som en dogmatisk referanse (Carlsen & Moe, 2019).

3. Metode

Valg av metode henger tett sammen med valg av problemstilling. Det er problemstillingen som legger føringer for hvilke metoder man kan velge å bruke for å finne svar på problemstillingen (Jacobsen, 2016).

Jeg har valgt å starte med en spørreundersøkelse for å finne ut hvilke kartleggingsverktøy som brukes av populasjonen jeg er interessert i: skoler som har innføringstilbud i tidligere Hordaland fylke. I tillegg til spørreskjemaet gjorde jeg en såkalt systematisk gjennomgang av måleinstrumentene (kartleggingsverktøyene) som brukes i skolene. For å gjøre denne gjennomgangen så brukte jeg et verktøy, *EFPA Test Review Model*, utviklet av European Federation of Psychologists' Association (EFPA), for å analysere validiteten og reliabiliteten i testinstrumentene (Evers, et al., 2013). Jeg har ikke selv analysert reliabilitet og validitet i instrumentene, men altså brukt informasjonen om reliabilitet og validitet som er gjort tilgjengelig i testmateriellet med veiledninger og fra testutviklere.

3.1. Spørreskjema, utvalg og prosedyre for datainnsamling

I mitt masterprosjekt har jeg sett på språkkartlegging i innføringstilbud og validiteten og reliabiliteten i instrumentene som brukes til denne kartleggingen. Min populasjon har vært lærere ved skoler med innføringstilbud i tidligere Hordaland fylke. Spørreundersøkelsen ble pilotert til fem lærere som jobber med elevgruppen på min arbeidsplass, før den ble sendt ut til skolene.

Spørreskjemaet ble sendt til 25 skoler, og etter flere epost-purringer satt jeg igjen med 10 svar, som bare er 40% svar. Jeg valgte da å ta kontakt med skolene per telefon, og endte opp med svar fra totalt 17 skoler, som tilsvarer 68%. Tabell 1 viser fordelingen av svar:

Tabell 1	Antall svar
Barneskole	9
Ungdomsskole	4
Videregående skole	4
Totalt	17

Jeg valgte å starte med et spørreskjema (vedlegg 1) og et informasjonsskriv (vedlegg 2) som ble sendt digitalt via epost til 25 skoler som har innføringstilbud i tidligere Hordaland fylke. Spørreundersøkelsen var frivillig og anonym, og tok for seg bruk av kartleggingsprøver i innføringsklasser. Spørreskjemaet var enkelt og inneholdt bare fire hovedspørsmål:

1. Innføringstilbudet er lokalisert på: barneskole, ungdomsskole, videregående skole

2. Hvilke kartleggingsmateriell bruker skolen for å kartlegge minoritetsspråklige elever i innføringstilbud?
3. Hva brukes kartleggingsprøven til?
4. Hvor mange ganger i året kartlegges elevene?

I første spørsmål skulle lærerne krysse av for om de arbeidet på en barneskole, ungdomsskole eller videregående skole. I andre spørsmål fikk de listet et utvalg kartleggingsprøver de kunne krysse av for, inkludert om de brukte en lokal utviklet prøve og et åpent alternativ om de brukte andre prøver enn de som var oppgitt.³ I spørsmål tre var det fem svaralternativ: å plassere eleven i nivådelt gruppe, å tilpasse opplæringen til eleven, å undersøke om eleven har vansker på andrespråket, vet ikke, og et åpent spørsmål der lærerne kunne skrive inn om de brukte prøvene til noe annet. Fjerde spørsmål var åpent og ba lærere skrive antall ganger elevene blir kartlagt.

Spørreskjemaet skal fortelle meg hvilke kartleggingsprøver som faktisk brukes i innføringstilbud, og her vil jeg også se om det er tendenser til at noen favoriseres fremfor andre på ulike «trinn»⁴. Spørsmål 2 er derfor det viktigste for meg i spørreskjemaet fordi det er dette som er grunnlaget for den videre analysen min av validitet og reliabilitet i kartleggingsverktøyene som brukes i skolene. I tillegg spør jeg *hva* prøvene brukes til og *hvor ofte* de brukes. Dette spørsmålet får de for hver kartleggingsprøve de bruker fordi det kan være at ulike prøver brukes til ulike formål. Bruker man prøvene til å nivådele, tilpasse opplæringen, finne ut om vansker på andrespråket eller andre ting?

For å kunne gjennomføre spørreundersøkelsen, måtte man samtykke til å delta (vedlegg 3). Spørreundersøkelsen er godkjent hos NSD med innsamling av e-postadresser for å enkelt kunne fjerne deltakere som ønsker å trekke seg fra deltakelse i prosjektet (vedlegg 4). Når masteroppgaven er godkjent vil alle data fra spørreundersøkelsen slettes. Dette er viktig i forhold til personvernet (Jacobsen, 2016), samtidig som det er viktig for meg å sjekke at jeg ikke får flere svar fra samme respondent. Målet med spørreundersøkelsen er ett svar fra hver skole. I presentasjon av data i oppgaven vil ingen respondenter være identifiserbare og all data vil være anonymisert.

³ Hvilke prøver som ligger i spørsmålet, kan ses i spørreskjemaet i vedlegg 1

⁴ Med «trinn» menes her kategoriene «barneskole 1.-7.trinn», «ungdomsskole 8.-10.trinn» eller «videregående skole 11.-13.trinn».

Det er også veldig viktig at resultatene fra undersøkelsen blir presentert fullstendig og i riktig sammenheng. Jacobsen (2016) sier at dette kan gjøres ved å presentere råutskriftene av alle tabellene man bruker i spørreskjemaet. Alle analyser man gjør av disse dataene vil være en reduksjon av detaljer og mangfold, og derfor er en presentasjon av rådataene er nødvendig for å sikre en riktig presentasjon av dataene. Det vil også tydeliggjøre at vi ikke tar resultater ut av sammenhenger, og åpenhet gjennom eksplisitte valg i forskningen er et krav Jacobsen trekker frem som veldig viktig (Jacobsen, 2016).

3.2. Systematisk litteraturgjennomgang av andrespråkstester

I denne oppgaven gjorde jeg en systematisk litteraturgjennomgang for å kunne identifisere, evaluere og sammenstille eksisterende forskning og empiri, som kartleggingsprøver som brukes i innføringstilbud (Fink, 2014; Booth, Sutton & Papaioannou, 2016). Dette ses ofte på som en kontorbasert metode der man gjør en sekundæranalyse for å utforske abstrakte konsepter og kunnskaper (Jesson, Matheson & Lacey, 2011). Denne litteraturgjennomgangen kan finne ut om resultater er konsistente, om det er svakheter i empirien og videre om det er behov for mer forskning på området. En slik litteraturgjennomgang vil gjøre informasjonen om kartleggingsprøvene mer tilgjengelig og resultatene blir mer transparente.

For å gjøre en god litteraturgjennomgang er det viktig å ha kontroll på datamaterialet, og her har jeg valgt å bruke en matrise for å stille opp resultatene (Garrard, 2011). Dette er en måte å strukturere datainnsamlingen på, en indeksering. Denne måten å strukturere informasjonen på skal altså gi orden og oversikt over innholdet i datamaterialet slik at det kan analyseres effektivt.

3.3. EFPA Test Review Model

EFPA Test Review Model har som mål å gi både en beskrivelse av måleinstrumenter og en detaljert vurdering av tester som brukes i bl.a. arbeidsmarkedet, utdannings- og helsesektoren (Evers, et al., 2013). Dette er et verktøy som er utviklet av European Federation of Psychologists' Association (EFPA) for å teste kvaliteten ved instrumenter som måler hva som helst innenfor utdanningsmessige og psykologiske tester. De hevder dette fordi de følger *Standards for Educational and Psychological Testing*, der begrepet «test» innbefatter alle typer tester, målinger og prøver, uavhengig av hva den kalles (Evers, et al., 2013).

EFPA Test Review Form er bygd opp i tre deler. Den første delen er en beskrivelse av instrumentet, den andre delen er en vurdering av instrumentet og den tredje delen er referanser (Evers, et al., 2013). Beskrivelsen inneholder følgende deler: (1) Generell beskrivelse

(instrumentnavn, forfattere, utgiver, publikasjonsdato), (2) Klassifisering (innholdsdomene, populasjon, skalaer og variabler som måles, responsmodus, administrasjonstid), (3) Måling og skåring (skåringsprosedyre, skalaer som brukes, transformasjon fra standard skårer), (4) Automatiserte digitale rapporter (tilgjengelighet, media, kompleksitet, struktur, mottakelighet til kontekst, muligheter for modifikasjon, transparenthet, språk og tone, tilsiktet mottakere), og (5) Vilkår og kostnader (dokumentasjon, publiseringsmetoder, start- og gjentagende kostnader, pris på rapporter, nødvendige testrelaterte kvalifikasjoner).

Vurderingen inneholder: (1) Kvalitet i logisk begrunnelse, forklaring, presentasjon og informasjon, (2) Kvalitet i testmateriellet (papir og blyant-prøver, digitale prøver), (3), Norm (norm-referert tolkning, kriterie-referert tolkning), (4) Reliabilitet (tilgjengelige data, indre konsistens, test-retest, ekvivalens i form av parallell eller alternate forms, item response teori-basert metode [IRT], inter-rater reliabilitet), (5) Validitet (begrepsvaliditet, kriterievaliditet, total validitet), (6) Kvalitetsrapporter generert av datasystem (ramme eller dekning, reliabilitet, relevans eller validitet, rettferdighet, akseptabelhet, lengde, total tilstrekkelighet), og (7) Endelig vurdering (konklusjon og anbefaling). Disse punktene vurderes etter denne skalaen:

Rating	Forklaring
N/A	This attribute is not applicable to this instrument
0	Not possible to rate as no, or insufficient information is provided
1	Inadequate
2	Adequate
3	Good
4	Excellent

Figur 5. Modell som forklarer vurderingene i EFPA's Test Review Model (Evers et al., 2013, s.25)

Veiledningen til *EFPA Test Review Form* forklarer at instrumenter som får en eller flere vurderinger på [0] eller [1] på områder som anses som kritiske for riktig og sikker bruk av instrumentet, ikke skal vurderes til å ha møtt minstestandarden (Evers, et al., 2013). Her kan man gjerne hevde at reliabilitet og validitet er det viktigste i kvalitetssikring av en prøve: Å sikre at prøven måler det man tenker den skal måle, og at samme deltaker kan forvente å få samme resultat uavhengig av hvor og når prøven tas.

3.4. Validitet og reliabilitet

Validitet og reliabilitet i kartleggingsprøvene skolene bruker til språkkartlegging av minoritetsspråklige elever i innføringstilbud er det viktigste jeg skal se på i min masteroppgave. Den teoretiske forankringen i forståelsen av begrepene validitet og reliabilitet, som jeg har gjort rede for i teorikapitlet, er derfor en viktig del av analysearbeidet.

Validitet og reliabilitet i kartleggingsprøvene vil bli undersøkt ved å bruke *EFPA Test Review Model*. Som jeg beskriver i 3.3, er dette en grundig gjennomgang av alle prøvens sider, fra beskrivelser, teoretisk bakgrunn, selve prøven, normering, reliabilitet og validitet.

3.5. Analyse

I dette prosjektet skal jeg gjøre flere analyser. Den største analysen er å måle validitet og reliabilitet i kartleggingsprøvene som brukes i tidligere Hordaland fylke. Til dette arbeidet bruker jeg *EFPA Test Review Model*, et anerkjent verktøy for å måle validitet og reliabilitet i tester, bl.a. for utdanningsformål.

I spørreskjemaet har jeg, i tillegg til spørsmål om hvilke kartleggingsprøver som brukes, spørsmål om på hvilket trinn, hvor ofte og til hvilket formål de bruker kartleggingsprøvene. Her skal jeg se om det finnes sammenhenger mellom om enkelte prøver brukes mer på noen trinn enn andre og om noen prøver brukes mer i forhold til noen formål.

4. Resultater

Resultatene presenteres i to deler: I delkapittel 4.1. *Hvilke prøver bruker skolene?* presenteres resultatene for forskningsspørsmålene «Hvilke kartleggingsprøver brukes til å kartlegge norskferdigheter i innføringstilbud i tidligere Hordaland fylke?» og «Hvor ofte og til hvilket formål brukes kartleggingsprøvene?» I delkapittel 4.2. *Vurdering av validitet og reliabilitet i instrumentene* presenteres resultatene for forskningsspørsmålet «Hvilken validitet og reliabilitet har måleinstrumentene?»

4.1. Hvilke prøver bruker skolene?

I tabell 2 ser vi antall skoler som bruker de ulike kartleggingsverktøyene. Vi ser at *Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk* er den som er mest utstrakt i bruk, men vi ser også at den brukes mest i grunnskolen og lite i videregående skole. *Trondheimstesten* og lokalt utviklede prøver brukes også i stor grad i grunnskolen. *Migranorsk, Norskprøven* og *Vestlandsverktøyet* er mest brukt i videregående skole.

Tabell 2 Antall innføringstilbud som bruker instrumenter for å kartlegge språkferdigheter hos minoritetspråklige elever (N = 17)

	Barneskole	Ungdomsskole	Videregående skole	Totalt
Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk (Udir)	8	3	1	12
Den europeiske språkpermen*				0
FLORO	1	1		2
TOSP	1			1
Migranorsk			3	3
Kartleggingsverktøy for minoritetspråklige elever i VGS (Vestlandsverktøyet)		1	2	3
Norskprøven for voksne**			2	2
Kartleggingsprøve for flerspråklige elever (Trondheimstesten)***	6	1		7
Lokalt utviklet prøve	6	1		7
Andre prøver	1			1
Totalt****	16	6	8	38

*Aldersdifferensiert 6-12 og 13-18 år

**Både gammel og ny versjon

***Nivådifferensiert: 5-6 år, 3.-4.trinn, 5.-7.trinn

****Totalantall tilsvarer ikke antall skoler fordi det er mulig med flere svar

Tabell 3 viser at alle skoler kartlegger *minst* én gang i året, og flere kartlegger mer enn tre ganger i året. Flere rapporterer også at elever kartlegges kontinuerlig gjennom året, særlig de som rapporterer å bruke *Kartleggingsmateriell: Språkkompetanse i grunnleggende norsk*. Noen av skolene sier de formelt registrerer disse resultatene 1-2 ganger i året, men at kartleggingen skjer kontinuerlig gjennom året. En skole rapporterer at de *må* bruke *Kartleggingsmateriell: Språkkompetanse i grunnleggende norsk*, og kan velge å bruke andre som supplement. Her har sannsynligvis skoleledelsen lagt føringer for hvordan kartlegging skal foregå siden andre skoler i samme kommune ikke melder dette, og det finnes skoler i kommunen som ikke bruker dette kartleggingsverktøyet.

Tabell 3	Måling av frekvens av kartlegging av språkferdigheter i innføringstilbud (N = 17)	
Frekvens av kartlegging av språkferdigheter	En gang i året	3
	To ganger i året	4
	Tre ganger i året	3
	> Tre ganger i året	7
	Totalt	17

I tabell 4 kan vi se at 16 kartleggingsprøver brukes for å nivådele i grupper, 21 kartleggingsprøver brukes for å tilpasse opplæringen, 7 kartleggingsprøver brukes for å avdekke vansker på andrespråket, og 11 kartleggingsprøver brukes til andre formål, som å kartlegge vansker på førstespråket og som dokumentasjon på opplæring eller til overføring til ny klasse/skole. Flere skoler oppgir å bruke prøvene som dokumentasjon på opplæring og/eller som grunnlag for å gi elever rett til særskilt språkopplæring, jf. §2-8 eller §3-12 i Opplæringslova. En skole oppgir at elever som etter kartlegging viser seg å være svake i forhold til klassen de ordinært er plassert i, får tilbud om å være 'hjelpelærer' for yngre elever som en del av undervisningen i noen timer i uken, og at dette er et tilbud de ser virker godt for disse elevene.

