


Tom Hilding Skoglund

# Egnet som **POLITI?**



Seleksjonsmetodenes  
prediktive validitet ved  
PHS sin egnethetsvurdering

Tom Hilding Skoglund

# Egnet som **POLITI?**

Seleksjonsmetodenes  
prediktive validitet ved PHS sin egnethetsvurdering

© Politihøgskolen, Oslo 2018

PHS Forskning 2018:1

ISBN 978-82-7808-132-7 (trykt utgave)

ISBN 978-82-7808-133-4 (elektronisk utgave)

ISSN 0807-1721

Det må ikke kopieres fra denne boka i strid med åndsverkloven og fotografiloven eller i strid med avtaler om kopiering inngått med Kopinor, interesseorgan for rettighetshavere til åndsverk.

Alle henvendelser kan rettes til:

Politihøgskolen

Forskningsavdelingen

Slemdalsvn. 5

Postboks 5027, Majorstuen

0301 Oslo

[www.phs.no](http://www.phs.no)

Omslag/foto: Eileen Schreiner Berglie/colourbox

Sats og layout: Eileen Schreiner Berglie

Trykk: Staples

# Innhold

<b>Forord</b>	<b>5</b>
<b>Sammendrag</b>	<b>7</b>
<b>1 INNLEDNING</b>	<b>9</b>
1.1 Om politiets rolle	11
1.2 Om Politihøgskolen	12
1.2.1. Fem hovedområder i politibachelorutdanningen	12
1.2.2. Pedagogisk modell	13
1.2.3. Utvikling av kompetanser for å vurdere personlig egnethet	14
1.3 Opptaksprosessen	15
1.3.1. Medisinske krav	15
1.3.2. Skikkethetsvurdering	16
1.3.3. Opptaksnemdene og vurdering av personlig egnethet	16
1.4 Formålet med undersøkelsen	18
1.4.1. Problemstilling	19
<b>2 Forskning med relevans for Politihøgskolens seleksjonsmodell</b>	<b>20</b>
2.1 Internasjonale meta-analyser	21
2.2 Studier fra Forsvaret	23
2.3 Prediktiv validitet ved den svenske politiseleksjonen	25
2.4 Evaluering fra Agenda Kaupang	28
<b>3 Metode</b>	<b>30</b>
3.1 Utvalg	30
3.2 Prosedyre	30
3.3 Prediktorer	31
3.3.1. Karakterpoeng	31
3.3.2. Psykologiske tester fra cut-e AS	31
3.3.3. Opptaksnemndenes vurderinger	35
3.4 Kriterier	36
3.4.1. Kriterier for kull 2013	37
3.4.2. Kriterier for kull 2014	38
3.5 Statistiske analyser	39

<b>4</b>	<b>Resultater</b>	<b>41</b>
4.1	Kull 2013	41
4.1.1.	Korrelasjonsanalyse	41
4.1.2.	Multipel hierarkisk regresjonsanalyse	42
4.1.3.	Egnethetsvurdering og integritet	43
4.2	Kull 2014	44
4.2.1.	Korrelasjonsanalyse	44
4.2.2.	Multipel hierarkisk regresjonsanalyse	45
4.2.3.	Egnethetsvurdering og integritet	46
4.2.4.	Egnethetsvurdering og fullføring av studiet på normert tid	46
<b>5</b>	<b>Diskusjon</b>	<b>48</b>
5.1	Egnethetsvurderingen og faglige prestasjoner på grunnutdanningen	49
5.1.1.	Videregående skolekarakterer	49
5.1.2.	Psykologiske tester	50
5.1.3.	Intervjuet	54
5.1.4.	Gruppeøvelsen	56
5.2	Egnethetsvurderingen og bedømmelsen av integritet	56
5.3	Egnethetsvurderingen og fullføring av studiet på normert tid	58
5.4	Begrensninger ved undersøkelsen	59
5.5	Styrker ved undersøkelsen	62
<b>6</b>	<b>Konklusjon</b>	<b>64</b>
<b>7</b>	<b>Referanseliste</b>	<b>67</b>
	<b>Vedlegg</b>	<b>71</b>

# Forord

Hva gjør at noen er egnet til en jobb, mens andre ikke? Jeg har tidligere hatt gleden av å undervise i psykologi ved Politihøgskolen, og ble etter hvert nysgjerrig på hva som gjør at noen blir bedømt som dyktige politifolk, mens det for andre var spørsmål om hvorfor de i det hele tatt ble vurdert som egnet. De siste årene har jeg jobbet med seleksjon i Forsvaret, og erfart ulike opptaksprosedyrer ved ulike avdelinger. I begge etater høres ofte mange meninger om hvordan man skal selektere operativt personell, og ikke rent sjelden er det svært ulike meninger ute og går. Nesten alle mener seleksjon er viktig, og er enige i at rekruttering, seleksjon og opplæring utgjør grunnpilarene i den operative evnen.

Seleksjon handler om prediksjon, og er en øvelse i beslutningstaking om mennesker – av og til basert på mindre informasjon enn det man ønsker. Ofte er det en slags kamp mot mangfoldet, i den tro at en utdanningsposisjon eller jobb ikke passer for alle. Samtidig vil gode metoder for å identifisere egnethet og talent begrense feilaktige konklusjoner om andre basert på forutinntatte holdninger og stereotype forestillinger. Slik sett er profesjonell seleksjon en kamp mot overdrevent subjektivt skjønn hos de som selekterer.

Takk til Politihøgskolens rektor Nina Skarpenes og leder for forskningsavdelingen Haavard Magne Reksten som lot en ekstern få mulighet til å studere den nye egnethetsvurderingen som trådte i kraft i 2012. Takk til dere på Studieavdelingen som alltid var støttende, og helt nødvendige sparringspartnere og ressurser for at undersøkelsen kunne realiseres. Takk også til cut-e Norge AS som fra første stund var imøtekommende og behjelpelige.

En takk går også til sjefpsykolog i Forsvaret Ole Christian Lang-Ree og professor i psykologi Monica Martinussen for god fagveiledning knyttet til denne rapporten. Avslutningsvis vil jeg takke kolleger i Forsvaret og politiet for fruktbare løpende diskusjoner om seleksjon, i mer eller mindre (u)formell form.

Hovedhensikten med rapporten er å tilkjenne funnene i en fundert og strukturert form. Faguavhengig lesbarhet har vært en målsetning, men enkelte avsnitt vil nok fremstå som tyngre enn andre. Rapporten er ugradert, og slik kunne nå ut til de som måtte ha interesse av den utover Politihøgskolen og cut-e AS.

Rapporten er bygget opp ved hjelp av kapitler som er vanlige i en kvantitativ forskningsrapport. Mer tekniske og formaliserte forskningsartikler basert på datamaterialet vil eventuelt bli til ved en senere anledning.

Tom H. Skoglund  
Oslo, desember 2017

# Sammendrag

Politiutdanningen har i mange tiår hatt et mer omfattende opptakssystem enn det som har vært vanlig ved universitets- og høyskoleutdanninger i Norge, hvor søkerne blant annet skal vurderes for personlig egnethet. I 2012 ble det innført en mer fundert, strukturert og forskningsforankret egnethetsvurdering enn den som gjaldt tidligere. Vurderingen skal nå ta stilling til nivået på kompetansene *åpenhet og inkludering, modenhet, handlekraft, samhandling, analytisk evne og integritet*. Disse seks kompetansene definerer innholdet i hva personlig egnethet er. Tre seleksjonsmetoder brukes til dette formålet: psykologiske tester, intervju og gruppeøvelse. I tillegg kommer informasjon fra skikkethetsvurderingen. Denne undersøkelsen tar sikte på å kartlegge den prediktive validiteten til de nye seleksjonsmetodene. Følgende problemstilling ble valgt:

*Viser grad av egnethetsvurderingen sammenheng med faglige prestasjoner på grunnutdanningen, bedømmelsen av integritet og om studiet fullføres på normert tid?*

Problemstillingen står i stil med tilgjengelige registerdata for kullene som begynte grunnutdanningen i 2013 og 2014. De faglige prestasjonene utgjorde de karaktergivende fagemnene. Skikkethetsvurderingen av søkerne resulterte i om kompetansen *integritet* var akseptabel eller ikke, samt at det for kull 2014 var mulig å ta ut data på hvem som ikke hadde fullført studiet innen normert tid. Tre hovedfunn trekkes ut av korrelasjons- og regresjonsanalyser:

*Første hovedfunn viser at karakterer fra videregående skole predikerer faglige prestasjoner i grunnutdanningen på en god måte. Egnethetsvurderingen treffer i riktig retning, men sammenhengene mellom seleksjonsmetodene og faglige prestasjoner er svake og til dels ikke-eksisterende. En økning i skårer*



*på egnethetsvurderingen indikerer derfor ikke en økning i senere faglige prestasjoner på en klar måte.*

*Andre hovedfunn viser at egnethetsvurderingen ikke har en klar sammenheng med integritetsbedømmelsen. Dette til tross for at det inngår en integritetstest/sårbarhetstest i egnethetsvurderingen (for kull 2014).*

*Tredje hovedfunn viser at egnethetsvurderingen ikke har en klar sammenheng med om studiet fullføres på normert tid eller ikke.*

Funnene har både likheter og ulikheter med andre tilsvarende studier, hvor det legges vekt på en internasjonal meta-analyse om seleksjon, studier fra Forsvaret og en undersøkelse om politistudent-opptaket i Sverige. En begrensning i undersøkelsen knyttet seg til manglende prestasjonsdata fra praksisåret B2, og fra det senere yrkeslivet. Samtidig var det en styrke at karakterene i de faglige prestasjonene på grunnutdanningen hvilte på ulike, men standardiserte, vurderingsformer, samt at løsning av mange eksamensoppgaver krevde en polisiær tilnærming.

Rapporten anbefaler ikke konkrete endringer av egnethetsvurderingen, men retter kritikk mot deler av egnethetsvurderingen basert på funnene. Spesielt kritiseres bruken av de psykologiske testene. Det poengteres at rapporten ikke først og fremst berører negativ seleksjon.

# 1 INNLEDNING

Personellseleksjon handler om å velge ut de best egnede og å unngå å velge de uegnede. Selve kongstanken i seleksjonsfeltet er at mennesker er ulike på en del egenskaper, og at trenbarheten av enkelte egenskaper er begrenset. Eksempelvis vil mange hevde at grunnleggende kognitive evner og personlighetstrekk er stabile egenskaper hos voksne (Furnham, 2008). Blant de som jobber innenfor dette feltet antas det at den enkelte søker til en utdanning eller jobb innehar personlige forutsetninger, eller et kompetansepotensial, som kan kartlegges. Disse individuelle egenskapene, som ofte kartlegges gjennom ulike psykologiske tester, intervjuer og praktiske øvelser, antas videre å være verdifulle når man skal forutsi hvor godt utdanningen eller jobben vil bli gjennomført. I tillegg antas det at kartlegging av individuelle forhold er verdifullt for å unngå å velge uegnet personell, en antagelse som har paralleller til mer medisinsk orienterte arbeidspsykologiske vurderinger. Det finnes flere norske oversiktsbøker på rekrutteringsområdet (e.g. Skorstad, 2015) og på uegnet personell i arbeidslivet (Kile, 1990), og temaet rekruttering og seleksjon har etter hvert blitt et populært fagområde, spesielt innenfor HR-fag og organisasjonspsykologien. Det har vært en enorm internasjonal forskningsproduksjon innen området, hvor omfangsrike meta-analyser har vist at en rekke psykologiske kjennetegn ved individet har en betydelig prediktiv verdi opp mot prestasjoner i utdanning og arbeid (Schmidt & Hunter, 1998; Schmidt, Oh, & Shaffer, 2016). Blant annet viser disse meta-analysene at nivåene på kognitivt evnenivå og personlighetstrekket samvittighetsfullhet ofte er tydelig relatert til det å få til en utdanning eller jobb på en god måte. Slike funn betyr imidlertid ikke at kvaliteten på utdanning og praksiserfaring ikke er viktige faktorer, eller at eksterne faktorer som eksempelvis arbeidsmiljø ikke påvirker prestasjoner. Kvaliteten på samspillsmønstre og ledelse fremheves ofte som betydningsfullt for jobbutførelse (Einarsen & Skogstad, 2012). Seleksjon må derfor sees i sammenheng med andre faktorer for å gi det beste personellet.

Både internasjonalt og nasjonalt er det militæret som har de lengste tradisjonene innenfor systematisk personellseleksjon (Laurence & Matthews, 2012). I Norge var Forsvaret i gang med sesjonsordningen ganske raskt etter andre verdenskrig, hvor det blant annet ble gjennomført mannskapsklassifisering basert på psykologiske tester. I årene som fulgte ble det et økende fokus på systematikk i utvelgelse av militærflygere og annet spesialisert personell (Hansen, 2006). Det er således ikke overraskende at det har vært produsert en del forskning innenfor norsk militærseleksjon, hvor det eksempelvis de siste årene har kommet flere masteravhandlinger som undersøker opptaket til befalsskole (Kjenstabakk, 2012), offisersutdanning (Vik, 2013) og Luftforsvarets flygerskole (Harlem, 2016; Svensson, 2013).

Mange vil mene at slektskapet mellom Forsvaret og politiet, som først og fremst går på opplæring i og muligheter for bruk av legitim voldsmakt, tilsier at det også bør stilles krav til politipersonell. Samtidig er det rimelig å anta at politiet har noen særegenheter i grunnutdanningen og yrkesutøvelsen, slik at en kopiering av Forsvarets seleksjonsmetoder ikke nødvendigvis vil være treffende. En kopiering har det heller ikke vært. Politihøgskolen (PHS) har i mange tiår gjennomført vurderinger av personlig egnethet og skikkethet hos søkerne sine, først og fremst gjennom fysiske tester, gjennomføring av intervju og kontroll av vandel. Fra år 2012 ble også psykologiske tester og en gruppeøvelse brukt, noe som var nytt for opptaksprosedyrene ved PHS.

I dette innledningskapitlet vil politiets rolle i samfunnet kort beskrives med utgangspunkt i lovverket og generalistmodellen. Videre gjennomgås innholdet i og oppbyggingen av politibachelorutdanningen. Deretter presenteres de seks kompetansene PHS legger vekt på i vurderingen av personlig egnethet ved opptak til utdanningen. Sentrale momenter ved opptaksprosessen beskrives, før formålet og problemstillingen til undersøkelsen presenteres.

## 1.1 Om politiets rolle

Politi-loven § 1 beskriver politiets rolle (Politi-loven, 1995):

Politi-et skal gjennom forebyggende, håndhevende og hjelpende virksomhet være et ledd i samfunnets samlede innsats for å fremme og befeste borgernes rettsikkerhet, trygghet og alminnelige velferd for øvrig.

Det kan kanskje være lett å assosiere politiyrket med tilspissede arbeidskrav innenfor en kontekst preget av farlige gjerningsmenn, uoversiktlige omstendigheter og uskyldige personer i behov for sikkerhet. Fra tid til annen forekommer en slik stridslignende arbeidskontekst, hvor konsekvensene av feilbeslutninger og feilhandlinger blir store. Samtidig gir beskrivelsen av politiets rolle assosiasjoner til andre yrkesscenarioer som ikke nødvendigvis bærer preg av en operativ kritikalitet, for eksempel det å gi tilpasset hjelp til mennesker i sårbare faser. Den varierte rollen til politiet, med sannsynlighet for operativ kritikalitet og nødvendig bruk av legitim voldsmakt, har nok gjort at vurdering av personlig egnethet har blitt ansett som like viktig i politiet som i Forsvaret. For å belyse videre hva som forventes av politipersonell, og at tanken om personlig egnethet for yrket strekker seg lenger enn rene operative kvaliteter, er det også nyttig å se at PHS skal utdanne «generalister» (Politi-høgskolen, 2014):

En generalist er en tjenestemann/-kvinne som besitter grunnleggende kunnskaper og ferdigheter i politiets forebyggende, kriminalitetsbekjempende og trygghetsskapende arbeid. I løsning av oppgavene skal generalisten ha kompetanse til å utføre grunnleggende politiarbeid, foreta helhetsvurderinger, se sitt arbeid i en bredere samfunnsmessig sammenheng og trekke inn relevant spesialkompetanse og samarbeidspartnere ved behov. Generalisten skal ha utviklet grunnlag for fortsatt læring og utvikling gjennom utøvelsen av yrket. (s. 2).

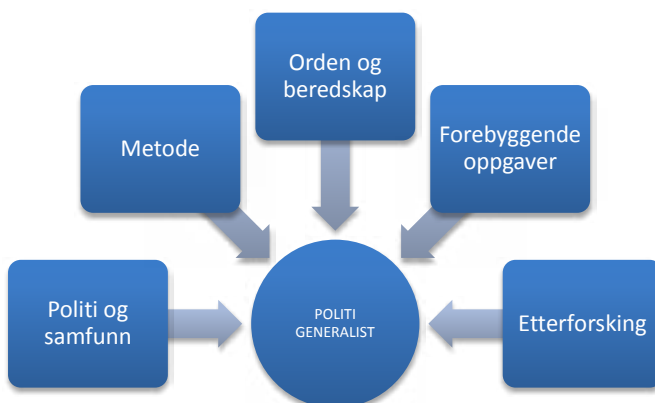
Ikke bare beskrives det her en omfattende ferdighetsforventning, men også en betydelig forventning om kunnskapen til politigeneralisten. Når løsning av arbeidsoppgavene skal innebære mer enn den isolerte operative oppdragsløsning, eksempelvis det og «se sitt arbeid i en bredere samfunnsmessig sammenheng», legger det føringer på hva politiutdanningen skal inneholde, og ikke minst hvilken kompetanseprofil som er ønskelig for personellet.

## 1.2 Om Politihøgskolen

Det er rettslig forankret at PHS er den sentrale utdanningsinstitusjonen for politiet (Politi-loven, 1995). Siden 1992 har PHS vært en treårig høgskoleutdanning, og fra 2004 av ble utdanningen akkreditert med en bachelorgrad (Hove, 2012). Skolen holder i skrivende stund til i Oslo, Stavern, Kongsvinger og Bodø. Siden 2010 har PHS tatt opp om lag 700 studenter på hvert kull, hvor det har vært litt over 4000 søkere i 2013 og 2014, og litt over 5000 søkere i 2015 (Politihøgskolen, 2016).

### 1.2.1. Fem hovedområder i politibachelorutdanningen

Utdanningen er bygget opp etter fem hovedområder som hver for seg inneholder mer spesifikke fagemner (Politihøgskolen, 2014). De fem hovedområdene gjenspeiler hovedoppgavene til en politigeneralist og er gjengitt i figur 1.1.



Figur 1.1: Hovedområdene i politibachelorutdanningen

Hovedområdet *Politi og samfunn* har til hensikt å utvikle samfunnsforståelsen til studentene, slik at de skal kjenne til overordnede forhold i og rundt egen organisasjon. Spesielt skal studentene forstå at politiet er del av den utøvende

makt, og at politiets inngripen skal ha hjemmel i lov. En forståelse av at kunnskap opparbeides gjennom systematisering av erfaring, gjennom rettsvitenskapelig tenkning og gjennom forskningsmetoder er kjernen i hovedområdet *Metode*. Studentene skal lære å utføre kunnskapsbasert politiarbeid, både gjennom skriftlige rapporteringer og i avgrensede oppdragsløsninger i politihverdagen. Hovedområdet *Orden og beredskap* skal gi studentene kunnskaper og ferdigheter for å løse ordensoppdrag. Hovedfokuset er «første enhet på åstedet», og studentene skal utvikle en nødvendig handlingskompetanse, samt bli bevisst egne holdninger og atferd, for å mestre ulike typer oppdrag. Det kriminalitetsforebyggende arbeidet er fokuset i hovedområdet *Forebyggende oppgaver*. Studentene skal kunne begrunne sine valg av ulike strategier og metoder for målrettet forebyggende arbeid. Straksetterforskning skal vektlegges i hovedområdet *Etterforskning*, hvor studentene skal kunne utføre både tekniske og taktiske strategier for å stanse, avdekke og etterforske straffbare forhold. Presisjon i muntlig og skriftlig kommunikasjon vektlegges.

### **1.2.2. Pedagogisk modell**

Studiet utgjør 180 studiepoeng, og det legges til rette for kobling mellom erfaringsbasert og forskningsbasert kunnskap gjennom alle tre år. Rammeplanen beskriver organiseringen av studiets tre år på følgende måte (Politi-høgskolen, 2014):

- Første studieår (B1) gjennomføres ved høgskolen. Studentene skal tilegne seg kunnskaper, ferdigheter og generell kompetanse i det enkelte emne, og gradvis bli i stand til å se tverrfaglige sammenhenger innen hovedområdene. B1 skal danne en faglig plattform for videre læring i praksisstudiet (B2).
- Andre studieår (B2) gjennomføres ved en opplæringsenhet i et politidistrikt. Studentene skal under veiledning utføre studiearbeid i henhold til læringsutbyttebeskrivelsene i fagplan for B2. I B2 skal studentene anvende tilegnet kunnskap fra B1, og tilegne seg ny kunnskap gjennom ulike typer politiarbeid i samhandling med erfarne polititjenestepersoner. Ferdighetstrening skal vektlegges spesielt i B2.

- Tredje studieår (B3) gjennomføres ved høyskolen. I dette studieåret vektlegges integrering av kunnskaper, ferdigheter og generell kompetanse innen hovedområder fra alle studieårene. Studentene skal bearbeide og videreutvikle erfaringer, samt teoretisk og praktisk kunnskap. Studentene skal gis mulighet for faglig videreutvikling, kritisk refleksjon og bevisstgjøring av holdninger og verdier. Det skal legges vekt på studentaktive læringsformer og samarbeidslæring.

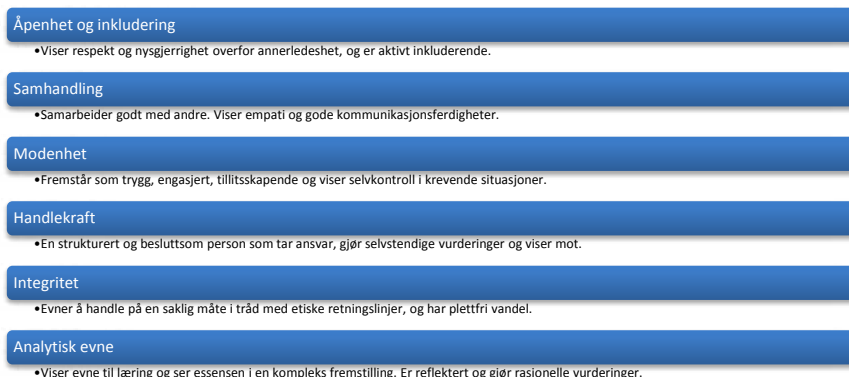
### 1.2.3. Utvikling av kompetanser for å vurdere personlig egnethet

PHS har i mange tiår hatt et mer omfattende opptakssystem enn det som har vært vanlig ved universitets- og høyskoleutdanninger i Norge. I 2012 ble det innført en mer omfattende vurdering av personlig egnethet enn den som tidligere gjaldt. Prosessen bak utviklingen av et nytt system for egnethetsvurdering er beskrevet av Abraham (2012). Tidligere baserte egnethetsvurderingen seg på en skriftlig egenerklæring fra søkeren, samt et 20 minutters langt panelintervju. Abraham (2012) påpeker at det i det tidligere opptakssystemet ikke var en felles forståelse av hva «personlig egnethet» egentlig innebar, samt at det var uklarthet i hvorvidt man primært selekterte til utdanning eller til selve polityrket.

I forlengelsen av disse observasjonene ble en omfattende jobbanalyseprosess iverksatt i samarbeid med konsulentselskapet cut-e Norge AS<sup>1</sup>. Dette arbeidet resulterte i formulering av kriterier, som ble kalt kompetanser, som ansees å være dekkende for å vurdere personlig egnethet. Så vidt forfatteren har forstått, ansees kompetansene som dekkende for å vurdere egnethet til så vel utdanning som til det senere politigeneralist-yrke. Begrepet «kompetanser» blir som regel i organisasjonspsykologien definert som de observerbare og målbare kunnskapene, ferdighetene, evnene og personlige egenskapene som en må ha for å kunne gjøre en jobb på en god måte; de er definert av kravene i jobben, og ikke nødvendigvis av kjennetegn på de som har jobben fra før (Mansikka, Harris, & Virtanen, 2017). Kompetansene til PHS er gjengitt i figur 1.2.