Tabell 4

Antall innføringstilbud som bruker informasjon fra resultatene av kartleggingsprøvene for

	Å plassere elevene i nivådelte grupper	Å tilpasse opplæringen til den enkelte elev	Å finne ut om eleven har vansker på andrespråket	Annet
Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk (Udir)	4	7	3	4
Den europeiske språkpermen*				
FLORO			1	2
TOSP			1	
Migranorsk	3	2		
Kartleggingsverktøy for minoritetsspråklige elever i VGS (Vestlandsverktøyet)	4	2		
Norskprøven for voksne**	2	1		
Kartleggingsprøve for flerspråklige elever (Trondheimstesten)***	1	3	1	
Lokalt utviklet prøve	1	3		1
Andre prøver	1	3	1	4
Totalt	16	21	7	11

*Aldersdifferensiert 6-12 og 13-18 år

**Både gammel og ny versjon

***Nivådifferensiert: 5-6 år, 3.-4.trinn, 5.-7.trinn

I spørreundersøkelsen skilles det ikke mellom *Norskprøven for voksne A1-B2* og *Norskprøve 2 (og 3) for voksne*. I oppfølgingssamtaler med skolene ble det klart at begge deler ble brukt.

4.2. Vurdering av validitet og reliabilitet i instrumentene

Tabell 5 inneholder en generell oversikt over måleinstrumentene og utgivers beskrivelse av instrumentet.

Måleinstrumenter og tilhørende publikasjoner inkludert i the European Federation og Psychologists' Associations review					
Instrument (siste versjon)	Forfatter	Type publikasjon	Publisert år	Alders-gruppe	Utgivers beskrivelse av instrumentet
Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk (2008)	Utdanningsdirektoratet	Digital nedlastning (pdf)	2008	1.-13. trinn	Et hjelpemiddel for læreren for å kartlegge elevens kompetanse og vurdere om læringsmål er nådd. Det kan også fungere som hjelp for læreren når det gjelder progresjon og tilpassing av opplæringen.
Den europeiske språkpermen 6-12 (2006)	Torill Simonsen og Heike Speitz	Digital nedlastning (pdf)	2006	6-12 år	Et hjelpemiddel for å lære språk gjennom å stimulere til refleksjon om egen læring, refleksjon omkring interkulturelle aspekter, utvikling og bevisstgjøring av flerspråklighet, elevautonomi og livslang læring av språk og dokumentere språkkompetanse etter en felles mal i alle europeiske land
Den europeiske språkpermen 13-18 (2006)	Torill Simonsen og Heike Speitz	Digital nedlastning (pdf)	2006	13-18 år	Et hjelpemiddel for å lære språk gjennom å stimulere til refleksjon om egen læring, refleksjon omkring interkulturelle aspekter, utvikling og bevisstgjøring av flerspråklighet, elevautonomi og livslang læring av språk og dokumentere språkkompetanse etter en felles mal i alle europeiske land
FLORO (2013)	Utdanningsdirektoratet og NAFO	Papirformat + CD	2013	1.-10. trinn	Inneholder tre kognitive delprøver: Ordminne, RAN og Ordrepetisjon. Prøvene skal bidra til å finne elever med språk-, evt. spesifikke, lese-/skrivevansker. Språk og lærevansker er kompliserte og sammensatte fenomen. Denne testen kartlegger flere sider ved elevens kognitive utvikling, og kan tas på flere språk
TOSP (2008)	Sunil Loona	Papirformat	2008	1.-10. trinn	Et hjelpemiddel for å vurdere elevens ferdigheter på tre områder i dens språklige- og kognitive utvikling: Begrepsforståelse, ordforråd og leseferdighet. Det er en tospråklig prøve som gjør at eleven kan testes både på norsk og sitt morsmål for å vurdere elevens "helhetlige verbale repertoar"
Migranorsk (2020)	Fagbokforlaget	Digital nettbasert prøve	2020	Min.spr. barn, ungdom og voksne	Adaptiv språktest i norsk som andrespråk som tester lesing, lytting, grammatikk og skriftlig produksjon.
Kartleggingsverktøy for minoritetsspråklige elever i VGS, aka. Vestlandsverktøyet (2020)	Vestland fylkeskommune og Høgskulen på Vestlandet	Papirformat	2020	VGS	En skriftlig prøve og en muntlig prøve som skal kartlegge om eleven har rett på særskilt språkopplæring etter §3-12 i Opplæringslova.
Norskprøve 2 for voksne (2011)	Folkeuniversitetet, Norsk språktest	Papirformat + CD	2011	>18 år	Kartleggingsprøve for å kartlegge lytte-, lese- og skriveferdigheter på norsk. Gir språknivå.

Norskprøve 3 for voksne (2011)	Folkeuniversitetet, Norsk språkstest	Papirformat + CD	2011	>18 år	Kartleggingsprøve for å kartlegge lytte-, lese- og skriveferdigheter på norsk. Gir språknivå.
Norskprøven for voksne, nivå A1-B2 (2021)	Kompetanse Norge	Digital nettbasert prøve	2021	16-67	Kartleggingsprøve for å kartlegge lytte-, lese- og skriveferdigheter på norsk. Gir språknivå.
Kartleggingsprøve for flerspråklige elever, aka. Trondheimstesten (2013)	Berit Brubakk, Sissel de la Cruz og Mary Landsen (Trondheim kommune)	Papirformat	2013	5-6 år	Kartleggingsmateriale for å kartlegge elevens læringsutbytte i mottaksgruppe. Måler muntlige ferdigheter: Tale/kommunikasjon.
Kartleggingsprøve for flerspråklige elever, aka. Trondheimstesten (2013)	Berit Brubakk, Sissel de la Cruz og Mary Landsen (Trondheim kommune)	Papirformat	2013	3.-4. trinn	Kartleggingsmateriale for å kartlegge elevens læringsutbytte i mottaksgruppe. Måler muntlige og skriftlige ferdigheter: Tale/kommunikasjon, skriving, lesing og lytting.
Kartleggingsprøve for flerspråklige elever, aka. Trondheimstesten (2013)	Berit Brubakk, Sissel de la Cruz og Mary Landsen (Trondheim kommune)	Papirformat	2013	5.-7. trinn	Kartleggingsmateriale for å kartlegge elevens læringsutbytte i mottaksgruppe. Måler muntlige og skriftlige ferdigheter: Tale/kommunikasjon, skriving, lesing og lytting.

Tabell 6 viser en oversikt over karakteristikkene til kartleggingsprøvene. Den eldste kartleggingsprøven er fra 2006 og de nyeste fra 2020, så det er stor spredning i publikasjonene. De fleste prøvene kan administreres i gruppe, mens tre må administreres individuelt (*Trondheimstesten 5-6 år*, *TOSP*, *FLORO*) og to av dem har en del som kan administreres som gruppe og en del som må administreres individuelt (*Trondheimstesten 3.-4.trinn* og *5.-7.trinn*). Alle prøvene med unntak av *Migranorsk* og *Norskprøven*, er papir og blyant-prøver. De fleste av disse er lærerskårt, to skåres av datasystemer (*Norskprøven lytte & lese* og deler av *Migranorsk*) og to er elevskårt (*Den europeiske språkpermen 6-12 år* og *13-18 år*). De fleste prøver måler lytte-, lese- og skriveferdigheter. Noen måler også muntlig kommunikasjon, grammatikk eller ordminne. I de fleste prøvene er det brukt flervalgsspørsmål eller åpne svar, men noen prøver bruker også likert-skalaer. Bare tre prøver er normert, *FLORO*, *TOSP* og *Norskprøven*. *FLORO*s normering er gjort på de tilgjengelige elevene på tilgjengelige skoler i forsøket. Den norske versjonen av *FLORO Ordminneprøve* er normert på 245 elever fordelt på 1.trinn-VGS (vedlegg 5), *FLORO RANprøve* har ikke oppgitt antall for normering, og *FLORO Ordrepetisjonsprøve* er normert med 30 elever på 2.-6.trinn og 20 elever på VGS (vedlegg 5). Normeringen i *TOSP* er basert på alle elevene på 3., 5. og 7.trinn ved fem skoler i Skien kommune (vedlegg 6). Normeringen i *Norskprøven* er koblet til nivåbeskrivelsene i *Rammeverket*, og de ferdighetene som er beskrevet på de ulike nivåene der, altså en kriteriebasert normering (vedlegg 7).

Tabell 6

Beskrivelse og karakteristikk for måleinstrumenter for språkkartlegging av minoritetsspråklige

Instrument (revisjonsår)	Formål med instrumentet	Type måling	Aldersgruppe	Responsmodus	Antall subskalaer	Antall testledd (parenteser skaler nr)	Skåring	Testledd-format	Administrasjonstid	Test relatert kvalifikasjoner nødvendig	Antall studier	Studiestørrelse
Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk, Udir (2007)	Måler ferdigheter i lytting, lesing, muntlig produksjon, skriving, språklæring og språk og kultur	Gruppe; Individuell	1.-13.trinn	Direkte observasjon	6	21 (1); 39 (2); 32 (3); 30 (4); 19 (5); 18 (6)	Lærerskårt	Likert (3)	Ingen info	Nei	0	N/A
Den europeiske språkpermen 6-12 (2006)	Måler ferdigheter i lytteforståelse, leseforståelse, skriving og muntlig produksjon og muntlig samhandling	Gruppe; Individuell	6-12 år	Papir og blyant; Direkte observasjon	5	22 (1); 24 (2); 26 (3); 24 (4); 25 (5)	Annet: Eleven skårer selv (evt. lærer)	Likert (3); Ja/Nei	Ingen info	Nei	0	N/A
Den europeiske språkpermen 13-18 (2006)	Måler ferdigheter i lytteforståelse, leseforståelse, skriving og muntlig produksjon og muntlig samhandling	Gruppe; Individuell	13-18 år	Papir og blyant; Direkte observasjon	5	42 (1); 52 (2); 50 (3); 41 (4); 57 (5)	Annet: Eleven skårer selv (evt. lærer)	Likert (3)	Ingen info	Nei	0	N/A
FLORO (2013)	Måler evne til å lære/huske; kognitive stiler	Individuell	1.-10.trinn	Muntlig intervju	3	8 (1); 6 (2); 28 (3)	Lærerskårt; Rett/feil; ikke-normerte standard-avvik	Diktat; Likert (5); Åpen	Ingen info	Ja	0	N/A

TOSP (2008)	Måler evne til å lære/huske; kognitive stiler; verbale ferdigheter	Individuell	1.-10.trinn	Muntlig intervju	4	30 (1); 52 (2); 40 (3); 75 (4)	Lærerskårt; Rett/feil; Stanine; Normer	Åpen	Ingen info	Ja	1	150-177 (pr normert trinn - 3.trinn, 5.trinn, 7.trinn)
Migranorsk (2020)	Måler ferdigheter i lytting, lesing, grammatikk og skriving	Gruppe; Individuell	Min.spr. barn, ungdom og voksne	Digital nettbasert prøve	4	40 (1); 40 (2); 40 (3); 1 (4)	Digital rapport (lytte, lese, grammatikk); Rett/Feil; Lærerskårt (skriving)	Flervalg; Åpen	Testtid: 120 min; Administrasjonstid: Ingen info	Nei	Ingen info	N/A
Kartleggingsverktøy for minoritetspråklige elever i VGS, aka. Vestlandsverktøyet (2020)	Måler muntlig og skriftlig produksjon	Gruppe; Individuell	VGS	Papir og blyant; Muntlig intervju	2	4 (1); 3 (2)	Lærerskårt	Åpen	Ingen info	Ingen info	0	N/A
Norskprøven 2 (2011)	Måler ferdigheter i lytting, lesing og skriving	Gruppe; Individuell	>18 år	Papir og blyant	3	20 (1); 23 (2); 2 (3)	Lærerskårt; Rett/feil	Flervalg; Åpen	Testtid: 125 min; Administrasjonstid: Ingen info	Ja	Ingen info	N/A
Norskprøven 3 (2011)	Måler ferdigheter i lytting, lesing og skriving	Gruppe; Individuell	>18 år	Papir og blyant	3	23 (1); 25 (2); 2 (3)	Lærerskårt; Rett/feil	Flervalg; Åpen	Testtid: 125 min; Administrasjonstid: Ingen info	Ja	Ingen info	N/A
Norskprøven for voksne, lytte og lese (2021)	Måler ferdigheter i lytting og lesing	Gruppe; Individuell	16-67	Digital nettbasert prøve	12	7 (1); 7 (2); 8 (3); 16 (4); 22 (5); 23 (6); 7 (7); 7 (8); 7 (9); 16 (10); 20 (11); 17 (12)	Digital rapport; Rett/Feil	Flervalg; Sett i rekkefølge	Testtid: 135 min; Administrasjonstid: Ingen info	Ja	5	>1000

Norskprøven for voksne, skrive (2021)	Måler ferdigheter i skriving	Gruppe; Individuell	16-67	Digital nettbasert prøve	3	3 (1); 3 (2); 2(3)	Digital rapport; Lærerskårt	Åpen	Testtid: 90-120 min; Administrasjonstid: Ingen info	Ja	4	<100
Kartleggingsprøve for flerspråklige elever, aka. Trondheimstesten (2013)	Måler muntlige ferdigheter	Individuell (M)	5-6 år	Muntlig intervju	1	26 (1)	Lærerskpr; Rett/feil	Flervalg; Likert; Åpen	Ingen info	Nei	0	N/A
Kartleggingsprøve for flerspråklige elever, aka. Trondheimstesten (2013)	Måler muntlige ferdigheter og ferdigheter i lesing, lytting og skriving	Gruppe/Individuell (PB); Individuell (M)	3.-4.trinn	Papir og blyant; Muntlig intervju	2	26 (1); 26 (2)	Lærerskpr; Rett/feil	Flervalg; Likert; Åpen	Ingen info	Nei	0	N/A
Kartleggingsprøve for flerspråklige elever, aka. Trondheimstesten (2013)	Måler muntlige ferdigheter og ferdigheter i lesing, lytting og skriving	Gruppe/individuell (PB); Individuell (MI)	5.-7.trinn	Papir og blyant; Muntlig intervju	2	16 (1); 16 (2)	Lærerskårt; Rett/feil	Flervalg; Likert; Åpen	Ingen info	Nei	0	N/A

NOTAT: DNP = Digital nettbasert prøve; DO = Direkte observasjon; FV = Flervalg; Ingen info = Ingen informasjon å vurdere; OS = Observasjonsskjema; MI = Muntlig intervju; PB = Papir og blyant; R/F = Rett/Feil; LS = Lærerskårt

I tabell 7 kan vi se informasjonen fra *EFPAs Oversikt over materialkvalitet og dokumenterte psykometriske egenskaper for målingsinstrumentet for språkkartlegging av minoritetsspråklige elever*.

Først ble kvaliteten i de logiske begrunnelsene, forklaringene, presentasjonen og informasjonen vurdert. To prøver skårer inadequate (*Kartleggingsmateriell: Språkkompetanse i grunnleggende norsk, FLORO*), en ble vurdert til adequate (*TOSP*) og to prøver ble vurdert til good (*Norskprøven for voksne⁵ lytte & lese, Norskprøven for voksne skrive*). De resterende prøvene hadde ingen eller for lite informasjon til å bli vurdert.

Så ble kvaliteten på testmaterialet vurdert. Bare én ble vurdert til excellent (*FLORO*), mens fem ble vurdert til good (*TOSP, Norskprøve 2, Norskprøve 3, Norskprøven for voksne lytte & lese, Norskprøven for voksne skrive*), seks adequate (*Kartleggingsmateriell: Språkkompetanse i grunnleggende norsk, Den europeiske språkpermen, Vestlandsverktøyet, Trondheimstesten⁶*) og en inadequate (*Migranorsk*). Datagenererte rapporter finner man i både *Migranorsk* og *Norskprøven*, og de skårer adequate (*Migranorsk*) og excellent (*Norskprøven*).

Til slutt ble instrumentets dokumenterte psykometriske egenskaper vurdert i form av normer, reliabilitet og validitet. Av de 11 kartleggingsprøvene er det to som er good på normering, reliabilitet og validitet (*Norskprøven for voksne lytte & lese, Norskprøven for voksne skrive*), en som er adequate på normering og reliabilitet (*TOSP*), en er inadequate på normering og reliabilitet (*FLORO*), mens resten av kartleggingsprøvene ikke har informasjon om normering, reliabilitet og validitet. Fagbokforlaget ville ikke utlevere informasjon om validitets- og reliabilitetsanalyser av sitt kartleggingsverktøy *Migranorsk*, til tross for flere henvendelser, og de begrunner dette med 'kommersielle hensyn'.

⁵ I vurderingen av testen er denne delt inn i lytte & lese og skrive. Dette har jeg gjort fordi de bruker ulike analysemodeller både for reliabilitet og validitet. I skriveprøven er det hovedsakelig sensorsvurderingene som analyseres.

⁶ 3 tester for 3 ulike trinn

Tabell 7

Oversikt over materialkvalitet og dokumenterte psykometriske egenskaper for målingsinstrumentet for språkkarlegging av minoritetsspråklige elever

Instrument (revisjonsår)	Kvalitet i logisk begrunnelse, forklaring, presentasjon og informasjon	Kvalitet av testmaterialet			Norm	Reliabilitet	Validitet
		Papir og blyant	Digital (lokal eller web)	Kvalitets- rapporter generert av datasystem			
Kartleggingsmaterieill: Språkkompetanse i grunnleggende norsk (2008)	Inadequate	Adequate	N/A	N/A	Ingen info	Ingen info	Ingen info
Den europeiske språkpermen 6-12 (2006)	Ingen info	Adequate	N/A	N/A	Ingen info	Ingen info	Ingen info
Den europeiske språkpermen 13-18 (2006)	Ingen info	Adequate	N/A	N/A	Ingen info	Ingen info	Ingen info
FLORO (2013)	Inadequate	Excellent	N/A	N/A	Inadequate	Inadequate	Ingen info
TOSP (2008)	Adequate	Good	N/A	N/A	Adequate	Adequate	Ingen info
Migranorsk (2020)	Ingen info	N/A	Inadequate	Adequate	Ingen info	Ingen info	Ingen info
Kartleggingsverktøy for minoritetsspråklige elever i VGS, aka. Vestlandsverktøyet (2020)	Ingen info	Adequate	N/A	N/A	Inadequate	Ingen info	Ingen info
Norskprøve 2 for voksne (2011)	Ingen info	Good	N/A	N/A	Ingen info	Ingen info	Ingen info
Norskprøve 3 for voksne (2011)	Ingen info	Good	N/A	N/A	Ingen info	Ingen info	Ingen info
Norskprøven for voksne, lytte og lese (2021)	Good	N/A	Good	Excellent	Good	Good	Good
Norskprøven for voksne, skrive (2021)	Good	N/A	Good	Excellent	Good	Good	Good
Kartleggingsprøve for flerspråklige elever 5- 6 år, aka. Trondheimstesten (2013)	Ingen info	Adequate	N/A	N/A	Ingen info	Ingen info	Ingen info
Kartleggingsprøve for flerspråklige elever 3.-4.trinn, aka. Trondheimstesten (2013)	Ingen info	Adequate	N/A	N/A	Ingen info	Ingen info	Ingen info
Kartleggingsprøve for flerspråklige elever 5.-7.trinn, aka. Trondheimstesten (2013)	Ingen info	Adequate	N/A	N/A	Ingen info	Ingen info	Ingen info

N/A = Not Applicable; Ingen info = Ingen eller for lite informasjon å vurdere

5. Drøfting

I denne masteroppgaven har jeg undersøkt hvilke kartleggingsprøver som brukes for å kartlegge norskspråklige ferdigheter i innføringstilbud i tidligere Hordaland fylke, og hvilken reliabilitet og validitet disse kartleggingsprøvene (måleinstrumentene) har. I drøftingen ønsker jeg å ta opp hvilke kartleggingsprøver som brukes, hvor ofte de bruker prøvene og hva prøvene brukes til. Videre tar jeg opp reliabilitet og validitet i måleinstrumentene og mulige konsekvenser resultatene i denne masteroppgaven har for praksisfeltet.

5.1. Hvilke kartleggingsprøver brukes, hvor ofte og til hvilket formål?

Det brukes 14 kartleggingsprøver i innføringstilbud i tidligere Hordaland fylke, pluss sju lokalt utviklede prøver. Lokalt utviklede prøver er ikke tatt med i analysen av reliabilitet og validitet, men de er med i undersøkelsen om hva kartleggingsprøvene brukes til og hvor ofte. De 14 kartleggingsprøvene som er en del av analysen er som følger: *Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk, Den europeiske språkpermen* (i to versjoner: 6-12 år og 13-18 år), *FLORO, TOSP, Migranorsk, Kartleggingsverktøy for minoritetsspråklige elever i VGS (Vestlandsverktøyet), Norskprøve 2 for voksne, Norskprøve 3 for voksne, Norskprøven for voksne A1-B2* (i to versjoner: lytte og lese, og skrive) og *Trondheimstesten* (i tre versjoner: 5-6 år, 3.-4.trinn og 5.-7.trinn).

Alle skolene svarer at de kartlegger elevene minst én gang hvert skoleår, og flere sier de kartlegger to, tre og fire ganger i året. Et par skoler oppgir også at de bruker *Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk* til kontinuerlig kartlegging gjennom året, men at ulike deler av kartleggingsmateriellet da brukes på ulike tidspunkt, ofte knyttet til tema/sjanger som det jobbes med i undervisningen.

Dersom elever bare kartlegges en gang i året, og denne prøven brukes til å plassere eleven i nivådelte grupper, så vil denne eleven sannsynligvis være i denne gruppen resten av skoleåret. Er det hensiktsmessig for elevens læring? Et hovedargument for å nivådele elever er for å gi bedre tilpasset opplæring slik at de kan lære mer. Da blir det et paradoks om man tenker at alle i en slik gruppe har lik læringsutvikling resten av skoleåret.

Disse kartleggingsprøvene brukes i stor grad til å tilpasse opplæringen, dele elever inn i nivådelte grupper og avdekke vansker på andrespråket. Som jeg trekker frem i 2.2.5. så er ikke 'vansker på andrespråket' nødvendigvis språk- eller lærevansker, men at eleven har svake ferdigheter på andrespråket, altså det språket skolen bruker, og at man da bruker kartleggingsprøver for å finne ut hva som er vanskelig for eleven. Men hva er det man

kartlegger for å finne ut dette? Flere prøver ser på forventet eller anbefalt ordforråd for ulike trinn uten at det spesifiseres kriterier for hvilket ordforråd elever bør ha på de ulike trinnene. Hvordan har man kommet frem til at det er disse ordene elevene må mestre på disse trinnene? Er det hentet fra ordbruk i lærebøker, dagligtale eller språk læreren tenker eleven har bruk for? Her burde man kanskje gått inn for en normering av ordforrådet den generelle populasjonen har på de ulike trinnene, og tatt utgangspunkt i dette.

En skole oppgir at elever som etter kartlegging har svakere resultater enn andre i samme gruppe, får tilbud om å være 'hjelpelærer' for yngre elever som en del av undervisningen noen timer i uken. De melder selv at dette fungerer godt som tiltak for elevene, og gir dem både motivasjon, mestringsfølelse og ekstra læring.

Mange prøver ser på bruk av sjangere, forståelse (i både lytting, lesing og samtale) og grammatiske strukturer og kobler dette mot *Rammeverket*. Da ser man dette som noe som kommer *før* ordinær læreplan i norsk, og ikke som en del av det eleven ordinært skal lære. Flere av kartleggingsprøvene brukes også som utgangspunkt for å avgjøre om eleven har rett til særskilt språkopplæring, jf. §2-8 eller §3-12 i Opplæringslova. Når prøvene brukes til dette formålet, så er det snakk om 'high stakes' tester, altså tester som har store konsekvenser for elevens videre opplæring. Det samme kan man gjerne si om prøver som brukes for å plassere elever i nivådelte grupper. Prøvene blir 'high stakes' fordi det har stor påvirkning på elevens opplæring. Om prøvene i tillegg bare brukes til dette én gang i året, så blir denne konsekvensen enda større fordi eleven da plasseres i et opplæringsløp for hele skoleåret uten mulighet for endring, selv om elevens forutsetninger og ferdigheter kan endre seg i større eller mindre grad gjennom året. Så kan man gjerne si at en erfaren lærer vil kunne se dette underveis uten nye kartlegginger, men ser man på forskning så vet man at lærers innsikt i og fortolkning av elevs ferdigheter er begrenset (Antoniazzi, Snow & Dickson-Swift, 2010; Südkamp, Kaiser & Möller, 2012).

5.2. Hvilken validitet og reliabilitet har måleinstrumentene?

Norskprøven for voksne A1-B2 er den eneste prøven som informantene i mitt utvalg nevner, som i utstrakt grad er undersøkt for reliabilitet og validitet. I lytteprøven får man høre teksten to ganger, med unntak av tekstene på B2 nivå, de får man bare høre én gang. Det er likevel viktig å merke seg at svaralternativene med tekst ikke leses opp for innlæreren. Flere av spørsmålene på A1-A2 er at man skal trykke på element i et bilde eller en illustrasjon, men de resterende oppgavene krever at man kan lese og forstå alternativene man kan velge mellom. Dette er en svakhet ved lytteprøven fordi den da også krever andre kompetanser enn den

ferdigheten den sier den måler: lytteforståelse, og svekker på den måten begrepsvaliditeten til lyttetesten. Vi ser ikke det samme i lesetesten fordi både oppgaver og svar krever leseferdigheter, som da gjør at begrepsvaliditeten i lesetesten er god. Reliabiliteten og validiteten i lytte- og lesetesten undersøkes med IRT-analyser og skrivetesten undersøkes med MFRM-analyser.

TOSP er normert og *FLORO* er delvis normert, men ingen av disse er testet for reliabilitet og validitet. *TOSP* bruker gjennomsnitt, standardavvik og stanine-skårer med fem prøveklasser i sin normering (vedlegg 6). Normeringen er basert på alle elever på 3.trinn, 5.trinn og 7.trinn fordelt på fem grunnskoler i Skien kommune, og så har de overført normeringsverdiene fra norsk til de andre språkene (Loona, 2008a). Dette begrunner de med at elever som har tatt testen på to språk, i stor grad viser de samme variasjonene i begge språk. Det blir ikke presentert dokumentasjon på dette i materialet, som da blir en svakhet ved normeringen for andre språk enn norsk. Det finnes også en normering for 9.trinn i instruksjonsheftet, men det står ingenting om hvordan denne normeringen er gjennomført. *FLORO*s veiledende normering er først prøvd ut på elever med norsk som morsmål før man har testet det ut på andre språk (Utdanningsdirektoratet & NAFO, 2013b). Normeringen er ikke standardisert, men basert på den tilgjengelige populasjonen på tilgjengelige språk og skoler, og på den måten blir det ikke basert på et generelt populasjonsgrunnlag. Testutviklerne hevder likevel at normeringen har en viss representativhet fordi skolene har rekruttert elever fra ulike bomiljøer og familier med ulik sosial bakgrunn (Utdanningsdirektoratet & NAFO, 2013b). Selv om det har deltatt elever fra større steder som Bergen, Drammen, Trondheim, Stavanger og Bærum, så er det varierende hvilke sosioøkonomiske bakgrunner de ulike deltakerne kommer fra (Utdanningsdirektoratet & NAFO, 2013b). I begge disse prøvene kan man stille spørsmål ved om det er nødvendig å normere på andre språk enn norsk. Normeringen av *FLORO* på andre språk blir veldig begrenset. Noen av språkene har så få deltakere, at når man fordeler dem ut på de ulike trinnene, så blir den ytre validiteten lav (vedlegg 5). Når man for thai har 23 elever fordelt på 3 trinn, eller 31 arabisktalende elever på tre ulike trinn, så blir det vanskelig å fastslå hvor representative resultatene av normeringen er.

Målet med *TOSP* er å kartlegge begrepsforståelse, ordforråd og leseferdighet, og prøven brukes ofte på elever som strever på andrespråket, for å se om eleven har de samme vanskene på førstespråket. Dette kan være interessant for å kartlegge generelle ferdigheter i norsk, men det er likevel uklart hva som er viktig i normering av testen på andre språk enn norsk.

*FLORO*s formål er å måle ordminne, RAN og ordrepetisjon. Dette er gjerne ment for å

fokusere på andre ting ved språk- og leseferdighet enn forståelse. I så tilfelle blir formålet med testen noe annet enn å kartlegge norskferdigheter, slik de andre testene i denne undersøkelsen. Dersom dette er hovedmålet for prøven, så trenger man gjerne ikke normering på førstespråkene. Hvis man tenker at normering på alle språk er viktig, så er det flere spørsmål å ta hensyn til: Skal prøven normeres for språkets populasjon i Norge? For generell populasjon i sitt hjemland? Vil en normering i andre land være gyldig i Norge? Uansett hvilket utgangspunkt man velger i normeringsarbeidet, er det viktig å ta høyde for konteksten til normeringen og begrensninger denne kan føre med seg.

Det er overraskende at *Trondheimstesten, Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk* og *Den europeiske språkpermen* ikke er normert eller undersøkt for reliabilitet og validitet, særlig siden de to første er de mest brukte kartleggingsverktøyene i innføringsklasser i tidligere Hordaland fylke. *Trondheimstesten* fokuserer på begreper og språk man tenker er nødvendig for bruk på ulike trinn. Å ha prøver som forteller om eleven har et tilstrekkelig språk for å klare seg i et opplæringsløp er en veldig god tanke, men arbeidet må ikke avsluttes når prøven er ferdiglaget. Dette er et godt utgangspunkt for videre arbeid med prøven for å få på plass normeringer og deretter undersøke for reliabilitet og validitet. Det er viktig å sjekke at det man tenker at man trenger av begreper og språk på et nivå er det eleven faktisk trenger i praksis.

I den skriftlige delen av *Trondheimstesten* må eleven mestre både lesing og skriving for å kunne svare på oppgavene. Dette er et utgangspunkt som kan svekke validiteten i prøven fordi det krever at eleven både skal kunne lese og forstå oppgavene og skrive svar på dem, slik Weir (2005) peker på. I både lyttedelen og lesedelen av *Norskprøve 2 og 3* kreves det både at eleven skal krysse av på flervalgsspørsmål med tre svaralternativer og skrive svar på åpne spørsmål. Som Hargreaved (2010) trekker frem, så vil tre svaralternativer svekke reliabiliteten i en test fordi det er så stor sjanse for å gjette seg frem til riktig svar. Hun mener at flervalgsspørsmål må ha fire eller fem svaralternativer for at reliabiliteten skal være god. I tillegg er det uheldig at det er åpne svar på disse prøvene, fordi skriftlige svar måler en annen ferdighet enn den prøven måler, som igjen vil svekke den ytre validiteten i prøvene.

Selv om *Vestlandsverktøyet* ikke er undersøkt for reliabilitet og validitet pr. nå, er det planer om å gjøre reliabilitetsvurderinger for sensorvurderingene på et utvalg lærere i *Vestlandsverktøyet* (Carlsen, 2021, e-postutveksling), men det er ikke avklart når dette skal skje. Samtidig er dette et verktøy som er ment at alle lærere selv skal kunne gjennomføre på elever ved sin skole, og når prøven er papirbasert er det ingen kontroll på hvordan elevene

vurderes i praksis, selv om reliabilitetsvurderinger gjøres på et utvalg lærere. Om reliabilitetsvurderinger av sensorvurderinger skal gjennomføres, så bør dette gjøres kontinuerlig med alle lærere som bruker prøven og deres vurderinger. Verktøyet sier heller ingenting om det er krav til opplæring før bruk. Når prøveresultatet er en tolkningsvurdering fra en lærer, så gjør det at det blir enda viktigere å sikre at alle som vurderer tolker vurderingskriteriene og nivåbeskrivelsene på samme måte. Dette er jo også hovedargumentet for sensoropplæring på sentralgitte eksamener: Man skal sikre at elever vurderes så likt som mulig, uavhengig av hvem sensoren er.