---

<sup>1</sup> cut-e AS har konsulenter med organisasjonspsykologisk spisskompetanse som leverer tjenester innenfor flere HR-felt (<http://www.cut-e.no/hjem/>)



Figur 1.2: Beskrivelse av kompetansene som egnethetsvurderingen baserer seg på.

Kompetansene er ytterligere operasjonalisert i manualer og observasjonsskjemaer opptaksnemndene benytter seg av. De mer konkrete operasjonaliseringene gjengis imidlertid ikke her på grunn av hensyn til testsikkerhet. Ved å gå ut med detaljer i seleksjonssystemer avsløres innholdet på et slikt nivå at søkere kan komme til å forberede seg på en uheldig måte før opptaket begynner.

### 1.3 Opptaksprosessen

Reglene om opptak til PHS er forankret i lov, hvor det forutsettes at søkeren har plettfriandel, oppfyller fastsatte formelle krav og fremstår som skikket for tjeneste i etaten (Politoloven, 1995). Sett bort fra fysiske opptakskrav, skiller det mellom medisinsk undersøkelse, skikkethetsvurdering og vurdering av personlig egnethet i opptaksprosessen. Selve opptaksprøvene knytter seg til fysiske tester og seleksjonsmetoder for egnethetsvurderingen.

#### 1.3.1. Medisinske krav

For å dokumentere innfrielse av medisinske krav må alle søkere sende inn komplett utfylt legeerklæring. I det medisinske regelverket (Politidirektoratet, 2005) nevnes en rekke somatiske forhold som diskvalifiserende, samt at «psykisk helsekrav og sosial tilpasning» beskrives i punkt 3.1:



«Pågående psykisk lidelse medfører udyktighet. Både aktuell og tidligere alvorlig sinnslidelse/psykotisk lidelse som schizofreni eller bipolar affektiv lidelse medfører udyktighet. Sykehistorie med annen psykisk lidelse som eksempelvis angstlidelse, depressiv lidelse eller spiseforstyrrelse medfører udyktighet dersom dette vurderes å kunne medføre betydningsfull forhøyet sårbarhet, og/eller at tjenesten ikke kan utføres tilfredsstillende. Alvorlig personlighetsforstyrrelse medfører udyktighet. Misbruk av medikamenter og/eller alkohol medfører udyktighet. Bruk av illegale substanser medfører udyktighet.»

Søkerne møter ingen direkte medisinsk evaluering ved selve opptaket, men ulike leger som leies inn av PHS gjennomgår erklæringene søkerne har sendt inn.

### **1.3.2. Skikkethetsvurdering**

Når det gjelder «plettfriandel» og «skikket for tjeneste i etaten» siktes det her til en skikkethetsvurdering av søkerne som skilles fra vurdering av personlig egnethet. PHS har hjemmel i politiloven §24 b til å innhente opplysninger om henlagte straffesaker eller andre forhold som kan ha betydning for søkerens skikkethet (Politiloven, 1995). En slik vurdering gjøres på alle søkere, uavhengig av resultat på øvrige formelle krav, og utføres av politipersonell. Skikkethetsvurdering gjøres via politiets sentrale systemer, og er et omfattende vurderingsarbeid. Vandel vurderes på grunnlag av uttømmende og utvidet politiattest, som søker selv sender inn i forbindelse med søknad om opptak. Vandelsmerknader knyttes til kompetansen *integritet* i egnethetsvurderingen, og denne informasjonen gis til opptaksnemdene.

### **1.3.3. Opptaksnemdene og vurdering av personlig egnethet**

Opptaksnemdene består av personellet som foretar den praktiske seleksjonen for å undersøke personlig egnethet. PHS har fire opptaksnemnder basert på geografisk tilhørighet. Hver nemnd har fire faste medlemmer med et varamedlem. Lederen skal være jurist med erfaring fra politiet, to medlemmer skal være politi med minst fem års erfaring, samt at siste medlem skal være ansatt i universitets- og høyskolesektoren. Både kvinner og menn skal være representert, samt at det er et mål at ett av medlemmene skal ha etnisk minoritetsbakgrunn. Nemndene møtes gjennom året i opplærings- og evaluerings-

øyemed. cut-e AS har en sentral rolle i opplæring av opptaksnemdene, samt at PHS har rådgivere med kompetanse innen rekruttering og seleksjon.

For å vurdere personlig egnethet bruker PHS tre metoder; psykologiske tester, gruppeøvelse og intervju. De psykologiske testene kartlegger ulike kognitive evner og språk, personlighetstrekk, samt integritet. Både intervjuet og gruppeøvelsen har en høy grad av struktur med tilhørende observasjonsmanualer på atferdsindikatornivå. På kompetansen *integritet* foretas det som nevnt over også en skikkethetsvurdering. Tabell 1.1 viser hvilke metoder som svarer til hvilke kompetanser.

Tabell 1.1: Tabellen viser hvilke seleksjonsmetoder som brukes for å måle de ulike kompetansene som egnethetsvurderingen baserer seg på

	Evne- og språktester	Integritets-test	Personlighets-test	Intervju	Gruppeøvelse
Åpenhet og inkludering				X	
Modenhet			X	X	X
Handlekraft			X	X	X
Samhandling	X		X	X	X
Analytisk evne	X				
Integritet (brukes også vandelskontroll)		X			

For den enkelte søker vil de psykologiske testene bli utført på forhånd før de kommer på opptaksdagen for intervju og gruppeøvelse. Når søkeren har gjennomgått det hele, møtes opptaksnemnda til felles diskusjon om hver enkelt søker, og tar i fellesskap en endelig beslutning om personlig egnethet. Til syvende og sist er det nemndas beslutning som er gjeldende. Rangering av søkere foretas etter forskrift fra Kunnskapsdepartementet, hvor karakter fra videregående skole, alder og tilleggs poeng vektet blant de personlig egnede søkerne. I metodekapitlet i denne rapporten vil seleksjonsmetodene gjennomgå mer detaljert.

## 1.4 Formålet med undersøkelsen

Den nye prosedyren for vurdering av personlig egnethet har gjort det mulig å gjennomføre en prediktiv valideringsstudie. En slik studie ser på sammenheng mellom seleksjonen og senere prestasjoner hos de som velges ut.

Opptaksnemda loggfører en tallverdi på kompetansene som vurderes. Resultatene fra de psykologiske testene er også loggført med tallverdier. Dette gjør det mulig å statistisk koble egnethetsvurderingen opp mot utdanningsresultatene (karakterer på en fem-delt skala; A-E). Antakelsen i undersøkelsen er formet slik at det er forventet at en økning i skårer på egnethetsvurderingen skal indikere en økning i senere prestasjoner. Dersom det er en klar sammenheng mellom rangeringen av individuelle egenskaper i seleksjonen og rangeringen i utdanningsresultatene, vil det indikere at seleksjonsmetodene virker. Det bør ikke være slik at differensieringen i seleksjonen ikke viser noen, eller i verste fall negativ, sammenheng med utdanningsresultatene. Dette innebærer at seleksjonsmetodene ikke kan forutsi hvem som vil klare utdanningen best.

I praksis er opptaksnemdene avhengig av et så godt som mulig beslutningsgrunnlag å vurdere egnethet ut fra. Det ligger også i dette at opptaksnemda må sørge for at ikke-egnede søkere ikke kommer inn. Dette kalles negativ seleksjon, og det er ikke nødvendigvis slik at svake skårer på definerte kompetanser dekker alle former for uegnethet (King, 2014). Her vil det være gråsoner opp mot den medisinske seleksjonen og det som angår «psykisk helsekraft og sosial tilpasning». Et perspektiv på hva som er fornuftige kriterier på negativ seleksjon og uønskede kandidater kan imidlertid være en undersøkelse i seg selv. Det er viktig å poengtere at denne undersøkelsen ikke berører temaet negativ seleksjon på PHS i særlig grad. Undersøkelsens statistiske analyser er rettet mot den positive seleksjonen som finner sted.

For å få et enda mer dekkende bilde på treffsikkerheten til egnethetsvurderingen, vil også enkelte tilleggsanalyser foretas. Utover seleksjonens evne til å forutsi utdanningsresultater, vil egnethetsvurderingen sees i sammenheng med utfallet av integritetsvurderingen, og med studiefullføring på normert tid. Sammenhengen mellom vurdering av egnethet og integritet vil riktignok

være et mål på samtidig kriterievaliditet, fordi disse vurderingene gjøres på samme tid.

Undersøkelsen gjennomføres ved hjelp av tilgjengelige registerdata for de to kullene som begynte sin utdanning i 2013 og 2014, heretter benevnt som kull 2013 og kull 2014. Disse to kullene var uteksaminert i henholdsvis 2016 og 2017.

#### **1.4.1. Problemstilling**

Undersøkelsen har ikke tilgang på data fra fysiske tester, medisinsk evaluering og innholdet i skikkethetsvurderingene for disse to kullene. I analysene som denne rapporten baserer seg på, er det data fra egnethetsvurderingen som benyttes.

Undersøkelsens problemstilling formuleres slik: Viser grad av egnethetsvurderingen sammenheng med faglige prestasjoner på grunnutdanningen, bedømmelsen av integritet og om studiet fullføres på normert tid?

## 2 Forskning med relevans for Politihøgskolens seleksjonsmodell

Innen generell personellseleksjon (Furnham, 2008; Ployhart, 2012; Schmidt et al., 2016), og innen politiseleksjon mer spesielt (Annell, 2015; Aamodt, 2004a), finnes det svært mye forskning. Dette kapitlet tar sikte på å presentere et lite utvalg av denne forskningen, for deretter å gå forholdsvis nøye gjennom funn som har relevans for Politihøgskolens seleksjonsmodell.

Internasjonalt finnes det omfangsrike meta-analyser som kartlegger den prediktive validiteten til ulike seleksjonsmetoder gjennom å summere opp den forskningen som er gjort basert på en rekke enkeltstudier. Nasjonalt finnes det en del seleksjonsstudier fra opptak til utdanninger og tjenester i Forsvaret, noe som er interessant på grunn av slektskapet mellom etatene.

Ut fra det forfatteren kjenner til har det ikke vært gjort prediktive valideringsstudier på PHS sitt opptakssystem tidligere. Derimot ble det i Sverige forholdsvis nylig avlagt en doktorgrad om seleksjonen til politiutdanningen der (Annell, 2015). De Nordamerikanske studiene om politiseleksjon som er oppsummert av Aamodt (2004a) er problematisk å sammenligne med, da politiutdanningene i Norge og Nord-Amerika kan være ganske så ulike.

I løpet av 2012 og 2013 ble det gjennomført evalueringer av opptaksprosessen ved PHS fra ulike aktører<sup>2</sup>. Disse rapportene dreier seg i hovedsak om kvalitative observasjoner omkring forbedringspunkter for et gryende seleksjonssystem. Ut fra forfatterens skjønn er rapporten fra Agenda Kaupang AS<sup>3</sup> spesielt grundig; formålet var å undersøke hvordan opptaksnemda oppfattet de nye seleksjonsmetodene, og om de nye metodene bidro til et bedre vurderingsgrunnlag for egnethet enn det som tidligere gjaldt (AgendaKaupang, 2012).

---

2 PHS internt ved studiekontoret, cut-e AS, Agenda Kaupang AS, og psykolog Dag Øyvind Engen Nilsen har alle evaluert opptaksordningen.

3 Agenda Kaupang er et konsulentselskap som arbeider med analyse, utredning og rådgiving innen områdene ledelse, styring, økonomi og organisasjonsutvikling (<http://agendakaupang.no/om-oss/om-agenda-kaupang/>).

I dette kapitlet vil relevante funn fra en omfattende internasjonal meta-analyse om seleksjon redegjøres for, før to studier fra Forsvaret blir presentert. Deretter vil funnene fra den svenske avhandlingen gjennomgås. Avslutningsvis vil hovedpoengene fra rapporten til Agenda Kaupang AS beskrives.

## 2.1 Internasjonale meta-analyser

Det er spesielt én meta-analyse som har vært svært innflytelsesrik i synet på «best practice» innen personellseleksjon. Artikkelen kom ut i 1998, og summerte opp 85 år med forskning på seleksjonsmetoder (Schmidt & Hunter, 1998). Nylig ble denne artikkelen fulgt opp, hvor forskning fra de siste 15 år til ble bakt inn i meta-analysen (Schmidt et al., 2016). Disse meta-analysene er svært omfattende og dokumenterer styrkene og begrensningene til mange seleksjonsmetoder på tvers av yrker og land.

Meta-analysene oppgir gjennomsnittlige validitetskoeffisienter basert på en rekke enkeltstudier og mange tusen deltakere. Dette er et statistisk mål på sammenhengen (korrelasjon) mellom prediktorer og kriterier, det vil si seleksjonsmetoder og mål på jobbutførelse eller prestasjoner i en opplærings-situasjon. Styrken på disse validitetskoeffisientene tallfestes, hvor 0 indikerer ingen sammenheng med kriteriene og 1 er en perfekt sammenheng. Den gjennomsnittlige korrelasjonen (validitetskoeffisienten) som oppgis i disse meta-analysene er ofte vektet i forhold til utvalgsstørrelsen i de respektive enkeltstudiene.

I tillegg oppgis ofte en korrigert gjennomsnittskorrelasjon mellom prediktoren og kriteriet. Dette innebærer at man korrigerer for usikkerhet i kriteriemålinger, og tar høyde for et fenomen som kalles begrenset spredning. Det første har å gjøre med at det er vanskelig å måle objektivt hva som er en god prestasjon i en jobb (eller av og til i en opplærings-situasjon), og det andre har å gjøre med at det ofte er kun de som skårer høyt på tester og andre seleksjonsmetoder man har kriteriedata på. Det er dermed lite variasjon i skåringer på seleksjonsmetodene, fordi de som skårer lavt som regel ikke blir ansatt i en jobb eller kommer inn i en opplærings-situasjon (vel og merke der det er foretatt en treffsikker seleksjon). Begge disse feilkildene fører til lavere kor-

relasjoner enn om de ikke hadde vært tilstede. Dette innebærer at de korrigerede korrelasjonene alltid vil være høyere enn de ukorrigerede.

I praksis sees aldri perfekte validitetskoeffisienter, rett og slett fordi prediksjon av menneskelig atferd er komplisert og bestemt av mange faktorer. En tommelfingerregel er at validitetskoeffisienter over 0,3 er høye (Kaplan & Saccuzo, 2009), men at også lavere verdier kan innebære at metodene er nyttige i praktisk bruk. Bedømmelsen av når en validitetsverdi blir praktisk nyttig er imidlertid ikke klar, og vil avhenge blant annet av kostnadseffektivitet og ressursbruk. I tillegg til å vurdere størrelsen på validiteten er det også vanlig å utføre signifikanstester for å kunne være sikre på at resultatene er systematiske og ikke uttrykk for tilfeldigheter.

I metaanalysen som oppsummerer 100 års forskning, sorteres ulike seleksjonsmetoder ut fra hvilken validitetskoeffisient/gjennomsnittskorrelasjon som beregnes. Tabell 2.1 gjengir verdiene på noen av de metodene PHS bruker, når «training performance» og «overall job performance» er kriteriemålet (oversatt til henholdsvis «opplærings situasjon» og «jobbutførelse»).

Tabell 2.1: Validitetskoeffisienter for fire ulike seleksjonsmetoder (Schmidt et al., 2016)

<i>Seleksjonsmetode</i>	<i>Opplærings situasjon</i>	<i>Jobbutførelse</i>
Evnetester	0,67	0,65
Integritetstester	0,43	0,46
Strukturerte intervjuer	0,41	0,58
Karakterer fra utdanning	-	0,34
Samvittighetsfullhet	0,25	0,22

Spesielt er evnetesters prediktive verdi svært høy. Schmidt et al. (2016) poengterer at tilleggsværdien til andre metoder som har høye samvariasjoner med kriteriet ikke nødvendigvis er så stor når evnetester også brukes. De ulike metodenes forklaringskraft når det kontrolleres for effekten av øvrige metoder, kan undersøkes ved hjelp av regresjonsanalyser. Ut fra slike analyser hevder forskerne at kombinasjonen mellom evnetester og integritetstester, eller evnetester og strukturert intervju, er det som stort sett gir det mest kostnadseffektive seleksjonsopplegget.

Det ser dermed ut til at evnetester er den desidert viktigste seleksjonsmetoden, i alle fall slik det kommer til uttrykk i oppsummerte sammenhenger. Enkelte har riktignok problematisert dette nærmest «hellige gral» funnet i seleksjonsfaget, og rapportert at evnetester ikke predikerer alle typer jobbferd. Deler man opp jobbutførelse i oppgaveløsning, ekstrarolle – og kontra-produktiv jobbferd, peker en annen meta-analyse på at evnetester har god prediksjonskraft på de to førstnevnte kategoriene, men ikke på den siste (Gonzales-Mulé, Mount, & Oh, 2014). Til den siste kategorien, som består av atferd som i mer eller mindre tydelig grad saboterer arbeidsplassens virksomhet, rapporterer disse forskerne at personlighetstrekk er en bedre prediktor enn evner. Andre har også pekt på at evners relevans økes med økende grad av kompleksitet i jobben (Schmidt, Oh, & Le, 2006), noe som kan sies å samsvare med en «common sense» forståelse av sammenhengen.

Ut fra det ovenstående ser vi at PHS benytter seg av seleksjonsmetoder som har en solid forskningsmessig forankring. Slik sett er den nye egnethetsvurderingen atskillig mer kunnskapsbasert, enn det den gamle ordningen var (før 2012). Uansett er det viktig å gjøre en lokal valideringsstudie, da vi ikke vet med sikkerhet om nettopp de seleksjonsmetodene som PHS bruker viser en prediktiv verdi.

## 2.2 Studier fra Forsvaret

Forsvaret har lange tradisjoner med bruk av psykologer og psykologiske metoder i seleksjon, og det har gjennom tiårene etter andre verdenskrig blitt produsert flere rapporter og interne dokumenter som tar for seg prediktiv validitet for ulike opptaksprosedyrer i Forsvaret (Hansen, 2006). Det har spesielt de siste årene blitt produsert en del masteravhandlinger omkring ulike opptak i Forsvaret. To av disse avhandlingene gjennomgås her da prediktorene som ble studert samsvarer med noen av de som brukes ved PHS sin egnethetsvurdering.

Kjenstabakk (2012) studerte den prediktive validiteten av opptaket til befalsskolen i Hæren ( $N=194$ ). Ved gjennomført ettårig befalsskole har elev-



ene oppnådd en snittkarakter i teorifag og fått en lederprestasjonsvurdering av instruktører, noe som slås sammen til en hovedkarakter. Resultatene viste at karakterer fra videregående skole, test på alminnelig evnenivå og en tredagers seleksjonsøvelse i felt var signifikante prediktorer opp mot hovedkarakteren, mens intervju ikke kom så godt ut. Seleksjonssystemet, hvor karakterer, evnenivå, intervju og seleksjonsøvelse inngikk, forklarte 23 % av variasjonen i hovedkarakteren de selekterte endte opp med.

I en masteravhandling som fokuserte på seleksjon til treårig krigsskole ( $N=161$ ) ble det stilt spørsmål om seleksjon har en betydning for å forutsi prestasjoner under utdanning, eller om resultater fra tidligere utdanning og praksis var tilstrekkelig (Vik, 2013). Krigskolen gjennomføres typisk etter at befalsskole og tjenesteperiode som befal er gjennomført, slik at det var en del tidligere prestasjonsvurderinger i form av militære tjenesteuttalelser på deltakerne i studien. Resultatene demonstrerte at tidligere prestasjoner i form av videregående skolekarakterer og militære tjenesteuttalelser hadde en god forklaringskraft opp mot både akademiske –og lederskapmessige prestasjoner under utdanningen. Det ble videre undersøkt om psykologiske tester knyttet til evner, kunnskap og personlighet, samt intervju, tilførte økt forklaringskraft overfor utdanningsresultatene utover tidligere prestasjoner. Svaret på dette ble statistisk positivt, med unntak av personlighetstesting. Seleksjonssystemet, som inneholdt karakterer fra videregående skole og befalsskole, militære tjenesteuttalelser, evne –og kunnskapstester, fysiske tester og intervju, forklarte totalt 42,1 % av variasjonen i totalprestasjonen på krigsskolen. Dette er en svært høy forklart varians, og må nok sees i sammenheng med at tidligere prestasjoner hadde en solid samvariasjon med kriteriene.

Oppsummert har det blitt demonstrert at evnenivå, seleksjonsøvelse i felt og tidligere prestasjoner i form av karakterer fra videregående og militære tjenesteuttalelser er gode prediktorer for prestasjoner i militære utdanninger. Videre kommer intervju variert ut, hvor det ikke hadde særlig prediktiv verdi ved opptak til befalsskole mens det ved opptak til krigsskole hadde en prediktiv verdi. Når det gjelder personlighetstesting, brukes ikke testresultatet direkte i Forsvarets seleksjoner (til forskjell fra PHS, hvor testskårer omgjøres

til en kompetanseskåre). Resultatet fra personlighetstester brukes typisk i intervjuer hvor militærpsykologer deltar, eller brukes til å lage mer omfattende profiler ved uttak til spesielle tjenester.

Ved egnethetsvurderingen ved PHS brukes det evnetester og intervju, men det brukes ikke en lengre seleksjonsøvelse utover den timelange gruppeøvelsen. Når det gjelder tidligere prestasjoner har undersøkelsen tilgang på karakterpoeng fra videregående skole. Funnene fra Forsvaret er nyttig å sammenligne med, men det må tas høyde for at seleksjonskriteriene (kompetansene) og detaljene i metodene er ulike i de to etatene. Det er også en forskjell at grunnutdanningen på PHS ikke utdanner ledere, noe både befalsskolen og krigskolen tar sikte på å gjøre.

### **2.3 Prediktiv validitet ved den svenske politiseleksjonen**

Annell (2015) fulgte søkere fra 2008-kullet ved politiutdanningen i Sverige ( $N=1344$ ) gjennom tre oppfølgingsperioder ( $N=717-729$ ): etter endt toårig politiutdanning, ved slutten av halvårig praksisperiode og ved slutten av første ansettelsesår. Avhandlingen benyttet seg av en rekke kriterier myntet på utdanning/jobb-prestasjoner, tilfredshet i utdanning/jobb, helse, samt ulike psykososiale arbeidsmiljøfaktorer som eksempelvis arbeidsbelastning. Seleksjonen i Sverige har flere likhetstrekk med den norske modellen, med blant annet bruk av psykologiske tester og intervju (Annell, 2012). Siden funnene fra Sverige er direkte relevante for funnene fra PHS, gjengis en del tallverdier i form av korrelasjoner i dette avsnittet. Korrelasjonene er hentet fra Annell, Lindfors, og Sverke (2015).

Evnetester samvarierte signifikant med utdanningsresultater ( $r = .19$ ), men ikke med prestasjoner i praksisperioden ( $r = -.01$ ) eller i første ansettelsesår ( $r = .07$ ). Forfatteren diskuterer dette funnet, som er motsatt av hva meta-analysene tilsier, ut fra en hypotese om begrenset spredning i testresultater samt at det er vanskelig å ha gode målbare kriterier på politiyrket. Videre hadde evnenivå en svak negativ sammenheng med tilfredshet ved alle tre oppfølgingsperioder ( $r = -.16 - -.07$ ), noe som betyr at de som fikk høye resultat på evnetestene mistrivdes mer. Ut fra dette funnet stiller han seg

spørsmål om hvordan kognitivt begavet personell blir tatt imot i politiet, men påpeker samtidig at det er en del studier som peker på at individer med høye evner ofte opplever lavere jobbtilfredshet (Ganzach, 1998).