Sett samlet kan en årsak til de nedslående resultatene handle om at dette er en veldig sammensatt elevgruppe, som gjør det utfordrende å lage kriterier til. Det gjør at det blir utfordrende å lage tester som tar opp i seg alle de ulike aspektene og utfordringene som denne elevgruppen er en del av. Man vil gjerne også tenke at det er ulikt hvordan de bruker morsmål og norsk ut fra både bosted og familieforhold. Er de den eneste personen eller familien som snakker språket på et sted er morsmål gjerne i mindre bruk enn hos andre som bor steder med mange andre som snakker samme morsmål. Dette gir også ulike utgangspunkter for andrespråklæring (Berggreen & Tenfjord (2007)). Hva skal være populasjonsgrunnlaget for en normering av slike tester? Hvordan kan man klare å lage gode normeringer for denne populasjonen? Mange av testene er systematisk lagt opp, og de fleste testene inneholder fokusområdene lytte, lese, snakke og skrive. Dette er gode utgangspunkter for testmaterieell, men man må samtidig være klar over hvilke begrensninger som ligger i testmateriellet man velger å bruke. Vi må også huske, som Carlsen og Moe (2019) hevder, viktigheten av at testutviklere kvalitetssikrer testene de utvikler. Arbeidet er ikke ferdig når en prøve er ferdiglaget. Man må fortsette å analysere reliabilitet og validitet i testmateriellet for å sikre at prøven er god og passer til formålet den er tenkt til.

5.3. Betydningen av få gode kartleggingsprøver

Denne studien avdekker at det er veldig få kartleggingsmaterieell som finnes på markedet som er testet for reliabilitet og validitet. Når det bare er én prøve som er undersøkt for reliabilitet og validitet (*Norskprøven*), én som er normert (*TOSP*) og én som hevder å ha en veiledende normering (*FLORO*), så betyr det at man fort kan stille spørsmål ved hvordan man da skal vite at prøvene måler det de sier de måler. Og risikerer man da at elever mister rettigheter de ellers skulle hatt fordi man ikke vet hva prøvene måler? Eller at elever som måles med en prøve får en rettighet som elever som måles med en annen prøve ikke får? Eller at en elev som har en sensor får en rettighet som en elev med en annen sensor ikke får? Man vet ikke om prøven

måler det den sier den måler uten at reliabilitet og validitet blir undersøkt. Hva skal man med kartleggingsverktøy i skolen som man ikke vet hva måler? Skolene sier de kartlegger for å se om elever har rett til særskilt språkopplæring, jf. §2-8 eller §3-12, for å nivådele elever, tilpasse opplæringen til den enkelte og for å avdekke vansker på andrespråket og evt. førstespråket. Hvis det er dette skolene trenger kartleggingsprøver til, så må de være sikre på at kartleggingsprøvene måler elementer som gjør dette mulig. Skal prøven brukes for å gi særskilt språkopplæring, så må prøven måle det som lovteksten definerer som 'tilstrekkelige ferdigheter i norsk'. Og når dette er et udefinert begrep, så blir det omtrent umulig å lage en slik test. Uansett hva målet med testen er, så er det viktig å velge en prøve som måler det man trenger. Trenger man ny skibukse, så går man ikke i butikken og ser på undertøy. De plaggene har ulike formål. På samme måte har ulike prøver ulike formål og måler ulike ferdigheter, og dersom dette ikke undersøkes i prøvene gjennom reliabilitets- og validitetsanalyser, så er det sannsynlig at det blir mange tilfeldige prøveresultater. Det er viktig at skoler og lærere vet hvilke prøver som måler hva slik at de kan gjøre gode valg når de velger kartleggingsverktøy til sine formål.

Det er ganske oppsiktsvekkende at de verktøyene som er mest brukt i skoleverket, *Kartleggingsmateriell: Språkkompetanse i grunnleggende norsk og Trondheimstesten*, ikke er normert eller testet for reliabilitet og validitet. Når prøvene ikke normeres, ikke prøves for reliabilitet og validitet, så vet vi ikke om de er gode eller dårlige, for det er ingen som har undersøkt dette. Min analyse av verktøyene forteller at det er et stort behov for slikt forskningsarbeid knyttet til de eksisterende prøvene slik at skolene reelt kan velge de prøvene som er gode på de områdene skolen ønsker å undersøke. Mange av prøvene ser ut til å måle ferdigheter som lesing, skriving og kommunikasjon, som er viktige ferdigheter i skolen. Men det er likevel forskjeller på prøvene. Hvordan vet vi hvilke prøver som egner seg best til ulike formål? Mange skoler bruker flere kartleggingsverktøy, og prøver som *TOSP* og *FLORO* nevnes stort sett for å se om eleven har vansker på første- eller andrespråket. Både *Kartleggingsmateriell: Språkkompetanse i grunnleggende norsk, Trondheimstesten, Migranorsk, Vestlandsverktøyet* og *Norskprøven* brukes både for å nivådele grupper og tilpasse opplæringen til den enkelte eleven. Flere av disse brukes også som grunnlag for å gi elever rett til særskilt språkopplæring, jf. §2-8 og §3-12 i Opplæringslova. Alle disse prøvene fokuserer på hva eleven kan, og alle av disse bortsett fra *Trondheimstesten* tar utgangspunkt i nivåbeskrivelsene *Rammeverket*. På den måten er det gjerne et nyttig verktøy for å nivådele eller tilpasse. Man kan lett sette sammen de som kan det samme eller har lignende ferdigheter,

men er verktøy som baserer seg på mestringsbeskrivelser gode å bruke for å gi elever rettigheter i henhold til lovverket? Dette kan man synse mye om, men man kan ikke gi konkrete svar uten at dette undersøkes i de enkelte prøvene. Jeg mener også at informasjon om normering, reliabilitet og validitet må inn som en del av veiledningsmaterialet til prøvene, slik at det blir tydelig for de som skal bruke prøvene hva som ligger til grunn for arbeidet, og slik at dette lett kan etterprøves i forskningsmiljøer. I tillegg vil det kunne gi mer informasjon om hva prøven måler fordi dette nødvendigvis må konkretiseres tydelig, som igjen kan gjøre lærere bedre i stand til å vurdere hvilket kartleggingsverktøy man ønsker å bruke til ulike formål i skolen.

5.4. Hva er tilstrekkelige norskferdigheter?

Som beskrevet i innledningen, er det ifølge Opplæringslovas §2-8 og §3-12 begrepet ‘tilstrekkelige ferdigheter i norsk’, eller mangelen på dette, som skal avgjøre om en elev har rett til særskilt språkopplæring. Det er uenighet i hva som skal defineres som ‘tilstrekkelige ferdigheter i norsk’, og det ser ut til at mye av usikkerheten her er knyttet til at begrepet ikke er avklart eller definert i lovtekst, forskrift eller tilsvarende. I *Læreplan i grunnleggende norsk for språklige minoriteter* står det at når eleven har nådd nivå 3, skal eleven over på ordinær læreplan i norsk. Ser vi dette i sammenheng med særskilt språkopplæring, så er særskilt språkopplæring noe man har rett på frem til man har ‘tilstrekkelige ferdigheter i norsk’ til å følge den ordinære læreplanen i norsk. Det kan da være logisk å konkludere med at nivå 3 i læreplanen tilsvarer ‘tilstrekkelige norskferdigheter’. Dette er likevel ikke en definisjonsavklaring, selv om man kan hevde at det er tydelig hva Utdanningsdirektoratet (2007b, 2009) mener eleven bør kunne før man går over på ordinær læreplan. Det er også viktig å peke på at denne formuleringen forsvant med innføringen av ny læreplan skoleåret 2020/2021. Skoler og skoleeiere står også helt fritt til om de ønsker å bruke *Læreplan i grunnleggende norsk for språklige minoriteter* eller å lage egne lokale læreplaner for denne gruppen (Rambøll Management, 2009).

Dermed blir begrepet ‘tilstrekkelige norskferdigheter’ et diffust begrep som egentlig ikke forteller noe om hva en elev skal kunne for å få gitte rettigheter ifølge Opplæringslova. Kunnskapsdepartementet (2007) mener dette skal vurderes av den enkelte lærer, men samtidig viser både studien til Antoniazzi, Snow og Dickson-Swift (2010) og metastudien til Südkamp, Kaiser og Möller (2012) at det bare er lite og noe samsvar mellom lærers vurdering av elevers ferdigheter og elevens faktiske ferdigheter. Dette blir dermed en stor utfordring når kriteriet for en lovfestet rettighet er noe som hver enkelt lærer skal vurdere individuelt. Forskning viser

at lærere ikke klarer å vurdere elevenes ferdigheter godt nok, og at lærere ofte overvurderer elevens ferdigheter, og da særlig de svake elevene som har størst behov for hjelp og støtte (Südkamp, Kaiser & Möller, 2012). I tillegg har man et udefinert begrep som skal gi elever rettigheter, og dette skal lærere altså vurdere individuelt. For lærere blir det en umulig oppgave, og man står da i fare for å sette lærere som portvoktere for hvilke elever som skal eller ikke skal få rettigheter jf. §2-8 og §3-12 i Opplæringslova. Er det denne rollen man vil at lærerne skal ha? Hvorfor er det ikke objektive kriterier som bestemmer når en elev har rett på særskilt språkopplæring? Hvorfor forventer man at andrespråklærere skal ha kompetanse til både undervisning og vurdering i det daglige undervisningsarbeidet og være eksperter innenfor språkkartlegging og andrespråkutvikling? Svarer disse forventningene til lærernes kompetanse til innholdet i lærerutdanningen og hva forskning viser? Det forventes ikke at lærere i 'ordinære' klasser skal ha spesialistkompetanse, og om elever i disse klassene strever, så henter man inn kompetanse, f.eks. fra spesialister på egen skole, Barne- og ungdomspsykiatrisk poliklinikk (BUP) eller pedagogisk-psykologisk tjeneste (PPT) der man finner spesialister som leger, psykologer, spesialpedagoger og logoped. Hvorfor forventer man at andrespråklærere skal ha *både* sin lærerkompetanse og denne spesialistkompetansen når man ikke forventer dette fra lærere i andre fag? Hvis dette er en kompetanse det er viktig at andrespråklærere har, bør vel dette også tas inn som en del av lærerutdanningen i disse fagene?

5.5. Hva er kriteriene for språkprøvene?

Et annet viktig spørsmål er hva kriteriene for å lage en språkprøve skal være. Hva skal prøven måle? Hva skal prøven brukes til? Hvem skal bestemme hvilke kriterier som skal gjelde? Dersom prøven skal brukes for å gi elever rett til særskilt språkopplæring, jf. §2-8 og §3-12, så er det nesten en umulig oppgave fordi lovteksten bruker et ikke-definert begrep som mål på når denne retten skal tilfalle eleven: 'tilstrekkelige ferdigheter i norsk'. Bruker man prøven for å dele elever inn i grupper eller tilpasse opplæringen til den enkelte? Uansett hva som er formålet til prøven, så må det være klare kriterier som prøven skal lages etter. Testutviklerne må vite både hva prøven skal brukes til og hva som skal måles før de lager prøven. Uten slike kriterier blir det umulig å lage en test. En historieprøve om 2.verdenskrig må nødvendigvis ha spørsmål om 2.verdenskrig. Men det er ikke nok å vite at prøven skal handle om 2.verdenskrig. Det må være flere kriterier for den som lager prøven. Handler den om 2.verdenskrig i Norge, Storbritannia eller Tyskland? Hvordan vektlegges de ulike elementene under krigen, som f.eks. jødeutryddelsen, okkupasjonen, diktaturet, kampene,

motstandsbevegelsen, osv. Slike hensyn må også tas når det er snakk om en språkprøve. Det er ikke nok å si at testen måler lesing eller skriving. Det må være konkrete kriterier tilgjengelig for hvert område som måles. En prøve kan aldri bli bedre enn kriteriene som ligger til grunn for prøveutviklingen (Bachman & Palmer, 1996), og da er det kanskje verre med diffuse kriterier enn manglende reliabilitet og validitet.

Man kan hevde at *Rammeverket* skal være utgangspunktet for en språkprøve fordi nivåbeskrivelsene ikke endrer seg og gir beskrivelser for hva eleven skal kunne for å oppnå ulike språknivåer. Dette kan være et greit utgangspunkt hvis målet med prøven er å fastslå at eleven mestrer x og y. Det er mer problematisk å hevde at dette skal være et utgangspunkt som skal være allmenngyldig til alle typer andrespråkspartlegging. Kartlegger man for å se om eleven har språkferdigheter til å følge et visst opplæringsløp, så må kartleggingsprøven også være lagt opp slik at den måler om eleven mestrer det som trengs for akkurat dette opplæringsløpet. Skal man kartlegge for å se om eleven har rett på særskilt språkopplæring, jf. §2-8 og §3-12, så må man ha kartleggingsprøver som måler det man tenker er 'tilstrekkelige ferdigheter i norsk', og for å få dette til må først begrepet avklares og deretter må andre kriterier på plass som konkretiserer begrepsavklaringen.

Videre kan det også være problematisk når Carlsen og Moe (2019) skriver at nivåbeskrivelsene i *Vestlandsverktøyet* er laget med utgangspunkt i *Rammeverket*, men tilpasset en yngre brukergruppe. Hva betyr egentlig det? Nivåbeskrivelsene i *Rammeverket* er tenkt som statiske beskrivelser av språkferdigheter, helt uavhengig av alder. Kognitive ferdigheter gjør at man ikke kan forvente det samme fra en 10-åring som man kan fra en 25-åring, men betyr det at man står fritt til å endre nivåbeskrivelsene i *Rammeverket*?

Rammeverket har vært et felles prosjekt i Europa, og ledende språkforskere fra mange land har bidratt i arbeidet. Carlsen og Moe (2019) forklarer selv at det er viktig at innholdet i nivåbeskrivelsene ikke endres for å tilpasses innlærergupper fordi man da snakker om nivåer med ulikt innhold, og målet med *Rammeverket* er at alle skal ha den samme forståelsen av språknivåene. Man kan stille spørsmålsteget ved hvorfor testutviklerne da har valgt å utvikle egne beskrivelser for å tilpasse *Rammeverket* til en yngre brukergruppe. Det er også uklart hva det menes med å tilpasse *Rammeverket* til en yngre brukergruppe når nivåbeskrivelsene i *Rammeverket* skal være uavhengig av alder og bare beskrive ferdigheter man mestrer. I den sammenhengen blir det vanskelig å forstå hva testutviklerne har gjort for å tilpasse nivåbeskrivelsene til en 'yngre brukergruppe'. Et annet alternativ hadde gjerne vært å se til nivåbeskrivelsene i *Den europeiske språkpermen*. Der finnes det nivåbeskrivelser både for

aldersgruppen 6-12 år og 13-18 år. Jeg har ikke sett på hvilke endringer eller tilpasninger som evt. er gjort i disse nivåbeskrivelsene.

5.6. Implikasjoner for praksis og mulige konsekvenser

Det har vært utfordrende å få tak i informasjon om reliabilitet og validitet til noen av testene, og noen vil heller ikke utgi denne informasjonen, og begrunner det med 'kommersielle hensyn'. Slik jeg ser det, vil det heller være kommersielle hensyn som argumenterer for at denne informasjonen *skal* være tilgjengelig sammen med prøvematerialet fordi det gir sterke argumenter for kvaliteten til prøven. Et eksempel på dette er evnetesten WISC-IV, der det gis informasjon om både normering, reliabilitet og validitet i veiledningsmaterialet til testen (Wechsler, 2003). På den måten kan resultatene etterprøves. Denne prøven brukes til et annet formål enn andrespråkstester, men det illustrerer at det ikke er problematisk å tilgjengeliggjøre informasjon om normering, reliabilitet og validitet i testmateriellet, selv i et kommersielt perspektiv. Det er viktig at forskningsmateriale er etterprøvbart, og da må denne typen informasjon være tilgjengelig. Når prøveutviklere velger å ikke utlevere slik informasjon, så kan det argumenteres for at det undergraver god forskningsetikk.