En svensk språkprøve hadde en signifikant sammenheng med utdanningsresultater ( $r = .26$ ), men ikke med praksisperiode og yrket ( $r = .04$  og  $.07$ ). Annell tar opp at dette funnet har samsvar med tidligere forskningsfunn fra politiorganisasjoner, hvor verbale tester ser ut til å predikere prestasjoner bedre enn andre kognitive tester (Aamodt, 2004b). Han poengterer at funnet bekrefter betydningen av god kommunikasjon i politiyrket, men ser ikke ut til å problematisere at den signifikante sammenhengen mellom språkprøven og prestasjoner kun sees opp mot utdanningsresultater.

Når resultater fra en personlighetstest ble koblet med kriteriene, ble det funnet at personlighetstrekket *medmenneskelighet* samvarierte signifikant med tilfredshet i alle tre oppfølgingsperioder ( $r = .11 - .21$ ), samt at trekket hadde positiv sammenheng med god helse. Funnet drøftes med henblikk på at *medmenneskelighet* kan være betydningsfullt for hvordan nye politimenn – og kvinner tilpasser seg de ulike jobbkravene i yrket. Videre hadde *samvittighetsfullhet* signifikante sammenhenger med utdanningsresultater ( $r = .11$ ), samt trivsel i utdanning og praksisperiode ( $r = .08$  og  $.17$ ). *Emosjonell stabilitet* hadde signifikante sammenhenger med prestasjon i praksisperiode ( $r = .12$ ), og trivsel ved alle tre oppfølgingsperioder ( $r = .09 - .21$ ). Personlighetstesten var bygget rundt noen av trekkene som inngår i femfaktormodellen i personlighetspsykologien. Femfaktormodellen er en svært utbredt modell for å beskrive individuelle forskjeller innen personlighet, og kategoriserer menneskets personlighet ut fra fem overordnede faktorer: Ekstrovertsjon, Medmenneskelighet, Samvittighetsfullhet, Nevrotisisme (omvendt Emosjonell stabilitet) og Åpenhet for opplevelser (John, Naumann, & Soto, 2008). Den svenske testen inneholdt faktorene medmenneskelighet, samvittighetsfullhet og nevrotisisme.

Søkerne intervjues på slutten av opptaket ved to anledninger, av henholdsvis politipersonell og psykolog. Politiintervjuet viste signifikante sammenhenger med prestasjon og trivsel i praksisperiode ( $r = .08$  og  $.10$ ), mens psykologintervjuet hadde signifikant sammenhenger med utdanningsresultater

og prestasjon i praksisperiode ( $r = .11$  og  $.07$ ) og med helse under utdanning ( $r = .08$ ). De to intervjupanelene slår etter hvert sammen sitt inntrykk av søkerne, hvor begge har forsøkt å kartlegge søkerens personlige egenskaper. Formatet på intervjuet er semistrukturert. Et interessant funn var for øvrig at medisinsk undersøkelse ved opptaket ikke viste sammenheng med helse ved alle tre oppfølgingsperioder.

Fysiske kondisjonstester hadde signifikante sammenhenger med utdanningsresultater og yrkesprestasjon ( $r = .16$  og  $.18$ ), samt med god helse ved alle tre oppfølgingsperioder. Forfatteren argumenterer for at fysiske tester trolig er den beste prediktoren i det svenske opptaket.

Avhandlingen til Annell (2015) er meget solid med tanke på kriterier å måle seleksjonsresultater opp mot. Korrelasjonene som oppgis blir relevante å sammenligne med for funn fra PHS. Basert på hovedtrender i de svenske funnene anbefaler Annell at intervjupraksisen heller bør fokusere på å velge bort uønskede kandidater, fremfor å søke å belyse positive prediktive forhold (negativ versus positiv seleksjon). Blant annet begrunner han det med at det finnes mye informasjon i tester som ikke utnyttes fullt ut, slik at intervjuet blir overflødig i positiv seleksjonsøyemed. En optimal praksis ut fra hans vurdering vil være å kombinere informasjon fra verbale tester, personlighetstest og kondisjonstester. Om vi går tilbake til anbefalingen fra Kaplan og Saccuzo (2009) om at validitetskoeffisienter (det vil si korrelasjonsverdier mellom mål gjort på ulik tidspunkt) over 0,3 er høye, blir imidlertid Annell sine funn forholdsviss beskjedne. Det er vanskelig å få øye på den praktiske nytten når samvariasjonene blir under 0,3, selv om de er statistisk signifikante.

Et viktig avsluttende poeng med avhandlingen til Annell var at hans overordnede mål var å belyse faktorer som antas å være viktige for holdbar rekruttering, altså at politipersonellet blir i jobben over tid. Til dette formålet trengtes kartlegging av ytre organisatoriske faktorer i tillegg til de individuelle. Når betydningen av seleksjon og psykososiale forhold på jobben ble veid opp mot hverandre, viste det seg at sistnevnte forhold forklarte mer variasjon i en rekke forhold ved første ansettelsesår enn det seleksjonsdataene evnet å gjøre. Opp mot medarbeideres variasjon i jobbtilfredshet, ekstrarolle jobbatferd, vilje til å fortsette i yrket og helseforhold ble organisatoriske forhold på jobben

betydelig viktigere som forklaringsvariabler enn seleksjonen medarbeiderne hadde gått gjennom. Dette siste poenget fra avhandlingen er nyttig å vurdere når et seleksjonssystem skal gjennomgås.

## 2.4 Evaluering fra Agenda Kaupang

Av de aktørene som gjorde tidlige evalueringer av opptaksordningen ved PHS etter at den trådte i kraft i 2012, er Agende Kaupang sin rapport den mest grundige ut fra forfatterens skjønn (AgendaKaupang, 2012).

Mandatet ble beskrevet slik:

Agenda Kaupangs mandat har vært å måle opptaksnemdenes oppfatning av de nye verktøyene, og om de har bidratt til en forbedring i nemndenes grunnlag for skjønnsutøvelse av søkerens personlige egnethet. Det har vært utenfor vårt mandat å gi konkrete anbefalinger om hvordan de ulike metodene eventuelt burde forandres. (s.36)

Mandatet ble løst ved å benytte spørreundersøkelser, gjennomføre fokus-gruppeintervju og være tilstede ved et opplæringsseminar. Det ble dermed foretatt en bred og grundig datainnsamling som hovedkonklusjonen og enkelte konkrete forbedringspunkter baserte seg på.

En samlet vurdering i rapporten beskrives slik:

Med bakgrunn i de funn som er presentert i denne rapporten, kan vi konkludere med at innføring av ny opptaksform for PHS' opptaksnemnder har styrket grunnlaget for vurdering av søkerens personlige egnethet. De nye arbeidsformene gir mer tid med kandidatene, søkerne vurderes ved hjelp av flere metoder og etter kjente vurderingskriterier. Standardiserte metoder og arbeidsprosesser sikrer at utvelgelsen blir lik på tvers av de ulike opptaksnemndene. (s. 36)

På mer konkret anbefalingsnivå anbefaler rapporten en del endringer og forbedringspunkter, eksempelvis det å fjerne personlighetstesten og heller bygge ut intervjuet og gruppeøvelsen. Bruken av testen Shapes har skapt usikkerhet blant opptaksnemdene, og rapporten anbefaler i alle fall en mer omfattende testopplæring om den ikke fjernes. Det var også knyttet noen utfordringer til bedømmermøtet, hvor det ved tvilstilfeller hersket usikkerhet om beslutning om egnethet skulle avgjøres matematisk med gjennomsnitt av kompetanseskårene eller om det skulle være mulighet for justeringer når alle resultater sees under ett.

Generelt sett ser det ut som den nye opptaksprosessen har fått gode evalueringer. Det ble identifisert noen forbedringspunkter av mer konkret karakter, noe som ikke er overraskende for en gryende seleksjonssystem. Rapporten er nyttig å se til for å få et bilde av hvordan opptaksnemdene oppfatter den endrede arbeidsformen, og hvilke praktiske utfordringer det fort kan bli ved et nytt system. Det er imidlertid viktig å poengtere at et slikt her-og-nå statusbilde ikke sier noe om hvor godt seleksjonen virker. For å belyse seleksjonens prediktive verdi trengs det valideringsstudier som følger de samme personene over lengre tid, noe de to studiene fra Forsvaret og studien fra det svenske politioptaket er eksempler på.

## 3 Metode

### 3.1 Utvalg

Utvalget studien baserer seg på består av søkere som var gjennom opptaket til PHS i 2013 og 2014, og av disse de som ble studenter i årene 2013-2016 og 2014-2017. På grunn av ulikhet i enkelte prediktorer og i en del kriterier for de to kullene, ble det operert med to datasett i denne undersøkelsen.

I databasen for kull 2013 var det registrert 720 studenter, hvorav 444 (62 %) menn og 276 (38 %) kvinner. Gjennomsnittsalderen for disse var 22,9 år i opptaksåret 2013 ( $SD=2,9$ , aldersvariasjon 20-38 år). Det var manglende karakterregistreringer for 2 til 56 blant disse. Antallet med registrerte psykologiske testresultater og opptaksnemdvurderinger var henholdsvis 1128 og 1142 personer, og av disse var 41 registrert med uakseptabel integritet.

For kull 2014 var det registrert 671 studenter, hvorav 441 (62 %) menn og 257 (38 %) kvinner. Gjennomsnittsalderen i opptaksåret 2014 var 22,6 år ( $SD=2,7$ , aldersvariasjon 20-41). Det var manglende registreringer av karakterer for 53 til 158 blant disse. Antallet med registrerte resultater fra psykologiske tester og opptaksnemdvurderinger var henholdsvis 1315 og 1270 personer, og av disse var 47 registrert med uakseptabel integritet. For kull 2014 var det også registrert 97 studenter som ikke hadde fullført studiet innen normert tid.

### 3.2 Prosedyre

Undersøkelsens formål var å statistisk koble tilgjengelige prediktorer med tilgjengelige kriterier for å undersøke den prediktive validiteten til egnethetsvurderingen. Datautlevering ble godkjent av rektor PHS ved prosjektoppstart i november 2014, og en styringsgruppe med internt personell ved PHS ble opprettet. Spesialist i organisasjonspsykologi (cand.psychol.) Tom H. Skoglund ble midlertidig deltidsansatt som forsker ved PHS for å gjennomføre studien.

Seksjon for opptak og rekruttering ved studiekontoret Oslo koblet prediktorer med kriterier, og utleverte en anonymisert database til undersøkelsen. Studien ble vurdert å være unntatt meldeplikt av Personvernombudet for forskning (NSD) i mars 2015, da studien er del av internt system for kvalitets-sikring i henhold til universitets- og høyskoleloven § 1-6 (jf. personopplysningsforskriften § 7-20).

### **3.3 Prediktorer**

Prediktorene som brukes i denne studien er tre-delt: karakterpoeng fra videregående skole, psykologiske testresultater fra cut-e AS og opptaksnemndas vurderinger. Forfatteren kjenner ikke til bakgrunnen for valget og bruken av de psykologiske testene som ble benyttet. Forfatteren kjenner heller ikke til beslutningsgrunnlaget bak valg av metodikk for intervju og gruppeøvelse. I det følgende beskrives derfor bare innholdet i prediktorområdene, ut fra den informasjonen som har blitt oppgitt fra cut-e AS og PHS.

#### **3.3.1. Karakterpoeng**

Karakterpoeng er gjennomsnittet av karakterer fra videregående skole ganget med 10, til forskjell fra konkurransepoeng og skolepoeng hvor ulike tilleggspoeng er tatt med i beregningen av karaktersnittet.<sup>4</sup>

#### **3.3.2. Psykologiske tester fra cut-e AS**

Testbatteriet som ble benyttet for kull 2013 og 2014 besto av tester ment å mål verbale evner og evner på logisk resonnering, samt en personlighetstest. Kull 2014 ble i tillegg testet med en integritetstest. Alle testene er digitale, og ble tatt på PC-sal ved PHS (fra 2015 av ble testene tatt hjemme hos den enkelte søker). Testbeskrivelsene nedenfor er basert på interne dokumenter fra cut-e AS (cut-e, 2013, 2015). cut-e AS har rapportert enkelte reliabilitets- og validitetsmål på testene i disse dokumentene. Gjennomgangen her begrenses imidlertid til prediktiv validitetsverdier, siden det er de tallverdiene som blir direkte relevante å sammenligne med opp mot funn fra PHS.