Det er viktig at de som bruker prøvene, vanligvis lærere, vet hvilken prøve som er god til ulike formål, slik at de kan gjøre velbegrunnede valg når de velger kartleggingsmateriell. Ofte bruker man gjerne prøver man har hørt fra andre er gode, uten at det er gjort undersøkelser som støtter eller avkrefter dette. I min undersøkelse av kartleggingsprøvene som brukes i skolen, er det nedslående å se hvor få prøver som er testet for reliabilitet og/eller validitet, og hvor få som er normert. Når disse elementene ikke er undersøkt i prøvene, så kan man ikke objektivt si noe om kvaliteten i prøvematerialet. Det blir da umulig å si om prøven er god eller dårlig, eller om den egner seg til de formålene den brukes til.

Det er også et poeng å ha med seg at kartleggingsprøvene kanskje ikke forteller oss hva elevene kan hvis vi utelukkende kartlegger på et språk eleven kan dårlig (Shultz, Haugen & Støre, 2008). Kartleggingsprøvene forsøker å kartlegge de ulike språkferdighetene, men det er også en viss risiko for at elever gjetter seg frem til svar, særlig hvis det er tre eller færre svaralternativer (Hargreaves, 2010), og at det er utfordringer knyttet til kartlegging hvis eleven er god i såkalte reseptive ferdigheter (lese og lytte), men svak i produktive ferdigheter (skrive og snakke) på andrespråket (Shultz, Haugen & Støre, 2008).

Kartlegging av minoritetsspråklige elever i grunnskole og videregående skole kan få store konsekvenser for den enkelte elev fordi rettigheter kan gis eller falle bort. I teorikapitlet

trekker jeg frem begrepet 'tilstrekkelige norskferdigheter' og hvordan dette ser tydelig og avklart ut i læreplanen og kartleggingsverktøyet til Utdanningsdirektoratet, samtidig som Kunnskapsdepartementet (2007) påpeker at dette er et uavklart begrep som bestemmes av lærerens skjønnsmessige vurdering. Når man sitter med et begrep som skal avgjøre om eleven får eller ikke får en rettighet i lovverket, og samtidig sitter med kartleggingsprøver som i liten eller ingen grad er normert og undersøkt for reliabilitet og validitet, så kan man stille spørsmålstegn ved om ikke det er å trekke lodd om hvilke elever som får og ikke får tilbud om særskilt språkopplæring. Både Engen (2010), Rambøll Management (2016) og Carlsen og Moe (2019) stiller i tillegg spørsmålstegn ved om ikke elever sendes fra særskilt språkopplæring til 'ordinær' undervisning for tidlig. Så lenge begrepet 'tilstrekkelige norskferdigheter' ikke er avklart i lov, forskrift el.lign., og kartleggingsprøvene ikke normeres og undersøkes for reliabilitet og validitet, så er dette spørsmål vi ikke kan svare på. Dersom kartleggingsprøvene brukes til å gi eller frata elever rettigheter i forhold til §2-8 eller §3-12, så blir denne litteraturgjennomgangen veldig viktig fordi kartleggingsprøven i seg selv på den måten kobles sammen med et lovverk som kan gi eller ikke gi elever rettigheter. For minoritetsspråklige som kartlegges i skolen kan det gi utslag i om de får rettigheter jf. §2-8 eller §3-12, eller *hvilken type* særskilte tilbud for språkopplæring eller språkstøtte de får tilbud om. For voksne flyktninger kan det ha stor betydning fordi en av prøvene brukes som grunnlag for å søke permanent opphold og statsborgerskap i Norge, selv om prøven som brukes til dette er den som i størst grad undersøker reliabilitet og validitet.

Man kan også stille spørsmålstegn ved om det er kjennskap og troverdigheten til testutviklerne som bestemmer hvilke prøver man velger å bruke. Utdanningsdirektoratet, Trondheim kommune, Fagbokforlaget, Kompetanse Norge, Telemarksforskning Notodden, Vestland fylkeskommune og Høgskulen på Vestlandet. Alle disse er avsendere som tilsynelatende har god kunnskap om emnet og som gjerne er tilknyttet gode forskere og medarbeidere, men betyr det at prøven er god og måler det den sier den måler? Innen medisin ville man aldri godkjent en medisin bare fordi det var et anerkjent legemiddelfirma som stod bak. Alt må testes i flere runder før det godkjennes for distribusjon, og kravene for godkjenning er veldig strenge. Hvorfor er det annerledes med kartleggingsprøver som i denne sammenhengen kan ha stor betydning for elevens rett til særskilt språkopplæring (§2-8 og §3-12 i Opplæringslova) og organisering av opplæringstilbudet? Selv om ingen liv går tapt ved å bruke prøver som ikke er reliable og valide, så gir de store reelle konsekvenser for elevenes opplæring.

Alderson (2007) kritiserer også at *Rammeverket* ikke er bygd opp rundt mer teoretisk forskning, og at det utvikles og brukes i såkalte diagnostiske prøver der deltakerne får et stempel, eller en 'diagnose' med sitt språknivå, A2, B1, osv. Dette ser vi igjen i flere av kartleggingsprøvene som er koblet opp mot *Rammeverket*, selv om det kommer tydeligst frem i *Norskprøven* og *Migranorsk*. Det at man gjennom *Rammeverket* bare ser på språkkinnlærerens styrker og at man ikke får kartlagt de språklige svakhetene eller utfordringene, gjør det også vanskelig å vurdere hvor læringen skal fortsette. For eleven er det positivt å få vite hva man kan, men fra et lærerperspektiv er det gjerne viktigere å avdekke hva eleven strever med. Da kan man jobbe med det som eleven ikke får til slik at eleven kan mestre flere området. Med et utelukkende positivt syn på språklæring, så vil elever kunne mangle kunnskaper på områder som gjerne er viktig for videre læring. At en elev mestrer å lese e-post, skrive en fortelling eller delta i en samtale er gjerne mindre viktig enn å vite hvorfor en elev strever med disse tingene. Ser man på diagnostiske verktøy på andre områder, f.eks. psykologi eller førstespråkslesing, så er det vanligvis fokus på å avdekke eventuelle vansker. Hvorfor velges det å ikke gjøre dette i et andrespråksperspektiv? Hvorfor skal andrespråkstester fortelle eleven det han eller hun kan uten å si noe om elevens utfordringer, svakheter eller mangler i andrespråklæring, -lesing, eller det testen måler? Som elev vil man gjerne få positive tilbakemeldinger om hva man mestrer, og på den måten er nivåbeskrivelsene i *Rammeverket* gode for å gi elevene bekreftelse og mestringsfølelse. Samtidig er det viktig å huske på at læreren trenger å vite hva elevene ikke mestrer for å kunne tilpasse den videre opplæringen slik at eleven får best mulig læringsutbytte. Paradokset oppstår når man bruker elevsentrerte mestrings-tilbakemeldinger som grunnlag for å gi elever rettigheter til mer og tilpasset språkopplæring, som i seg selv sier at eleven har utfordringer. Er det ikke da viktig å avdekke utfordringene og bruke disse som utgangspunkt for å gi elever rett til særskilt språkopplæring?

6. Avslutning og veien videre

Det brukes mange ulike kartleggingsprøver for å kartlegge norskferdigheter i innføringstilbud i tidligere Hordaland fylke, både kommersielle prøver, ikke-kommersielle prøver og lokalt utviklede prøver. Alle skoler som har svart på undersøkelsen sier de kartlegger minst en gang i året, og flere skoler kartlegger tre-fire ganger i året. Prøvene brukes for å plassere elever i nivådelte grupper, tilpasse opplæringen til den enkelte og til å avdekke vansker på første- eller andrespråket. Noen oppgir også å bruke prøvene til å gi elever rett til særskilt språkopplæring, jf. §2-8 eller §3-12 i Opplæringslova, og/eller som dokumentasjon på opplæring til ny klasse eller ny skole.

Det er bare *Norskprøven* som vurderer reliabilitet og validitet i sin prøve. *TOSP* er normert på normalpopulasjon på 3., 5. og 7.trinn ved fem skoler i Skien kommune, mens *FLORO* har en veiledende normering på norsk fra 245 elever med norsk som morsmål, og fra tilgjengelige elever og språk fra innføringsklasser i Bergen, Drammen, Trondheim, Stavanger og Bærum. Elever skal være rekruttert fra områder og familier med ulik sosioøkonomisk bakgrunn, men det er likevel variasjoner i populasjonsgrunlaget på de ulike språkene. *Migranorsk* vil ikke gi ut informasjon om reliabilitet og validitet. De øvrige testene er ikke undersøkt for reliabilitet og validitet og er heller ikke normert. Som resultatene viser, så er de mest brukte kartleggingsverktøyene, *Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk og Trondheimstesten*, ikke undersøkt for reliabilitet og validitet.

Når det er få prøver som er undersøkt reliabilitet og validitet i forhold til, så vet vi ikke hvor gode disse testene er, eller om de egner seg til de formålene de brukes til. Jeg har gjennom diskusjonen prøvd å trekke frem at mange av prøvene brukes til flere ulike formål samtidig som vi ikke vet i hvilken grad de egner seg til disse formålene eller om prøvene i det hele tatt måler det de sier de måler. Når situasjonen er slik, så er det sannsynlig at det blir mange tilfeldige resultater på kartleggingsresultatene til elevene. Om vi i tillegg tar høyde for begrepet 'tilstrekkelige norskferdigheter', så forsterkes dette enda mer.

Begrepet 'tilstrekkelige ferdigheter i norsk' er uavklart og noe som Kunnskapsdepartementet mener lærer individuelt skal vurdere samtidig som Utdanningsdirektoratet gjennom *Læreplan i grunnleggende norsk for språklige minoriteter* sier at man etter nivå 3 skal gå over i ordinær undervisning og følge ordinær læreplan i norsk. Dette punktet forsvant med innføring av ny læreplan skoleåret 2020/2021. Det er viktig at begreper som er avgjørende for en elevs rettigheter er avklart i lovtekst, forskrift el.lign. Når det ikke gjøres, så blir kriteriene diffuse og det blir vanskelig får den enkelte lærer å vurdere når eleven kan nok til ordinær

undervisning. Og dette med kriterier er en viktig del av det å lage en kartleggingsprøve. Dersom kriteriene for språkprøven ikke er tydelige, så blir det nesten en umulig oppgave å lage en god prøve, og dette kan man gjerne hevde er verre enn om prøven ikke er undersøkt for reliabilitet og validitet.

For lærer som har elevgruppen, er det viktig å vite hvilke kartleggingsverktøy som er gode å bruke, og hvilke verktøy som passer til hvilke formål. Elevene kartlegges både for å nivådeles i gruppe, gi tilpasset opplæring, avdekke vansker og dokumentere språkkunnskaper. I tillegg brukes noen av prøvene til å gi elever rett til særskilt språkopplæring. Det er viktig at prøver som brukes til ulike formål er egnet til dette. Ser vi dette i sammenheng med manglende begrepsavklaring av 'tilstrekkelige norskferdigheter' og ellers manglende kriterier til språkprøver, så er det store utfordringer knyttet til språkprøvene.

Som jeg tar opp i drøftingen, så er det også viktig å merke seg at dette er en populasjon det er vanskelig å lage gode normeringer og kriterier til. Det er vanskelig å avgjøre hva man skal ta hensyn til når man lager kriteriene, og det gjør testutvikling til en veldig utfordrende oppgave. Dermed er det viktig at man har med seg begrensningene når man bruker testmateriellet som er tilgjengelig. Det er også viktig at man jobber med å normere testmaterieell og tester reliabilitet og validitet i de prøvene som brukes i skolen. Da kan vi få innsikt i hvilken kvalitet prøvene har. I tillegg kan man gjøre endringer i testmateriellet dersom man ser det som hensiktsmessig. Dette er særlig viktig for de prøvene som brukes for å gi eller ikke gi elever rett til særskilt språkopplæring jf. §2-8 eller §3-12 i Opplæringslova. Når prøvene brukes til dette formålet så blir prøvene 'high stakes'-prøver og har stor påvirkning på elevens videre opplæring. Da er det viktig at alle elever vurderes etter de samme kriteriene for å få denne rettigheten, og at ikke det er individuelle læreres vurderinger som skal være det som står mellom eleven og rettigheten.

Kilder

- Alderson, C. (2007). The Challenge of (Diagnostic) Testing: Do We Know What We Are Measuring? I red. J. Fox, M. Wesche, D. Bayliss, L. Cheng, C.E. Turner & C. Doe: *Language Testing Reconsidered* (ss. 21-40). University of Ottawa Press. Hentet fra: <https://jstor.org/stable/j.ctt1ckpccf.8>
- Amundsen, E. (2021). E-postutveksling om *Trondheimstesten*. Trondheim kommune.
- Antoniazzi, D.; Snow, P. & Dickson-Swift, V. (2010). Teacher identification of children at risk for language impairment in the first year of school. *International Journal of Speech-Language Pathology, Vol. 12, No. 3, ss. 244-252*. DOI: 10.3109/17549500903104447
- Arnesen, A., Braecken, J., Ogden, T. & Melby-Lervåg, M (2019). Assessing Children's Social Functioning and Reading Proficiency: A Systematic Review of the Quality of Educational Assessment Instruments Used in Norwegian Elementary Schools. *Scandinavia Educational Research, ss. 465-490*. DOI: <https://doi.org/10.1080/0031381.2017.1420685>
- Asdal, K. & Reinertsen, H. (2020). *Hvordan gjøre en dokumentanalyse. En praksisorientert metode*. Oslo: Cappelen Damm Akademisk.
- Bachman, L. (2002). Alternative Interpretations of Alternative Assessments: Some Validity Issues in Educational Performance Assessments. *Educational measurement, issues and practice, 2002-09, Vol. 21 (3), ss.5-18*. DOI: 10.1111/j.1745-3992.2002.tb00095.x
- Bachman, L. & Palmer, A. S. (1996). *Language Testing in Practice*. Oxford. Oxford University Press.
- Berggreen, H. & Tenfjord, K. (2007). *Andrespråklæring*. Oslo: Gyldendal Akademisk.
- Booth, A., Sutton, A. & Papaioannou, D. (2016). *Systematic approaches to a successful literature review*. Sage Publications.
- Brubakk, B., de la Cruz, S. & Landsem, M. (2009). *Kartleggingsprøver for flerspråklige elever. Fra mottaksgruppe til 5-7.trinn. Elevhefte*. Hentet fra Trondheim kommune: <https://www.trondheim.kommune.no/contentassets/da2bf07e93154a86b6c6c52a9589770b/pi2988-mottak-trinn-5-7.trinn-skriftlig-web-1.pdf>
- Brubakk, B., de la Cruz, S. & Landsem, M. (2013a). *Kartleggingsprøver for flerspråklige elever. Fra mottaksgruppe til 5-7.trinn. Muntlige oppgaver*. Hentet fra Trondheim kommune: <https://www.trondheim.kommune.no/contentassets/65100d1fdd574df4a3df199dd321e69a/pi9152-komplett-web-muntlig-5.-7.trinn.pdf>
- Brubakk, B., de la Cruz, S. & Landsem, M. (2013b). *Kartleggingsprøver for flerspråklige elever. Fra mottaksgruppe til 3-4.trinn. Elevhefte*. Hentet fra Trondheim kommune: <https://www.trondheim.kommune.no/contentassets/da2bf07e93154a86b6c6c52a9589770b/pi9180-komplett-skiftlig-web-3.-4.trinn.pdf>
- Brubakk, B., de la Cruz, S. & Landsem, M. (2013c). *Kartleggingsprøver for flerspråklige elever. Fra mottaksgruppe til 3-4.trinn. Muntlige oppgaver*. Hentet fra Trondheim