---

4 Se <http://www.samordnaopptak.no/info/opptak/poengberegning/index.html>



### *Tester på verbale evner*

**Scales verbal admin** er ment å måle evne til å finne og bruke konkret verbal informasjon i artikler, lister og instruksjoner. Scales It-no er ment å måle norskferdigheter gjennom tre deler: språkflyt, vokabular og staving. Resultater fra Scales It-no ga en skåre kompetansen samhandling. Testenes varighet er henholdsvis 10 og 12 minutter.

### *Tester på logisk resonnering*

**Scales lst og Scales sx (a)** er ment å måle deduktiv logisk resonnering. Oppgaven på lst er å identifisere hvilket symbol som mangler i en rute. På sx (a) er oppgaven å identifisere en riktig tallkombinasjon. Testenes varighet er henholdsvis 6 og 5 minutter. Scales ix er ment å måle induktiv logisk resonnering. Oppgaven består av å oppdage en hovedregel og deretter finne objektet som ikke passer til regelen. Testens varighet er 5 minutter.

### *Tester for å utlede skåre på kompetansen «Analytisk evne»*

For kull 2013 ble testene Scales verbal admin, Scales ix og Scales sx (a) brukt for å utlede en totalskåre på *analytisk evne*. For kull 2014 ble Scales verbal admin og Scales lst benyttet til samme formål.

### *Prediktiv validitet for Scales-testene*

I test-dokumentasjonen for **Scales sx (a)**, eller i annen dokumentasjon fra cut-e AS, er det ikke rapportert prediktiv validitet for denne testen (cut-e, 2014). For de andre Scales-testene kommer det frem følgende funn fra interne rapporter (cut-e, 2016, 2017a):

- **Scales lst** korrelerte fra 0,22-0,49 med ulike utfallsmål for en gruppe selgere (N=331), lærere (N=216) og helsearbeidere (N=150) i Sverige.
- **Scales verbal admin og Scales It-no** korrelerte henholdsvis 0,25 og 0,35 med oppgaveorienterte arbeidsoppgaver («task») for lærergruppen.
- **Scales ix** korrelerte fra 0,27-0,43 med ulike utfallsmål for «teknisk personell» (N=52), samt at testen korrelerte 0,64 med lederevaluering på prestasjoner for et lite utvalg (N = 23) i et norsk varehus.

Tabell 3.1: Shapes (basic) dimensjon, forklaring på disse og relasjon til PHS-kompetanse

Shapes (basic) dimensjon	Forklaring	PHS-kompetanse
Varsomhet	Ser større helheter; gjenkjenner mulig risiko; tar i betraktning konsekvensene av individuelle handlinger.	MODENHET
Opplevd mestring	Er overbevist om sin evne til å oppnå egne mål; stoler på egne ferdigheter; er overbevist om egne evner.	
Ivrighet	Går til jobben med lyst og energi; liker følelsen av at man er nødvendig; er entusiastisk og legger mye energi i arbeidet	
Ansvarsbevissthet	Arbeider samvittighetsfullt og pålitelig; mener det er viktig med orden, nøyaktighet og integritet.	HANDLEKRAFT
Selvstendighet	Handler godt på eget initiativ; tar ansvar; uttrykker fritt egne synspunkter; gjør sine meninger kjent for andre.	
Utholdenhet	Tilnærmer seg arbeidet med en grunnleggende positiv holdning; gir seg ikke etter nederlag; fortsetter med oppgaver selv om de er kjedelige.	
Harmoni	Unngår krangler og konflikter på jobben; synes det er viktig at det er harmoni, og at mennesker kommer godt overens med hverandre.	SAMHANDLING
Sosiale ferdigheter	Er pratsom og sosial; er komfortabel sammen med andre; kjenner mange mennesker; er livlig og aktiv i grupper.	
Samarbeid	Liker å jobbe sammen med andre; tror på teamarbeid; er hensynsfull mot andre; støtter og hjelper innad i teamet.	

### *Personlighetstest*

**Shapes (basic)** er en forenklet versjon av en mer omfattende personlighetstest (kalt Shapes management). Testen er ment å gi et bilde av en persons kompetanseprofil, det vil si hvordan personens yrkesrelaterte atferd kan beskrives. Testen gjennomføres ved at tre testledd presenteres om gangen for testtakeren, hvor oppgaven er å bedømme hvorvidt man er enig i at leddene beskriver en selv. Dette gjøres ved å fordele seks poeng til de tre leddene. Testens varighet varierer, men oppgis å være 9 minutter i gjennomsnitt. Ni ulike personlighetsdimensjoner/kompetanser fra denne testen ble brukt i seleksjonen for kull 2013 og 2014. De ni dimensjonene ble igjen fordelt til hvilken PHS-kompetanse de tilhørte. PHS-kompetansene modenhet, handlekraft og samhandling ble hver for seg belyst av tre Shapes (basic) dimensjoner. Tabell 3.1 (over) viser denne sammenhengen.

### *Prediktiv validitet for Shapes (basic)*

I testdokumentasjonen for Shapes (basic) er det rapportert signifikante korrelasjoner mellom dimensjonene Ivrighet, Ansvarsbevissthet, Utholdenhet, Harmoni og Samarbeid på den ene siden, og intervjuuskårer for lærlinger innen salg og teknisk bransje på den andre siden ( $N=483$ ) (cut-e, 2017b). Alle korrelasjonene er høyere enn 0,30, det anbefalte nivået fra (Kaplan & Saccuzo, 2009). Det kommer også frem i samme rapport at testen hadde et godt samsvar med overordnede vurdering av medarbeidere på en del ulike prestasjonskriterier ( $N=40$ , kommer ikke frem hvilken bransje dette var). Tidsrommet mellom Shapes (basic) testing og evalueringer gjort gjennom intervju og leder er imidlertid usikker. Om det er kort tid mellom de to målepunktene, er funnene uttrykk samtidig validitet fremfor en prediktiv validitet. Funnene viser uavhengig av dette at personlighetstesten har sammenheng med relevante kriterier.

### *Integritetstest for kull 2014*

**Squares** er en test som søker å måle sannsynlighet for kontraproduktiv jobbatferd. I testmanualen til Squares nevnes noen eksempler på denne type atferd, blant annet sabotasje, tyveri, bedrageri, rusmisbruk og trakassering. Testen søker å måle to faktorer, *impulskontroll* og *etisk bevissthet*, som i sin tur kobles til en skåre ment å indikere hvor *pålitelig* testtakeren er. Oppgaven i testen er å vurdere hvor man mener at man selv tilhører på en relativ skala sammenlignet med andre, ut fra en rekke forhold som går på atferd knyttet til de to ovennevnte faktorene. I tillegg til selve skalaposisjonen testtakeren bestemmer seg for, registreres også en antatt bedømmelsesprosess gjennom bruk av datamusen. Resultatet fra denne testen knyttes til PHS-kompetansen *integritet*, og testen benevnes som en *sårbarhetstest* i PHS sitt opptak. Testens varighet varierer, men oppgis å være gjennomsnittlig 7,5 minutter.

### *Prediktiv validitet for Squares*

På et lite utvalg ( $N=23$ ) i en norsk varehusbedrift viste Squares flere signifikante korrelasjoner med en 3-delt lederevaluering (lav, medium og høy jobbutførelse) (cut-e, 2017a). En total «safety score» fra testen korrelerte 0,62

med lederevalueringen. Det er også her usikkerhet med tanke på tidsrommet mellom testing og lederevaluering. Slik sett kan korrelasjonen være uttrykk for samtidig fremfor prediktiv validitet.

### 3.3.3. Opptaksnemndenes vurderinger

I dette prediktorområdet inngår bedømmelser av prestasjoner på intervjuet og gruppeøvelsen, hvor kompetansene som måles skåres på en 5-delt skala med følgende ordlyd: svært svak, svak, akseptabel, sterk, svært sterk.

#### *Intervjuet*

Alle kandidatene ble møtt med samme intervjustruktur uavhengig av hvem i opptaksnemnda som intervjuet. Intervjuet var kompetansebasert, ispedd spørsmål om kandidaten tidligere hadde søkt PHS. Hypoteser fra Shapes (basic) ble presentert, samt at hypoteser fra integritetstesten Squares ble tilkjennegitt for kull 2014. Dersom kandidaten hadde en lav konsistensskåre fra Shapes (basic) ble ikke hypoteser fra denne testen presentert (validitetsindeks i testen ment å indikere hvorvidt kandidaten har svart unyansert eller tilfeldig). Intervjuets varighet var om lag 40 minutter.

Det ble gitt spørsmål knyttet tilandel om intervjuer fikk merknad om dette fra skikkethetsvurderingen. Der det var vandelsmerknad ble kandidatens svar vurdert opp mot kompetansen *integritet*. For kull 2014 ble denne kompetansen søkt videre avdekket gjennom spørsmål knyttet til impuls kontroll og etisk bevissthet dersom kandidaten hadde lave skårer på sårbarhetstesten Squares.

Intervjuet søkte deretter i tur og orden å avdekke informasjon knyttet til kompetansene *åpenhet og inkludering, modenhet, handlekraft og samhandling*. De fleste spørsmålene i intervjuet var knyttet til åpenhet og inkludering. Forfatteren regner med at bakgrunnen for ulik vektning av spørsmål skyldtes at *modenhet, handlekraft og samhandling* også ble skåret ved hjelp av Shapes (basic) resultater. Intervjuer kunne justere testresultatet på de tre testutledede kompetanseskårene basert på intervjuet, men hadde kun søkerens intervju svar å basere seg på for skåring av åpenhet og inkludering. Samtlige kompetanser ble skåret på den 5-delte skalaen, med unntak av *integritet* som ble kategorisert som «Ok» eller «Uakseptabel».

### *Gruppeøvelsen*

Denne seleksjonsmetoden kan beskrives som en jobbsimulering, hvor kandidatenes atferd i et møtescenario søkes avdekket. Kandidatene ble samlet åtte personer om gangen og fikk presentert sakspapirer som skulle diskuteres. Det var ingen utpekt leder av gruppen, og kandidatens atferd innen kompetansene *samhandling, handlekraft og modenhet* ble vurdert. Hvert medlem av opptaksnemda vurderte to kandidater om gangen ut fra et skjema som beskrev atferdsindikatorer på de tre kompetansene. Hver kompetanse ble vurdert ut fra den 5-delte skalaen. Gruppeøvelsens varighet var om lag 1 time.

### *Beslutningsmøtet*

Opptaksnemdene gjennomførte et endelig beslutningsmøte hvor alle skårer fra tester, intervjuet og gruppeøvelsen ble gjennomgått. Nemda ble enige om en totalskåre på hver kompetanse, og hadde mulighet til å enten matematisk utregne den eller justere totalskåren ut fra andre aspekter ved totalvurderingen. Kompetansene *analytisk evne* og *integritet* ble ikke direkte berørt av opptaksnemdas seleksjonsmetoder, da vurdering her ble gjort direkte ut fra psykologiske tester og skikkethetsvurdering. Opptaksnemda hadde likevel mulighet til også å justere vurderinger på disse to kompetansene dersom det ble relevant. Som nevnt ble det stilt intervju spørsmål dersom skåre på sårbarhetstesten Squares var lav, eller i de tilfellene det forekom vandelsmerknad.

## **3.4 Kriterier**

Fagemner og praktiske oppgaver studentene fikk karakterer i gjennom politiutdanningen utgjør kriteriene. De enkelte fagene var strukturert inn i de fem hovedområdene som gjenspeiler hovedoppgavene til en politigeneralist. Hovedområdet *orden og beredskap* ble kalt *operative oppgaver* for kull 2013. Bacheloroppgaven for kull 2013 var knyttet delvis til hovedområdet *forebyggende* (med 4 av de totalt 6 studiepoeng som utgjorde oppgaven), mens bacheloroppgaven for kull 2014 var tilknyttet hovedområdet *metode* i sin helhet (økt til 10 studiepoeng). I praksis var temaene på bacheloroppgavene for begge kull selvvalgte, og kunne derfor ta for seg innhold fra alle fem hovedområder.