- kommune:
<https://www.trondheim.kommune.no/contentassets/65100d1fdd574df4a3df199dd321e69a/pi9151-komplett-muntlig-web-3.-4.trinn.pdf>
- Brubakk, B., de la Cruz, S. & Landsem, M. (2013d). *Kartleggingsprøver for flerspråklige elever. Fra mottaksgruppe til 5-7.trinn. Fasit, lærerveiledning og notathefte*. Hentet fra Trondheim kommune:
<https://www.trondheim.kommune.no/contentassets/65100d1fdd574df4a3df199dd321e69a/pi9155-komplett-web-5.-7.trinn-notathefte.pdf>
- Brubakk, B., de la Cruz, S. & Landsem, M. (2013e). *Kartleggingsprøver for flerspråklige elever. Fra mottaksgruppe til 3-4.trinn. Fasit, lærerveiledning og notathefte*. Hentet fra Trondheim kommune:
<https://www.trondheim.kommune.no/contentassets/65100d1fdd574df4a3df199dd321e69a/pi9154-komplett-web-3.-4.trinn-notathefte.pdf>
- Brubakk, B., de la Cruz, S. & Landsem, M. (2015a). *Kartleggingsprøver for minoritetsspråklige skolestartere. Muntlige oppgaver*. Hentet fra Trondheim kommune:
<https://www.trondheim.kommune.no/contentassets/da2bf07e93154a86b6c6c52a9589770b/4059-barnas-hefte-web1-1.pdf>
- Brubakk, B., de la Cruz, S. & Landsem, M. (2015b). *Kartleggingsprøver for minoritetsspråklige skolestartere. Fasit, lærerveiledning og notathefte*. Hentet fra Trondheim kommune:
<https://www.trondheim.kommune.no/contentassets/da2bf07e93154a86b6c6c52a9589770b/4095-lrerveiledning-040315-1.pdf>
- Carlsen, C. (2012). *NORSK PROFIL. Det felles europeiske rammeverket spesifisert for norsk. Et første sted*. Oslo: Novus Forlag AS.
- Carlsen, C. (2021). E-postutveksling om *Vestlandsverktøyet*. Høgskulen på Vestlandet.
- Carlsen, C. H. & Moe, E. (2019). *Vurdering av språkferdigheter*. Bergen: Fagbokforlaget.
- Carmines, E. G. & Zeller, R. A. (2011). *Reliability and Validity Assessment*. Thousand Oaks: SAGE Publications Inc.
- Colliver, J. A., Conlee, M. J. & Verhulst, S. J. (2012). From test validity to construct validity ... and back? *Medical Education*, Vol. 46, ss. 366-371. DOI: 10.1111/j.1365-2923.2011.04194.x
- Det kongelige barne-, likestillings- og inkluderingsdepartement. (2012-2013). *Meld.St.6 En helhetlig integreringspolitikk – Mangfold og fellesskap*. Oslo: Regjeringen.
- Domino, G. & Domino, M. L. (2006). *Psychological Testing. An Introduction*. Cambridge University Press. DOI: <https://doi-org.galanga.hvl.no/10.1017/CBO9780511813757>
- Downing, S. (2003). Validity: on the meaningful interpretation of assessment data. *Medical Education*, Vol. 37 (9), ss. 830-837. DOI: 10.1046/j.1365-2923.2003.01594.x
- Egeberg, E. (2016). *Flere språk – flere muligheter. Flerspråklighet, tilpasset opplæring og spesialpedagogisk metodikk*. Oslo: Cappelen Damm Akademisk.

- Engen, T. O. (2010). Viss den brede vei blir den smale sti. Om lovgrunnlaget for tilpasset opplæring for minoritetsspråklige elever. I red. H. Johansen, A. Golden, J. E. Hagen & A.-K. Helland: *Systematisk, variert, men ikke tilfeldig*. Ss. 24-37. Oslo: Novus Forlag.
- Evers, A., Hagemester, C., Høstmælingen, A., Lindley, P., Muñiz, J. & Sjöberg, A. (2013). *EFPA Review Model for the Description and Evaluation of Psychological and Educational Tests*. Hentet fra European Federation of Psychologist's Association: <http://assessment.efpa.eu/documents/>
- Fagbokforlaget. (2020a). *Brukerveiledning for tester (Migranorsk)*. Fagbokforlaget.
- Fagbokforlaget. (2020b). *Migranorsk test A1/A2/B1/B2 2020*. Fagbokforlaget.
- Fink, A. (2014). *Conducting Research Literature Reviews. From the Internet to Paper*. California, USA: SAGE Publications Inc.
- Folkeuniversitetet & Norsk språktest. (2011a). *Norsprøve 2 for voksne innvandrere (nivå A2)*. Folkeuniversitetet & Norsk språktest.
- Folkeuniversitetet & Norsk språktest. (2011b). *Norsprøve 3 for voksne innvandrere (nivå B1)*. Folkeuniversitetet & Norsk språktest.
- Folkeuniversitetet & Norsk språktest. (2011c). *Veiledning til Norsprøve 2 for voksne innvandrere (nivå A2)*. Folkeuniversitetet & Norsk språktest.
- Folkeuniversitetet & Norsk språktest. (2011d). *Veiledning til Norsprøve 3 for voksne innvandrere (nivå B1)*. Folkeuniversitetet & Norsk språktest.
- Garrard, J. (2011). *Health Sciences Literature Review Made Easy – The Matrix Method*. Sudbury, Massachusetts: Jones & Bartlett Learning.
- Goodwin, L. D. (2002). The Meaning of Validity. *Journal of Pediatric Gastroenterology and Nutrition, Vol. 35, Issue 1, ss. 6-7*. Hentet fra: https://journals.lww.com/jpgn/Fulltext/2002/07000/The_Meaning_of_Validity.3.aspx
14
- Goodwin, L. D. & Leech, N. L. (2017). The Meaning of Validity in the New Standards for Educational and Psychological Testing: Implications for Measurement Courses. *Measurement and Evaluation in Counselling and Development, Vol. 36, Issue 3, ss. 181-191*. DOI: 10.1080/07481756.2003.11909741
- Gough, P. B. (1996). How Children Learn to Read and Why They Fail. *Annals of Dyslexia, Vol. 46, ss. 3-20*. DOI: <https://doi.org/10.1007/BF02648168>
- Hargreaves, M. (2010). Citizenship Testing in the Anglophone Countries: The UK, Canada and the USA. I red. C. Slade & M. Martina: *From Migrant to Citizen: Testing Language, Testing Culture*. Ss. 101-124. New York: Palgrave Macmillan.
- Hasle, S. (2013). Sammenhengen mellom ordforråd og avkodingsferdigheter i teori og forskning. *Norsk Tidsskrift for Logopedi, Vol. 2, ss. 6-13*. Hentet fra: https://norsklogopedlag.no/Userfiles/Upload/Files/Logopeden_2-13_sammenhengen.pdf

- Heller, M. C. & Lervåg, A. (2020). Kartlegging av norskspråklige ferdigheter for elever på 1.-4.trinn: En vurdering av kartleggingsverktøyet Norsk Som Læringspråk (NSL). *Norsk Tidsskrift for Logopedi*, Vol. 3, ss. 6-17. Hentet fra: https://norsklogopedlag.no/Userfiles/Upload/files/Logopedi_3-20_HellerLerva%CC%8Ag.pdf
- Hoover, W. & Gough, P. (1990). The Simple View of Reading. *Reading and Writing: An Interdisciplinary Journal*, Vol. 2, ss. 127-160. DOI: <https://doi.org/10.1007/BF00401799>
- Husby, O. & Kløve, M. H. (2007). *Andrespråksfonologi. Teori og metodikk*. Oslo: Ad Notam Gyldendal.
- Jacobsen, D. I. (2016). *Hvordan gjennomføre undersøkelser? Innføring i samfunnsvitenskapelig metode*. Cappelen Damm Akademisk.
- Jesson, J. K., Matheson, L. & Lacey, F. M. (2011). *Doing Your Literature Review. Traditional and Systematic Techniques*. Sage Publications Inc.
- Jortveit, M. (2017). *Inkludering av minoritetsspråklige elever i skolen*. Kristiansand: Portal Akademisk.
- Kjelsaas, I., Eide, K., Vidarsdottir, A. S. & Hauge, H. A. (2020). Å føle seg angrepet intellektuelt: Språkkartlegging av minoritetsspråklige elever i videregående opplæring. *Acta Didactica Norden*, Vol. 14, Nr. 1, Art. 16. DOI: <http://dx.doi.org/10.5617/adno.7871>
- Kleven, T. A. & Hjordemaal, F. R. (2018). *Innføring i pedagogisk forskningsmetode. En hjelp til kritisk tolkning og vurdering*. Fagbokforlaget.
- Koda, K. (2012). Development of second language reading skills. Cross-linguistic perspectives. I red: S. M. Gass & A. Mackey: *The Routledge Handbook of Second Language Acquisition* (ss. 303-308). Routledge.
- Korsgaard, K., Hannibal, S. & Vitger, M. (2011). *Oppdagende skriving – En vei inn i lesingen*. Oslo: Cappelen Damm Akademisk.
- Kunnskapsdepartementet. (2007). *Likeverdig opplæring i praksis*. Hentet fra Regjeringen: https://www.regjeringen.no/globalassets/upload/kd/vedlegg/grunnskole/strategiplaner/udir_likeverdig_opplaering2_07.pdf
- Loona, S. (2008a). *Instruksjonshefte for bruk av TOSP*.
- Loona, S. (2008b). *TOSP. Tospråklige prøve norsk*.
- Melby-Lervåg, M. & Lervåg, A. (2014). Reading Comprehension and Its Underlying Components in Second-Language Learners: A Meta-Analysis of Studies Comparing First- and Second- Language Learners. *Psychological Bulletin*, Vol. 140, No. 2, ss. 409-433. DOI: 10.1037/a0033890
- Messick, S. (1994). Validity of psychological assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *ETS*

Research Report Series, Volume 1994, Issue 2, ss. i-28. DOI: [10.1002/j.2333-8504.1994.tb01618.x](https://doi.org/10.1002/j.2333-8504.1994.tb01618.x)

- Messick, S. (1996). Validity and Washback in Language Testing. *ETS Research Report Series, Volume 1996, Issue 5*. Hentet fra Educational Testing Service: <https://onlinelibrary.wiley.com/doi/pdfdirect/10.1002/j.2333-8504.1996.tb01695.x>
- Midtbø, T.; Rossow, A. & Sagbakken, B. (2018). Måling av sensorreliabilitet ved vurdering av norskprøve i skriftlig framstilling. *Acta Didactica Norge. Vol. 12, Nr. 4, Art. 12*. DOI: <http://dx.doi.org/10.5617/adno.6358>
- Opplæringslova. (u.d.). *Lov om opplæring i grunnskolen og den vidaregåande skolen*. Hentet fra Lovdata: <https://lovdata.no/dokument/NL/lov/1998-07-17-61>
- Pickering, L. (2012). Second language speech production. I red: S. M. Grass & A. Mickey: *The Routledge Handbook of Second Language Acquisition* (ss. 335-348). Routledge.
- Polio, C. (2012). The acquisition of second language writing. I red: S. M. Grass & A. Mickey: *The Routledge Handbook of Second Language Acquisition* (ss. 319-334). Routledge.
- Rambøll Management. (2006). *Evaluering av praktisering av norsk som andrespråk for språklige minoriteter i klasserommet*. Hentet fra Utdanningsdirektoratet: https://www.udir.no/globalassets/filer/tall-og-forskning/rapporter/5/evaluering_av_norsk_som_2_sprak.pdf
- Rambøll Management. (2009). *Evaluering av implementering av læreplaner i grunnleggende norsk og morsmål for språklige minoriteter*. Hentet fra Utdanningsdirektoratet: https://www.udir.no/globalassets/filer/tall-og-forskning/rapporter/5/delrapport1_lareplaner_spraklige_minoriteter.pdf
- Rambøll Management. (2016). *Evaluering av særskilt språkopplæring og innføringstilbud*. Hentet fra Utdanningsdirektoratet: <https://www.udir.no/globalassets/filer/tall-og-forskning/forskningsrapporter/evaluering-av-sarskilt-sprakopplaring-2016.pdf>
- Randen, G. T. (2013). *Tilstrækkelige ferdigheter i norsk? Kartlegging av minoritetsspråklige skolebarns språkferdigheter*. Hentet fra Brage: <https://brage.inn.no/inn-xmli/bitstream/handle/11250/132026/Randen.pdf?sequence=1&isAllowed=y>
- Ringdal, K. (2001). *Enhet og mangfold. Samfunnsvitenskapelig forskning og kvantitativ metode*. Bergen: Fagbokforlaget Vigmostad & Bjørke AS.
- Schultz, J.-H., Haugen, A.-M. & Støre, H. (2008). *Ingen ut av rekka går. Tilpasset opplæring for Shaza og Kristian*. Oslo: Universitetsforlaget.
- Simonsen, T. & Mikalsen, M.-L. (2008). *Europeisk språkperm 6-12. Lærerhåndbok*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/larerhandbok_smaperm_endelig_2016.pdf
- Simonsen, T. & Speitz, H. (2006). *Europeisk språkperm 13-18. Lærerhåndbok*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/larerveiledning_spraakperm_13-18.pdf

- Simonsen, T. & Speitz, H. (2014a). *Europeisk språkperm 13-18*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/1_sprakp_13-18_sprakbiografi_interaktiv_14_bm.pdf
- Simonsen, T. & Speitz, H. (2014b). *Sjekkliste A1 Europeisk språkperm 13-18*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/2_sprakp_13-18_sjekkliste_a1_interaktiv_14_bm.pdf
- Simonsen, T. & Speitz, H. (2014c). *Sjekkliste A2 Europeisk språkperm 13-18. Lærerhåndbok*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/2_sprakp_13-18_sjekkliste_a2_interaktiv_14_bm.pdf
- Simonsen, T. & Speitz, H. (2014d). *Sjekkliste B1 Europeisk språkperm 13-18. Lærerhåndbok*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/2_sprakp_13-18_sjekkliste_b1_interaktiv_14_bm.pdf
- Simonsen, T. & Speitz, H. (2014e). *Sjekkliste B2 Europeisk språkperm 13-18. Lærerhåndbok*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/2_sprakp_13-18_sjekkliste_b2_interaktiv_14_bm.pdf
- Simonsen, T. & Speitz, H. (2014f). *Sjekkliste C1 Europeisk språkperm 13-18. Lærerhåndbok*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/2_sprakp_13-18_sjekkliste_c1_interaktiv_14_bm.pdf
- Simonsen, T. & Speitz, H. (2014g). *Sjekkliste C2 Europeisk språkperm 13-18. Lærerhåndbok*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/2_sprakp_13-18_sjekkliste_c2_interaktiv_14_bm.pdf
- Südkamp, A; Kaiser, J. & Möller, J. (2012). Accuracy of teachers' judgements of students' academic achievements. A meta-analysis. *Journal of Educational Psychology, Vol. 104, No. 3, ss. 743-762*. DOI: <https://doi.org/10.1037/a0027627>
- Telemarksforskning Notodden (2014a). *Europeisk språkperm 6-12*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/sprakp_6-12_10_bm.pdf
- Telemarksforskning Notodden (2014b). *Sjekkliste A1 Europeisk språkperm 6-12*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/sprakp_6-12_10_bm-sjekkliste_a1.pdf
- Telemarksforskning Notodden (2014c). *Sjekkliste A2 Europeisk språkperm 6-12*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/sprakp_6-12_10_bm-sjekkliste_a2.pdf
- Telemarksforskning Notodden (2014d). *Sjekkliste B1 Europeisk språkperm 6-12*. Hentet fra Fremmedspråkssenteret: https://www.hiof.no/fss/leringsressurser/den-europeiske-sprakpermen/sprakp_6-12_10_bm-sjekkliste_b1.pdf

- Tornberg, U. (2000). *Språkdidaktikk*. Forlaget Fag og Kultur AS.
- Utdanningsdirektoratet. (2007a). *Kartleggingsmaterieell: Språkkompetanse i grunnleggende norsk*. Hentet fra Utdanningsdirektoratet: https://www.udir.no/globalassets/upload/kartleggingsprover/udir_kartleggingsmaterieell_bm_301007.pdf
- Utdanningsdirektoratet. (2007b). *Læreplan i grunnleggende norsk for språklige minoriteter (NOR7-01)*. Hentet fra Utdanningsdirektoratet: <https://www.udir.no/kl06/NOR7-01/Hele/Vurdering>
- Utdanningsdirektoratet. (2020). *Læreplan i grunnleggende norsk for språklige minoriteter (NOR07-02)*. Hentet fra Utdanningsdirektoratet: <https://www.udir.no/lk20/nor07-02>
- Utdanningsdirektoratet. (2009). *Veiledning. Språkkompetanse i grunnleggende norsk*. Hentet fra Utdanningsdirektoratet: https://www.udir.no/globalassets/upload/brosjyrer/5/veiledning_grunnleggende_norsk.pdf
- Utdanningsdirektoratet. (2011). *Det felles europeiske rammeverket for språk*. Hentet fra Språksenteret i Oslo skolen: <https://spraksenteret.oslo skolen.no/siteassets/vgs/det-felles-europeiske-rammeverket-for-sprak.pdf>
- Utdanningsdirektoratet. (2016). *Minoritetsspråklige – hva ligger i begrepet?* Hentet fra Utdanningsdirektoratet: <https://www.udir.no/laring-og-trivsel/minoritetsspraklige/minoritetsspraklige---hva-ligger-i-begrepet/>
- Utdanningsdirektoratet & NAFO. (2013a). *FLORO. Noteringshefte*. Utdanningsdirektoratet & NAFO.
- Utdanningsdirektoratet & NAFO. (2013b). *FLORO. Veiledning og instruksjoner*. Utdanningsdirektoratet og NAFO.
- Vestland fylkeskommune & Høgskulen på Vestlandet. (2020a). *Instruksjon for gjennomføring av skriftlig prøve*. Vestland fylkeskommune.
- Vestland fylkeskommune & Høgskulen på Vestlandet. (2020b). *Vurderingsskjema for muntlig produksjon*. Vestland fylkeskommune.
- Vestland fylkeskommune & Høgskulen på Vestlandet. (2020c). *Vurderingsskjema for skriftlig produksjon*. Vestland fylkeskommune.
- Vestland fylkeskommune & Høgskulen på Vestlandet. (2020d). *Vurderingsverktøy for kartlegging av norskferdigheter. Muntlig prøve*. Vestland fylkeskommune.
- Vestland fylkeskommune & Høgskulen på Vestlandet. (2020e). *Vurderingsverktøy for kartlegging av norskferdigheter. Skriveprøve*. Vestland fylkeskommune.
- Vestland fylkeskommune & Høgskulen på Vestlandet. (2020f). *Vurderingsverktøy for minoritetsspråklige elever i VGS. Forklaring til kriterier i vurdering av muntlig produksjon*. Vestland fylkeskommune.