### 3.4.1. Kriterier for kull 2013

Tabell 3.2: Hovedområde, fagemne/fagkode og vurderingsform på kriteriedata for kull 2013

Hovedområde og studiepoeng (sp)	Fagemne, fagkode og studiepoeng (sp)	Vurderingsform
Operative oppgaver 21,5 sp	Ordenstjeneste (ORD130) 4 sp	Muntlig eksamen
	Psykologi (PSY130) 4 sp	Skriftlig tre timers eksamen
	Vegtrafikkrett (VEG130) 3,5 sp	Skriftlig tre timers eksamen
	Arrestasjonsteknikk (AT330) 2 sp	Praktisk/muntlig eksamen
	Fysisk trening (FYS330) 2 sp	Praktisk eksamen
	Kommunikasjon og konflikthåndtering (KKH330) 6 sp	Muntlig eksamen
Metode 12,5 sp	Juridisk metode:strafferett, straffeprosess og ordensjuss (STR150) 12,5 sp	Skriftlig seks timers eksamen
Politi og samfunn 21 sp	Politolære (POL110) 6 sp	Hjemmeeksamen, 4 dager
	Sosiologi (SOS110) 4 sp	Muntlig eksamen
	Yrkesettikk (YRK310) 3 sp	Hjemmeeksamen, 2 dager
	Kriminologi (KRIM310) 3 sp	Skriftlig tre timers eksamen
	Forvaltningsrett og privatrettslige emner (FOP310) 5 sp	Skriftlig fem timers eksamen
Forebyggende oppgaver 9 sp	Forebyggende oppgaver: Kriminalitetsforebyggende politiarbeid, kriminologi, strafferett og straffeprosess, forvaltningsrett (FORB340) 9 sp	Tverrfaglig hjemmeeksamen, 2,5 dag
Etterforskning 15 sp	Kriminalteknikk (KRI150) 3 sp	Skriftlig 3 timers eksamen
	Etterforskning: strafferett og straffeprosess, kriminalteknikk, rapport- og etterforsklingslære, psykologi (ETF350) 12 sp	Muntlig eksamen
Bacheloroppgaven 6 sp	Bacheloroppgaven (OPPG300) 6 sp	Oppgave basert på ca. 240 sider selvvalgt litteratur

Det ble ikke gitt karakterer i alle fag og praktiske oppgaver, ei heller i praksisåret B2. For de fleste statistiske analysene som gjøres i denne studien trengs graderte mål på studentenes prestasjoner, noe som betyr at kriteriene prediktorene måles opp mot slankes en del sett opp mot den helhetlige bachelorutdanningen. En komplett bachelorgrad består av 180 studiepoeng. For kull 2013 gis det karak-

ter i 47 % av de produserte studiepoengene (85 av 180), for kull 2014 59 % (106 av 180). Det er en del forskjeller på hva studentene fikk karakterer i på de to kullene, noe tabellene gir en oversikt over.

### 3.4.2. Kriterier for kull 2014

Tabell 3.3: Hovedområde, fagemne/fagkode og vurderingsform på kriteriedata for kull 2014

Hovedområde og studiepoeng (sp)	Fagemne, fagkode og studiepoeng (sp)	Vurderingsform
Orden og beredskap 31 sp	Ordenstjeneste og ordensjus (DRFK10) 5 sp	Muntlig felleseksamen i emnene Ordenstjeneste og ordensjus, Kommunikasjon og konflikthåndtering, Fysisk trening og deler av Digitalt politiarbeid
	Fysisk trening (DRFK10) 2 p	Muntlig felleseksamen i emnene Ordenstjeneste og ordensjus, Kommunikasjon og konflikthåndtering, Fysisk trening og deler av Digitalt politiarbeid
	Kommunikasjon og konflikthåndtering (DRFK10) 2 sp	Muntlig felleseksamen i emnene Ordenstjeneste og ordensjus, Kommunikasjon og konflikthåndtering, Fysisk trening og deler av Digitalt politiarbeid
	Psykologi (PSYK10) 4 sp	Tre timers skriftlig eksamen
	Vegtrafikkrett (VEGT10) 3 sp	Tre timers skriftlig eksamen
	Kommunikasjon og konflikthåndtering (KKH30) 4 sp	Trekkeksamen mellom emnene Kommunikasjon og konflikthåndtering og Ordenstjeneste og ordensjus. (emnekode KORD30): Muntlig eksamen
Orden og Beredskap 31 sp	Ordenstjeneste og ordensjus (ORTJ30) 5 sp	Trekkeksamen mellom emnene Kommunikasjon og konflikthåndtering og Ordenstjeneste og ordensjus. (emnekode KORD30): Muntlig eksamen
	IP-trening (OPR30) 4 sp	Fire timers skriftlig eksamen.
	Arrestasjonsteknikk (AT30) 2 sp	Muntlig eksamen

Metode 10 sp	Bacheloroppgaven (BOPPG30) 10	6000 ord (pluss/minus 10 %). Samskriving to studenter 8000 ord (pluss/minus 10 %)
Politi og samfunn 18 sp	Sosiologi (SOPF10) 4 sp	Trekkeksamen for emnene Sosiologi, Politilære og Forvaltningsrett. Hjemmeeksamen.
	Politilære (SOPF10) 5 sp	Trekkeksamen for emnene Sosiologi, Politilære og Forvaltningsrett. Hjemmeeksamen.
	Forvaltningsrett (SOPF10) 2 sp	Trekkeksamen for emnene Sosiologi, Politilære og Forvaltningsrett. Hjemmeeksamen.
	Kriminologi (KRIM30) 4 sp	Trekkeksamen mellom emnene Kriminologi og Yrkesetikk. Skriftlig tre timers eksamen.
	Yrkesetikk (YRKE30) 3 sp	Trekkeksamen mellom emnene Kriminologi og Yrkesetikk. Hjemmeeksamen.
Forebyggende oppgaver 13 sp	Kriminalitetsforebyggende politiarbeid (KRIF10) 3 sp	Tre timers skriftlig eksamen
	Forebyggende oppgaver (KRIF30) 10 sp	Fem timers skriftlig eksamen
Etterforskning 34 sp	Strafferett og straffeprosess (STRA10) 10 sp	Seks timers skriftlig eksamen
	Rapport- og etterforsklingslære (REL10) 6 sp	Fem timers skriftlig eksamen som inkluderer deler av emnet Digitalt politiarbeid.
	Kriminalteknikk (KRIT10) 4 sp	Muntlig eksamen
Etterforskning 34 sp	Etterforskning: strafferett og straffeprosess, kriminalteknikk, rapport- og etterforsklingslære, psykologi (ETTERF30) 14 sp	Muntlig eksamen

### 3.5 Statistiske analyser

Alle statistiske analyser ble foretatt med SPSS, versjon 23. Sammenhengen mellom prediktorene og kriteriene ble belyst ved hjelp av korrelasjonsanalyser og regresjonsanalyser. I tillegg ble det undersøkt om det er sammenheng



mellom egnethetsvurderingen og integritetsbedømmelsen. For kull 2014 ble det i tillegg undersøkt om det er sammenheng mellom egnethetsvurderingen og studiefullføring på normert tid.

En korrelasjonsanalyse viser styrken på sammenhengene mellom de ulike prediktorene og kriteriene, mens en regresjonsanalyse er egnet til å vise i hvor stor grad flere prediktorer til sammen kan prediktere kriteriene. En hierarkisk multippel regresjonsanalyse ble valgt siden denne metoden også belyser de statistisk unike prediktive egenskaper ved de ulike gruppene av prediktorer, i tillegg til den samlede forklaringskraften, opp mot variasjonen i kriteriene.

Sammenhengen mellom integritetsbedømmelsen og studiefullføring på normert tid ble belyst ved hjelp av korrelasjoner. Disse analysene vil være verdifulle for å undersøke om prediktorene kan belyse kontraproduktive forhold både sikkerhetsmessig (integritet) og studiemessig (ikke forventet progresjon).

## 4 Resultater

I det følgende presenteres analyseresultatene for de to kullene. Vedlegg 1 og 2 viser korrelasjonstabellene for variablene som ble brukt for henholdsvis kull 2013 og 2014. Korrelasjonstabellene gjengir kriteriene i form av gjennomsnittskarakterer på de fem ulike hovedområdene i utdanningen, samt det totale karaktersnittet studentene oppnådde. I Vedlegg 3 (kull 2013) og 4 (kull 2014) vises beskrivende statistikk for prediktorene og kriteriene i det koblede datasettet. Disse vedleggene viser en god spredning på prediktorene (inkludert karakterpoeng fra videregående skole) for de som ble tatt inn som studenter, samt en god spredning på karakterene studentene fikk i løpet av studiet.

Som kriterium for korrelasjons- og regresjonsanalysene ble gjennomsnittskarakteren studentene fikk på PHS benyttet, heretter kalt karaktersnittet. Karaktersnittet ble beregnet ved at resultatet på samtlige fagemner ble delt på antall fagemner. For hovedområdet *Politi og samfunn* for kull 2014 var det slik at om lag halvparten av kullet hadde eksamen i *yrkesetikk* (YRKE30), og den andre halvparten hadde eksamen i *kriminologi* (KRIM30); dette ble tatt hensyn til, slik at den karakteren den enkelte student hadde produsert ble tatt med i beregningen av karaktersnittet. Kun de med registrerte karakterer i alle fagemner fikk en gyldig verdi på karaktersnitt.

Det ble også undersøkt mulige sammenhenger mellom prediktorer og integritetsbedømmelsen, og for kull 2014 ble det i tillegg undersøkt hvorvidt det var sammenheng mellom egnethetsvurderingen og fullføring av studiet på normert tid.

### 4.1 Kull 2013

#### 4.1.1. Korrelasjonsanalyse

Karaktersnittet viste en sterkest samvariasjon med karakterpoeng fra videregående skole, mens det var en svakere samvariasjon med evnetestene, inter-

vjuskåren på åpenhet og inkludering samt de tre kompetansene som måles i gruppeøvelsen. Språktesten Scales It-no og kompetanseskårene fra personlighetstesten Shapes viste ingen samvariasjon med kriteriene, ei heller intervju-skårene på *modenhet*, *handlekraft* og *samhandling*. Tabell 4.1. viser de signifikante korrelasjonene i rangert orden. For øvrige korrelasjoner henvises det til vedlegg 1.

Tabell 4.1: Kull 2013: Signifikante korrelasjoner i rangert orden mellom prediktorer og karaktersnitt (N=552)

Prediktor	Korrelasjonskoeffisient r
Karakterpoeng VGS	.42**
Intervju Åpenhet og inkludering	.18**
Scales sx (a)	.18**
Gruppeøvelse Modenhet	.15**
Scales verbal admin	.12**
Gruppeøvelse Handlekraft	.11*
Gruppeøvelse Samhandling	.10*
Scales ix	.10*

Note. \*p < .05; \*\*p < .01 (to-halet)

#### 4.1.2. Multipel hierarkisk regresjonsanalyse

Det neste steget i å analysere sammenhengen mellom prediktorer og kriterier besto av å se om de øvrige variablene hadde en prediksjonsverdi, når effekten av karakterpoeng fra videregående skole ble kontrollert for. Til dette formålet ble en gjennomsnittsskåre for gruppeøvelsen beregnet, siden interkorrelasjonene mellom de tre kompetansemålingene var høye for denne seleksjonsmetoden ( $r = .73-.78$ ). Gjennomsnittsskåren på gruppeøvelsen og karaktersnittet samvarierte signifikant ( $r = .13$ ). Scales-testene viste ikke en slik tydelig samvariasjon og ble derfor holdt atskilt i regresjonsanalysen.

Tabell 4.2. viser at karakterpoeng forklarer 17 % av variasjonen i karaktersnittet. Når de øvrige variablene som korrelerte signifikant med karaktersnittet tas med i steg 2, forklarer alle variablene til sammen 21 % av variasjonen. Variablene utover karakterpoeng fra videregående skole forklarte ytterligere 3 % av variasjonen. Når alle variablene sees samlet i steg 2, er

karakterpoeng fra videregående skole, intervju skåre åpenhet og inkludering, samt evnetesten Scales sx (a) signifikante, og økningen i forklart varians er signifikant fra steg 1 til steg 2.