Vestland fylkeskommune & Høgskulen på Vestlandet. (2020g). *Vurderingsverktøy for minoritetsspråklige elever i VGS. Forklaring til kriterier i vurdering av skriftlig produksjon*. Vestland fylkeskommune.

Weir, C. J. (2005). *Language Testing and Validation. An Evidence-Based Approach*. New York: Palgrave Macmillan.

Wechsler, D. (2013). *WISC-IV Manual. Norsk versjon*. Pearson.

Vedlegg 1: Spørreskjema

Spørreskjema om bruk av kartleggingsverktøy for kartlegging av språkferdigheter i innføringstilbud

Bakgrunnsinformasjon

1. Innføringstilbudet er lokalisert på en:
 - Barneskole
 - Ungdomsskole
 - Videregående skole

Om kartlegging

2. Hvilke kartleggingsmateriell bruker skolen for å kartlegge minoritetsspråklige elever i innføringstilbud?
 - Kartleggingsmateriell: Språkkompetanse i grunnleggende norsk (Udir)
 - Den europeiske språkpermen (6-12 og/eller 13-18)
 - FLORO (NAFO)
 - TOSP
 - Migranorsk
 - Norskprøven for voksne innvandrere
 - Vurderingsverktøy for minoritetsspråklige elever (kalt «Vestlandsverktøyet», utviklet av Hordaland Fylkeskommune)
 - Lokalt utviklet kartleggingsprøve
 - Annen prøve: _____
3. Hva brukes kartleggingsprøven til? (*denne kommer til hver av prøvene det krysses av for*)
 - Å plassere elevene i nivådelte grupper
 - Å tilpasse opplæringen av den enkelte elev i klassen
 - Å finne ut om elever har vansker på andrespråket
 - Vet ikke
 - Annet: _____
4. Hvor mange ganger i året kartlegges elevene? _____

Vedlegg 2: Informasjonsskriv til skolene

Kjære skole og lærere

Jeg skriver en masteroppgave i spesialpedagogikk om bruk av kartleggingsverktøy i innføringstilbud i tidligere Hordaland, ved Høgskulen på Vestlandet. Jeg ønsker å finne ut *hvilke* kartleggingsverktøy som brukes i innføringstilbud og *hva* de brukes til i disse klassene.

Dere mottar denne eposten fordi dere har eller har hatt innføringstilbud. I den forbindelse ønsker jeg at eposten videresendes til de som er norsklærer(e) i innføringstilbud slik at de kan svare på et spørreskjema knyttet til bruk av kartleggingsverktøy for språkferdigheter i innføringstilbud. Jeg ønsker ett svar per innføringstilbud/-klasse.

Undersøkelsen er anonym, det er frivillig å delta og det tar ca. 2-4 minutter å gjennomføre undersøkelsen. Om du velger å delta, kan du når som helst, og uten begrunnelse trekke din deltakelse.

Når du svarer på undersøkelsen blir du først bedt om å registrere en e-postadresse. Dette er for at vi skal kunne fjerne svarene dine dersom du trekker samtykket. Når du har registrert e-postadressen så vil du få tilsendt en ny e-post med en lenke til spørreundersøkelsen. Her kan du lese om studien og gi samtykke før du får gå videre til selve undersøkelsen.

Prosjektet ble godkjent hos NSD 13.11.2020 og vil følge alle gjeldende regler for datainnsamling og -håndtering. Det samlede datamaterialet vil ligge tilgjengelig for meg og veilederne mine på forskningsserveren til Høgskulen på Vestlandet til prosjektet avsluttes i juni 2021. Etter denne perioden vil alle data bli slettet. Deltakerne kan når som helst, og uten begrunnelse, trekke sin deltakelse i studien. Du trekker enkelt samtykke ved å ta kontakt med meg på oppgitt e-post nedenfor.

Jeg håper eposten videresendes til norsklærere i innføringstilbud.

Ved spørsmål om undersøkelsen ta gjerne kontakt.

Lenke til undersøkelsen: [lenke fjernet]

Med vennlig hilsen

Hege Starnesfet

Høgskulen på Vestlandet, campus Sogndal

Epost: 575333@stud.hvl.no

Veiledere:

Arne Ola Lervag, UiO, e-post: a.o.lervag@iped.uio.no

Hilde Hofslundsengen, HVL, e-post: Hilde.Hofslundsengen@hvl.no

Vedlegg 3: Samtykkeskjema

Vil du delta i forskningsprosjektet "Kartleggingsverktøy i innføringstilbud"?

Dette er et spørsmål til deg om å delta i et masteroppgaveprosjekt hvor formålet er å finne ut hvilke kartleggingsprøver som brukes for å kartlegge norskspråklige ferdigheter i innføringstilbud. I dette skrevet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

Formålet med mastergradsoppgaven om kartleggingsverktøy er å undersøke hva som brukes for å kartlegge norskspråklige ferdigheter i såkalte innføringstilbud i tidligere Hordaland fylke.

Opplæringslova §2-8 og §3-12 pålegger kommunene å kartlegge alle minoritetsspråklige elever for å avgjøre om de har «tilstrekkelige ferdigheter i norsk» til å følge ordinær læreplan. Dersom de ikke har det, utløser det ulike rettigheter til særskilt språkopplæring, som er særskilt norskopplæring, morsmålsundervisning og/eller tospråklig fagopplæring. I praksis tilfaller dette ansvaret på skolene. Mitt søkelys i oppgaven er retten til særskilt norskopplæring, siden det er det som gis i innføringsklasser i grunnskolen og den videregående skolen.

Det er ingen føringer for *hvordan* kommunene, eller skolene, skal kartlegge elevene, heller ikke hvilke kartleggingsmateriell som skal brukes. I dag er det mange ulike kartleggingsprøver å velge mellom, og skolene kan også velge å lage sine egne kartleggingsprøver. Det er også lite informasjon om kvaliteten på de kartleggingsprøvene som brukes.

Utvalget vil bestå av lærere fra barneskoler, ungdomsskoler og videregående skoler i hele det tidligere fylket. Undersøkelsen er helt frivillig, og lærerne velger selv om de vil svare på spørreundersøkelsen, som også er anonym.

Problemstillingen for prosjektet er «*Hvilke typer kartleggingsprøver brukes i innføringstilbud i tidligere Hordaland fylke, og hvilken kvalitet har måleinstrumentene som brukes?*»

Hvem er ansvarlig for forskningsprosjektet?

Høgskulen på Vestlandet, campus Sogndal, er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?

Utvalget av skoler er valgt ut gjennom søk på skolenes nettsider, og er sendt skoler som har innføringstilbud og som er lokalisert i tidligere Hordaland fylke.

Hva innebærer det for deg å delta?

Hvis du velger å delta i prosjektet, innebærer det at du fyller ut et spørreskjema. Det vil ta deg 1-2 minutter og inneholder 4 spørsmål. Dine svar fra spørreskjemaet blir registrert elektronisk.

Det er frivillig å delta

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle dine personopplysninger vil da bli slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrevet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

- De som vil ha tilgang til opplysningene er student Hege Stamnesfet (HVL) og veiledere Arne Ola Lervåg (UiO) og Hilde Hofslundsengen (HVL).
- Databehandlingsprogram som skal brukes for å samle inn, bearbeide og lagre data: SurveyXact
- Uvedkomne vil ikke ha tilgang til personopplysningene ved at programmet vi bruker, SurveyXact, til å sende ut spørreskjemaet erstatter dine personopplysninger med en kode som lagres kryptert på en forskningsserver.

Ingen deltakere vil kunne gjenkjennes i publikasjoner. Opplysninger som publiseres i mastergradsoppgaven vil være samlede og grupperte tall fra hele utvalget som beskriver hvilke kartleggingsprøver som brukes, hva de brukes til og hvor ofte de brukes. Ingen informasjon vil kunne spores tilbake til den enkelte skole eller lærer.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Opplysningene anonymiseres når prosjektet avsluttes/oppgaven er godkjent, noe som etter planen er juni 2021.

Dine rettigheter

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

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

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra *Høgskulen på Vestlandet* har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Hvor kan jeg finne ut mer?

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

- Høgskulen på Vestlandet ved
 - Student: Hege Stamnesfet, tlf: 477 54 054, e-post: 575333@stud.hvl.no
 - Veileder: Arne Ola Lervåg, tlf: 22 84 47 01, e-post: a.o.lervag@iped.uio.no
 - Veileder: Hilde Hofslundsengen, tlf: 57 67 60 59, e-post: Hilde.Hofslundsengen@hvl.no
 - Vårt personvernombud: *Trine Anniken Larsen*, tlf: 55 58 76 82 / 913 65 920, e-post: Trine.Anikken.Larsen@hvl.no

Hvis du har spørsmål knyttet til NSD sin vurdering av prosjektet, kan du ta kontakt med:

- NSD – Norsk senter for forskningsdata AS på epost (personvertjenester@nsd.no) eller på telefon: 55 58 21 17.

Med vennlig hilsen

Arne Ola Lervåg og Hilde Hofslundsengen
Prosjektansvarlig

Hege Stamnesfet
Student

(Forsker/veileder)

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet «Kartleggingsverktøy i innføringstilbud», og har fått anledning til å stille spørsmål.

Ved å svare på denne spørreundersøkelsen, samtykker jeg til

- å delta i spørreundersøkelsen
- at mine opplysninger behandles frem til prosjektet er avsluttet

Jeg har mottatt og forstått informasjon om prosjektet «Kartleggingsverktøy i innføringstilbud», og har fått anledning til å stille spørsmål. Ved å svare på denne spørreundersøkelsen, samtykker jeg til:

- å delta i spørreundersøkelsen og at mine opplysninger behandles frem til prosjektet er avsluttet

Vedlegg 4: Godkjenning fra NSD

Melding 13.11.2020 08:02

Det innsendte meldeskjemaet med referansekode 140298 er nå vurdert av NSD.

Følgende vurdering er gitt:

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 13.11.2020, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

MELD VESENTLIGE ENDRINGER Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde:

nsd.no/personvernombud/meld_prosjekt/meld_endringer.html Du må vente på svar fra NSD før endringen gjennomføres.

TYPE OPPLYSNINGER OG VARIGHET Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 30.06.2021.

LOVLIG GRUNNLAG

Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake. Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER

NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om: - lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen - formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke viderebehandles til

nye uforenlige formål - dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet - lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER

Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20). NSD vurderer at informasjonen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13. Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER

NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

SurveyXact er databehandler i prosjektet.

NSD legger til grunn at behandlingen oppfyller kravene til bruk av databehandler, jf. art 28 og 29. For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og eventuelt rådføre dere med behandlingsansvarlig institusjon.

OPPFØLGING AV PROSJEKTET NSD

vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!

Kontaktperson hos NSD: Jørgen Wincentsen Tlf. Personverntjenester: 55 58 21 17
(tast 1)

Vedlegg 5: Veiledende normering av *FLORO*

Under er normeringsgrunnlaget for *FLORO*. Alle tabeller er hentet fra instruksjonsheftet til testen (Utdanningsdirektoratet & NAFO, 2013b). Normeringsgrunnlaget for ordminneprøven på norsk er i tabell 8, for tyrkisk i tabell 8b, polsk i tabell 8c og somali i tabell 8d. I tabell 8a finner man gjennomsnittsverdiene for hvert enkelt språk.

Trinn	1	2	3	5	7	9	VGS
Antall = 245	44	39	37	57	29	estimat	39
Gjennomsnitt	2,5	2,7	2,9	3	3,2	3,3	3,4
Median	2,5	2,75	3	3	3	3,2	3,4
Standardavvik	0,4	0,4	0,4	0,3	0,5	0,4	0,4
Veiledende bekymrings-grense (ca. 1,75 standardavvik)	1,8	2	2,2	2,5	2,5	2,6	2,7

Trinn	1	2	3	5	6.-7.	9.-10.	VGS
Norsk N=245	2,5	2,7	2,9	3	3,2		3,4
Tyrkisk N=141	2,6	2,8	3,2	3,1	3	3,3	
Somali N=60				2,8	3		3
Albansk N=37 (+23*)		2,5		3	3,2		
Polsk N=65		2,7	2,9	3	3,4		3,5
Arabisk N=31		2,6		3			3,4
Thai N=23			2,8		3	3,4	

*tall fra tidligere utprøvinger av ordminneprøven

Trinn	1	2	3	5	6.-7.	9.-10.
Antall = 141	23	30	23	31	13	21
Gjennomsnitt	2,6	2,8	3,2	3,1	3	3,3
Median	2,3	2,9	3,3	3	3	3,4
Standardavvik	0,6	0,5	0,6	0,5	0,3	0,5

Trinn	2	3	4	5	6	7	8.-VGS
Antall = 65	6	11	9	13	16	9	9
Gjennomsnitt	2,7	2,9	2,9	3,1	3,2	3,4	3,5
Median	2,6	2,9	2,9	3	3,3	3,5	3,5
Standardavvik	0,4	0,5	0,3	0,5	0,3	0,4	0,7

Trinn	1	2.-3.	5	6.-7.	9.-10	VGS
Antall = 60		7	8	16		29
Gjennomsnitt			2,8	3		3
Median			2,8	3		2,9
Standardavvik			0,3	0,4		0,4

Tabell 9 viser gjennomsnitt, standardavvik og veiledende grense i RAN-testen (Rapid Automated Naming) på norsk.

Tabell 9		FLORO RAN (norsk)					
Trinn		Tall A	Tall B	Bokstav A	Bokstav B	Bilder A	Bilder B
Trinn 1	Gjennomsnitt i sekunder	34,01	34,26	38,82	35,55	34,31	37,65
	Standardavvik	17,67	12,04	21,36	13,74	12,93	18,09
	Veil. Grense	ca. 70	ca. 70	ca. 80	ca. 70	ca.70	ca.70
Trinn 2	Gjennomsnitt i sekunder	22,5	24,45	22,7	32,31	27,94	28,21
	Standardavvik	5,1	5,42	5,84	6,07	5,91	7,41
	Veil. Grense	ca. 32	ca. 34	ca. 34	ca. 44	ca. 48	ca. 41
Trinn 3	Gjennomsnitt i sekunder	20,5	21,95	20,22	21,39	24,38	24,46
	Standardavvik	6,88	6,35	7,72	7,74	4,61	5,69
	Veil. Grense	ca 32	ca. 32	ca. 34	ca. 36	ca. 31	ca. 35
Trinn 5	Gjennomsnitt i sekunder	14,9	16,5	13,9	15,3	20,3	20,3
	Standardavvik	2,2	2,7	3,4	2,4	5,1	4,3
	Veil. Grense	ca. 19	ca. 21	ca. 21	ca. 21	ca. 30	ca. 30
Trinn 7	Gjennomsnitt i sekunder	13,4	14	12,8	13,7	18,9	18,6
	Standardavvik	2,1	1,9	2,2	2,7	2,9	4,6
	Veil. Grense	ca.18	ca. 18	ca. 17	ca. 19	ca. 25	ca. 28
VGS	Gjennomsnitt i sekunder	11,44	12,15	11,33	11,84	15,62	14,56
	Standardavvik	1,7	2,1	1,8	2,3	3	3
	Veil. Grense	ca.15	ca. 16	ca. 14	ca. 17	ca. 22	ca. 22

Tabell 10 viser gjennomsnitt, standardavvik og veiledende grense for 2.-6.trinn og VGS i Ordrepetisjonstesten på norsk.

Tabell 10		FLORO Ordrepetisjon (norsk)	
Trinn 2.-6. N=30	Gjennomsnitt	22,4	
	Standardavvik	2,75	
	Veiledende grense 1,75 standardavvik	18 og under	
VGS N=20	Gjennomsnitt	24	
	Standardavvik	3	
	Veiledende grense 1,75 standardavvik	19 og under	

Vedlegg 6: Normering av TOSP

I tabell 11 er normeringsgrunnlaget for TOSP. Informasjonen i tabellen er hentet fra instruksjonsheftet til testen (Loona, 2008a). Det spesifiseres ikke et nøyaktig tall på hvor mange elever det er på hvert trinnsom ble kartlagt, men at hvert av trinnene 3., 5. og 7. kartla 150-177 elever fordelt på fem skoler i Skien kommune og at elevene kommer fra familier med ulik sosioøkonomisk bakgrunn.

Normeringsgrunnlaget for 9.trinn forklares ikke i instruksjonsheftet.

Stamineskårenes fordeling er som følger: Prøveklasse 1 = 11 %; Prøveklasse 2 = 17 %; Prøveklasse 3 = 44 %; Prøveklasse 4 = 17 %; Prøveklasse 5 = 11 % (Loona, 2008a, s.16). Prøveklasse 1 er de svakeste resultatene, mens prøveklasse 5 er de sterkeste resultatene.

Tabell

11 Normering av TOSP

Trinn		Gjennom snitt	Standard avvik	Stamine-skårer				
				Prøve- klasse 1	Prøve- klasse 2	Prøve- klasse 3	Prøve- klasse 4	Prøve- klasse 5
3.trinn	Informasjon	21,31	3,69	0-15,5	16-19	19,5-24	24,5-25,5	26-30
	Ord og begrepsrikdom	46,49	5,22	0-40	41-43	44-50	51-52	53-57
	Ord- gjenkjenning	56,54	2,25	0-53	54-55	56-58	59	60
	Ting som hører sammen	51,2	9,69	0-40	41-44	45-56	57-65	66-90
	Lesehastighet	64,06	25,43	0-36	37-50	51-72	73-101	102-
	Leseforståelse	10,68	3,33	0-6	7-8.	9-13.	14	15
	5.trinn	Informasjon	25,13	2,84	0-21,5	22-23,5	24-27	27,5-28
Ord og begrepsrikdom		50,38	4,41	0-45	46-48	49-54	54,5-55	55,5-57
Ord- gjenkjenning		58,39	1,47	0-56	57-58	58,5-59,5	60	60
Ting som hører sammen		59,65	8,2	0-48	49-53	54-64	65-69	70-78
Lesehastighet		113,51	36,88	0-73	74-93	94-127	128-170	171-226
Leseforståelse		13,41	1,9	0-10	11.des	13-14	15	15
7.trinn		Informasjon	17,14	4,07	0-12	12,5-14,5	15-20,5	21-22
	Ord og begrepsrikdom	39,79	4,96	0-32	33-37	38-43	43,5-45	46-51
	Ord- gjenkjenning	34,6	2,36	0-31	32-33	34-36	36,5-37	38-40
	Ting som hører sammen	41,64	10,23	0-27	28-35	36-48	49-54	55-70
	9.trinn	Informasjon	18,2	3,97	0-13,5	13,5-16	16-21	21-23
Ord og begrepsrikdom		41,1	4,5	0-35	35-39	39-44	44-47	47-52
Ord- gjenkjenning		34,9	3,8	0-32	32-34	34-37	37-38	38-40
Ting som hører sammen		44,5	11	0-29	29-38	38-53	53-58	58-75

Vedlegg 7: Det felles europeiske rammeverket for språk

Tabell 12 viser den generelle beskrivelsen av de ulike nivåene i *Rammeverket* mens tabell 13 viser mer detaljerte beskrivelser innenfor flere språkområder. Tabellene er hentet fra *Det felles europeiske rammeverket for språk* (Utdanningsdirektoratet, 2011, s. 28, s.32-33).

Tabell 12	Global skala	
Basisbruker	A1	Kan forstå og bruke kjente, dagligdagse uttrykk og svært enkle utsagn for å dekke grunnleggende behov. Kan presentere seg selv og andre og kan stille og svare på personlige spørsmål om f.eks. bosted, venner og bekjente og eiendeler. Kan delta i en samtale på en enkel måte hvis samtalepartneren snakker langsomt og tydelig og er innstilt på å hjelpe.
	A2	Kan forstå setninger og vanlige uttrykk knyttet til de viktigste områdene av dagliglivet, f.eks. svært enkel informasjon om en selv og familien, innkjøp, nærmiljø og arbeidsliv. Kan klare seg i enkle og rutinepregede samtalsituasjoner med direkte utveksling av informasjon om kjente og rutinepregede forhold. Kan med enkle ord beskrive visse sider ved sin egen bakgrunn og sitt nærmiljø og grunnleggende personlige behov
Selvstendig bruker	B1	Kan forstå hovedpunktene i klar, standard tekst og tale om kjente emner som en ofte møter i forbindelse med arbeid, skole, fritid osv. Kan klare seg i de fleste situasjoner som kan oppstå når en reiser i et område der språket snakkes. Kan skrive enkle, sammenhengende tekster om emner som er kjente eller av personlig interesse. Kan beskrive opplevelser og hendelser, drømmer, håp og planer, og kort forklare og begrunne meninger og planer
	B2	Kan forstå hovedinnholdet i komplekse tekster om både konkrete og abstrakte emner, også faglige drøftinger innenfor ens eget fagområde. Kan delta i samtaler med et så spontant og flytende språk at kommunikasjonen med morsmålsbrukere ikke blir anstrengende for noen av partene. Kan skrive klare, detaljerte tekster om et vidt spekter av emner, forklare et synspunkt på en aktuell sak og argumentere for og imot ulike alternativer.
Avansert bruker	C1	Kan forstå et bredt spekter av lengre, krevende tekster og oppfatte budskap som ikke er direkte uttrykt. Kan uttrykke seg flytende og spontant uten at det merkes noe særlig at en leter etter uttrykksmåter. Kan bruke språket fleksibelt og hensiktsmessig til sosiale, akademiske og yrkesrelaterte formål. Kan skrive klare, velstrukturerte og detaljerte tekster om komplekse emner og vise at en mestrer ulike setningsmønstre, bindeord og sammenbindende markører.
	C2	Kan uten problemer forstå praktisk talt alt en hører eller leser. Kan sammenfatte informasjon fra ulike muntlige og skriftlige kilder ved å gjengi argumenter og redegjørelser på en sammenhengende måte. Kan uttrykke seg spontant, med svært god flyt og presisjon, og få fram finere meningsnyanser selv i mer komplekse situasjoner.

Tabell 13 Skjema for egenvurdering

		Forståelse		Tale		Skrift
		Lytting	Lesing	Samhandling	Muntlig produksjon	Skriftlig produksjon
Basisbruker	A1	Jeg kan kjenne igjen vanlige ord og svært enkle uttrykk som har med meg selv, min familie og mine nære omgivelser å gjøre, når man snakker langsomt og tydelig.	Jeg kan oppfatte kjente navn og forstå ord og svært enkle setninger, for eksempel i annonser, på plakater eller i kataloger	Jeg kan delta i en samtale på en enkel måte hvis samtalepartnern er villig til å gjenta, omformulere, snakke langsomt og hjelpe meg med å formulere meg. Jeg kan stille og svare på enkle spørsmål knyttet til viktige behov eller til svært kjente emner	Jeg kan bruke enkle uttrykk og setninger for å beskrive stedet der jeg bor, og mennesker jeg kjenner.	Jeg kan skrive et enkelt postkort, f.eks. et julekort. Jeg kan fylle ut skjema med personlige opplysninger, som navn, nasjonalitet og adresse.
	A2	Jeg kan forstå vanlige ord og ytringer knyttet til områder som er viktige for meg personlig (f.eks. svært enkel informasjon som gjelder meg selv, min familie, nærmiljø, innkjøp og arbeidsliv). Jeg kan oppfatte hovedinnholdet i enkle, korte og tydelige beskjeder og meldinger.	Jeg kan lese svært korte, enkle tekster. Jeg kan finne konkrete, forutsigbare opplysninger i enkel, dagligdags informasjon som reklame, brosjyrer, menyer og rutetabeller. Jeg kan forstå korte, enkle, personlige brev.	Jeg kan delta i enkle og rutinepregede samtaler som innebærer enkel og direkte utveksling av informasjon om kjente emner og aktiviteter. Jeg kan bruke og forstå vanlige høflighetsfraser, men forstår vanligvis ikke nok til å holde samtalen gående.	Jeg kan bruke en rekke uttrykk og setninger for på en enkel måte å beskrive familie og andre mennesker, boforhold, min utdanningsbakgrunn, og hva jeg jobber eller har jobbet med.	Jeg kan skrive, korte, enkle beskjeder som gjelder de viktigste områdene av dagliglivet. Jeg kan skrive et svært enkelt personlig brev, f.eks. et takkebrev.
Selvstendig bruker	B1	Jeg kan forstå hovedinnholdet i klar, standardtale om kjente emner knyttet til arbeid, skole og fritid osv. Jeg kan forstå hovedinnholdet i mange radio- og TVprogram om aktuelle saker eller om emner som er av personlig eller faglig interesse for meg, når man snakker relativt langsomt og tydelig.	Jeg kan forstå tekster der språket i stor grad er preget av frekvente, dagligdags ord og uttrykk eller er knyttet til arbeidet mitt. Jeg kan forstå personlige brev med beskrivelser av hendelser, følelser og ønsker.	Jeg kan klare meg språklig i de fleste situasjoner som kan oppstå når man reiser i et område der språket snakkes. Jeg kan uforberedt delta i samtaler om kjente emner, emner av personlig interesse eller emner som er viktige i dagliglivet (f.eks. familie, hobby, arbeid, reise, aktuelle hendelser).	Jeg kan gi en enkel, sammenhengende framstilling av hendelser, erfaringer, drømmer, håp og planer. Jeg kan kort forklare og begrunne mine meninger og planer. Jeg kan fortelle en historie eller gjennfortelle handlingen i en bok eller film og beskrive reaksjonene mine.	Jeg kan skrive enkle, sammenhengende tekster om emner som er kjente eller av personlig interesse. Jeg kan skrive personlige brev der jeg beskriver opplevelser og inntrykk.

	B2	Jeg kan forstå lengre, sammenhengende tale og foredrag og følge komplekse argumentasjonsrekker når emnet er relativt kjent. Jeg kan forstå de fleste nyhets- og aktualitets-program på TV. Jeg kan forstå de aller fleste filmer der man bruker standard regionaltalemål.	Jeg kan lese artikler og rapporter om aktuelle problemstillinger der forfatteren gir uttrykk for spesielle holdninger eller synspunkter. Jeg kan forstå prosalitteratur fra vår egen tid.	Jeg kan bruke språket så flytende og spontant at samtaler med morsmålsbrukere ikke byr på særlige problemer. Jeg kan delta aktivt i diskusjoner i kjente sammenhenger, og jeg kan uttrykke og begrunne synspunktene mine.	Jeg kan gi klare og detaljerte beskrivelser innenfor et vidt spekter av emner knyttet til mitt interessefelt. Jeg kan forklare et synspunkt på en gitt problemstilling og gjøre rede for fordeler og ulemper ved ulike alternativer.	Jeg kan skrive klare, detaljerte tekster om et vidt spekter av emner knyttet til mine interesser. Jeg kan skrive en sammenhengende framstilling eller rapport for å formidle informasjon eller argumentere for eller imot et bestemt synspunkt. Jeg kan skrive brev som tydelig viser hvilken betydning visse hen-delser og erfaringer har for meg.
Avansert bruker	C1	Jeg kan forstå lengre, sammenhengende tale også når den ikke er tydelig strukturert, og når deler av innholdet er underforstått og ikke kommer tydelig fram. Jeg kan forstå fjernsynsprogram og filmer uten altfor store anstrengelser.	Jeg kan forstå lange, komplekse faktatekster og skjønnlitterære tekster og legge merke til trekk som er typiske for ulike stilnivå. Jeg kan forstå fagartikler og lengre, tekniske instruksjoner selv når de ligger utenfor mitt felt.	Jeg kan uttrykke meg flytende og spontant uten at det i særlig grad merkes at jeg leter etter uttrykksmåter. Jeg kan bruke språket fleksibelt og effektivt både i sosiale og faglige sammenhenger. Jeg kan formulere ideer og meninger presist og forholde meg til samtalepartnerne på en hensiktsmessig måte.	Jeg kan gi klare og detaljerte beskrivelser av komplekse emner der jeg trekker inn delmomenter, utdyper enkelte punkter og runder av med en passende konklusjon.	Jeg kan skrive klare og strukturerte tekster og gi uttrykk for mine syns-punkter på en utfyllende måte. Jeg kan skrive om komplekse emner (i brev, rapporter, artikler eller skolestiler) og understreke hva jeg mener er de viktigste punktene. Jeg kan velge en stil som passer til den som skal lese teksten.

<p>C2 Jeg forstår uten problemer alle former for muntlig språk både i det virkelige liv og i mediene, også når morsmålsbrukere snakker i et naturlig og raskt tempo, dersom jeg får litt tid til å gjøre meg kjent med dialekten.</p>	<p>Jeg leser uten problemer praktisk talt alle former for skriftlig språk, også tekster om abstrakte emner og tekster som er komplekse når det gjelder språk eller struktur, som instruksjoner, fagartikler og litterære verk.</p>	<p>Jeg kan uten anstrengelse delta i alle former for samtale og diskusjon og har god kjennskap til idiomatiske uttrykk og vanlige talemåter. Jeg kan uttrykke meg flytende og formidle finere meningsnyanser på en presis måte. Hvis jeg får et språklig problem, kan jeg omformulere meg slik at det nesten ikke merkes.</p>	<p>Jeg kan beskrive og drøfte/ argumentere på en klar og flytende måte, bruke et språk eller en stil som passer i sammenhengen, og strukturere det jeg sier, på en hensiktsmessig og logisk måte, slik at den som lytter, kan legge merke til og huske viktige poeng.</p>	<p>Jeg kan skrive klare tekster med god flyt i en passende stil. Jeg kan skrive komplekse brev, rapporter og artikler der jeg legger fram en sak på en hensiktsmessig, logisk og strukturert måte som hjelper mottakeren til å legge merke til og huske viktige poeng. Jeg kan skrive sammendrag og anmeldelser både av faglige og litterære arbeider.</p>
---	--	---	---	--