Tabell 4.2: Kull 2013: Hierarkisk regresjonsanalyse med totalt karaktersnitt på PHS som avhengig variabel

	Steg 1	Steg 2
Karakterpoeng VGS	.42**	.38**
Intervju åpenhet og inkludering		.10*
Scales sx (a)		.08*
Gruppeøvelse gjennomsnitt		.08*
Scales verbal admin		.06
Scales ix		.01
R <sup>2</sup>	.17**	.21**
R <sup>2</sup> forandring	-	.03**
F-verdi	114.63**	23.61**

Note. \*p < .05; \*\*p < .01

*er standardisert regresjonskoeffisient, R<sup>2</sup> er forklart varians i PHS-karaktersnittet, F-verdien indikerer stegenes signifikans for den avhengige variabelen. R<sup>2</sup> forandring Steg 2 er 0,034.*

#### 4.1.3. Egnethetsvurdering og integritet

For kull 2013 var 41 søkere registrert med uakseptabel integritet. For å undersøke den samtidige kriterievaliditeten til egnethetsvurderingen, ble korrelasjonsanalyser (punkt-biserial) gjennomført. Den dikotome variabelen integritet ble kodet slik at 0 indikerte «ok integritet» og 1 indikerte «uakseptabel integritet»:

0 = «ok integritet»

1 = «uakseptabel integritet»

I disse analysene vil negative korrelasjoner være å foretrekke. En negativ samvariasjon vil bety at jo høyere opp i verdi i gruppeverdi, altså at integriteten blir uakseptabel, vil skåringen på psykologiske tester, intervju og

gruppeøvelse gå ned. Skulle det forekomme klare positive korrelasjoner vil det bety det motsatte, noe som ville vært bekymringsverdig. Dette vil i så fall bety at høye skårer på tester, intervju og gruppeøvelse gjør det mer sannsynlig at integriteten er uakseptabel.

### *Psykologiske tester*

Psykologiske testresultater ble gjort om fra en 5-delt skala til en z-skåre fordeling for å få en mer finmasket fordeling i testskårer. Analysene viste en signifikant, men svak, negativ korrelasjon mellom Scales sx (a) og gruppeverdi ( $r = -.07$ ). Det var ellers ingen signifikante samvariasjoner med de andre Scales-testene og gruppeverdi ( $r = -.01 - .00$ ). Det var ingen signifikante korrelasjoner mellom Shapes-utledede kompetanser og gruppeverdi ( $r = -.02 - .04$ ).

### *Opptaksnemdvurderinger*

Det var signifikante, men svake, negative korrelasjoner mellom gruppeverdi på den ene siden og intervju *samhandling* ( $r = -.12$ ), intervju åpenhet og inkludering ( $r = -.14$ ) og gruppeøvelse *modenhet* ( $r = -.07$ ) på den andre siden. Det var ellers ingen signifikante samvariasjoner mellom gruppeverdi og med de andre kompetansene som måles i intervjuet og gruppeøvelsen ( $r = -.04 - -.01$ ).

## **4.2 Kull 2014**

### **4.2.1. Korrelasjonsanalyse**

Karakterpoeng fra videregående skole viste også her en sterkest samvariasjon med karaktersnittet. Intervjuskåren på åpenhet og inkludering og språktesten Scales lt-no korrelerte noe lavere men likevel signifikant. De øvrige Scales testene, kompetanseskårene fra Shapes og integritetstesten Squares viste ikke samvariasjon med karaktersnittet. Det ble heller ikke funnet en samvariasjon mellom karaktersnitt og intervjuskårer utover åpenhet og inkludering, ei heller for de tre kompetansene som måles i gruppeøvelsen. Tabell 4.3 viser de

signifikante korrelasjonene i rangert orden. For øvrige korrelasjoner henvises det til vedlegg 2.

Tabell 4.3: Kull 2014: Signifikante korrelasjoner i rangert orden mellom prediktorer og karaktersnitt (N=401)

Prediktor	Korrelasjonskoeffisient <i>r</i>
Karakterpoeng VGS	.41*
Intervju Åpenhet og inkludering	.14*
Scales It-no	.12*

Note. \* $p < .05$  (to-halet)

#### 4.2.2. Multipl hierarkisk regresjonsanalyse

Som for kull 2013, søkte regresjonsanalysen for kull 2014 å svare på om de øvrige variablene hadde prediksjonsverdi når effekten av karakterpoeng ble kontrollert for.

Tabell 4.4 viser at karakterpoeng fra videregående skole forklarte 17 % av variasjonen i karaktersnittet. Variablene samlet forklarte til sammen 18 % av variasjonen, hvor variablene utover karakterpoeng forklarte ytterligere 2 % av variasjonen i karaktersnitt. Når alle variablene sees samlet i steg 2, er karakterpoeng fra videregående skole og intervju-skåre på åpenhet og inkludering signifikante, og økningen i forklart varians er signifikant fra steg 1 til steg 2.

Tabell 4.4: Kull 2014: Hierarkisk regresjonsanalyse med totalt karaktersnitt på PHS som avhengig variabel

	Steg 1	Steg 2
Karakterpoeng VGS	.41**	.39**
Intervju åpenhet og inkludering		.11*
Scales It-no		.06
R <sup>2</sup>	.17**	.18**
R <sup>2</sup> forandring	-	.02*
F-verdi	76.78**	28.54**

Note. \* $p < .05$ ; \*\* $p < .01$

er standardisert regresjonskoeffisient,  $R^2$  er forklart varians i PHS-karaktersnittet,  $F$ -verdien indikerer stegenes signifikans for den avhengige variabelen.  $R^2$  forandring Steg 2 er 0,016.

### 4.2.3. Egnethetsvurdering og integritet

For kull 2014 var 47 søkere registrert med uakseptabel integritet. Samme prosedyre som for kull 2013 ble valgt for å undersøke mulige gruppeforskjeller (punkt-biseral korrelasjon). Negative korrelasjoner vil indikere at de med uakseptabel integritet gjør det svakere i egnethetsvurderingen, mens positive korrelasjoner vil indikere det motsatte.

#### *Psykologiske tester*

Psykologiske testresultater ble også for kull 2014 gjort om fra en 5-delt skala til en z-skåre fordeling for å få en mer finmasket fordeling i testskårer. Det var en signifikant, men svak, negativ samvariasjon mellom gruppeverdi og Scales 1st ( $r = -.06$ ). Det var ellers ingen signifikant samvariasjoner mellom gruppeverdi og de andre Scales testene ( $r = -.02 - .00$ ). Videre var det ingen signifikant sammenheng mellom gruppeverdi og resultater på Shapes-utlede kompetanser (Modenhet  $r = .04$ ; Handlekraft  $r = .00$ ; Samhandling  $r = -.01$ ). Sårbarhetstesten Squares evnet heller ikke å skille signifikant på de med uakseptabel og akseptabel integritet ( $r = .02$  for den totale pålitelighetskåren).

#### *Opptaksnemdvurderinger*

Analysene viste signifikante, men svake, negative korrelasjoner mellom gruppeverdi på den ene siden, og alle kompetansene som måles i intervjuet og på gruppeøvelsen ( $r = -.02 - .11$ ) på den andre siden.

### 4.2.4. Egnethetsvurdering og fullføring av studiet på normert tid

Det var 94 studenter som ikke hadde fullført studiet på normert tid for kull 2014. Som for integritetsvariabelen, ble denne dikotome variabelen kodet slik at 0 indikerte «fullført på normert tid» og 1 indikerte «ikke fullført på normert tid». Det vil også være slik for disse analysene at en negativ samva-

riasjon vil indikere at de som fullfører på normert tid får høyere skårer i egnethetsvurderingen. Skulle det forekomme klare positive samvariasjoner vil det være bekymringsverdig.

#### *Psykologiske tester*

Det var ingen signifikant samvariasjon mellom gruppeverdi og Scales testene ( $r = -.01 - .05$ ) eller gruppeverdi og Shapes-utledede kompetanseskårer ( $r = -.00 - .07$ ). For Squares ble det funnet en signifikant positiv, men svak, samvariasjon for totalskåren *pålitelighet* og gruppeverdi ( $r = .13$ ); begge underfaktorene *impulskontroll* og *etisk bevissthet* har en signifikant positiv, men svak, korrelasjon med gruppeverdi.

#### *Opptaksnemdvurderinger*

Analysene viste en signifikant, men svak, korrelasjon mellom intervju på åpenhet og inkludering og gruppeverdi ( $r = -.08$ ). Det var ellers ingen signifikante samvariasjoner mellom gruppeverdi på den ene siden, og intervju og gruppeøvelse på den andre siden ( $r = -.00 - .01$ ).



## 5 Diskusjon

Diskusjonen vil ta utgangspunkt i problemstillingen: *Viser grad av egnethetsvurderingen sammenheng med faglige prestasjoner på grunnutdanningen, bedømmelsen av integritet og om studiet fullføres på normert tid?*

Egnethetsvurderingen skal ta stilling til nivået på kompetansene åpenhet og inkludering, *modenhet, handlekraft, samhandling, analytisk evne og integritet*. Tre seleksjonsmetoder brukes til dette formålet: psykologiske tester, intervju og gruppeøvelse. I tillegg kommer informasjon fra skikkethetsvurderingen.

De faglige prestasjonene utgjorde de karaktergivende fagemnene. De enkelte fagemnene var strukturert inn i de fem hovedområdene som gjenspeiler hovedoppgavene til en politigeneralist; *politi og samfunn, metode, orden og beredskap (operative oppgaver for kull 2013), forebyggende oppgaver og etterforskning*. I tillegg inngår bacheloroppgaven for kull 2013, som i praksis var en hovedområdeuavhengig prestasjon. For kull 2014 var bacheloroppgaven det eneste karaktergivende bidraget til hovedområdet *metode*, men som likevel kunne inneholde tema fra de øvrige hovedområdene. For kull 2013 gis det karakter i 47 % av de produserte studiepoengene (85 av 180), for kull 2014 59 % (106 av 180). Det totale karaktersnittet gjenspeiler alle fagemner, samt ulike vurderingsformer (se Tabellene 3.2 og 3.3 for detaljer).

Endelig søkte problemstillingen å svare på om det var sammenheng mellom egnethetsvurderingen og bedømmelsen av integritet samt om studiet ble fullført på normert tid. Det foretas en skikkethetsvurdering av søkerne som er med å avgjøre om kompetansen *integritet* er akseptabel, samt at det for kull 2014 var mulig å ta ut data på hvem som ikke hadde fullført studiet innen normert tid.

## 5.1 Egnethetsvurderingen og faglige prestasjoner på grunntidningen

Resultatene viste at karakterpoeng fra videregående skole predikerte faglige prestasjoner i grunntidningen på en god måte. Egnethetsvurderingen traff i riktig retning som målt opp mot faglige prestasjoner, men det var svake og til dels ikke-eksisterende sammenhenger. Det var ingen funn som tydet på at egenskaper og prestasjoner som ble vurdert positivt i seleksjonen hadde en tydelig negativ sammenheng med de senere faglige prestasjonene.

Karakterpoeng hadde en forholdsvis høy samvariasjon med karaktersnittet. Når det ble tatt høyde for karakterpoeng, viste de andre seleksjonsmetodene med signifikante korrelasjoner noe ytterligere forklaringskraft opp mot karaktersnittet. Én av evnetestene, intervjudelen om åpenhet og inkludering, samt gruppeøvelsen bidro med unik prediktiv verdi når det ble kontrollert for prediksjonskraften til skolekarakterer. Imidlertid er disse prediktive verdiene svake. Det er derfor vanskelig å argumentere for en tydelig praktisk merverdi av disse seleksjonsmetodene opp mot positiv seleksjon, selv om funnene var statistisk signifikante.

*Første hovedfunn blir dermed at karakterer fra videregående skole predikerer faglige prestasjoner i grunntidningen på en god måte. Egnethetsvurderingen treffer i riktig retning, men sammenhengene mellom seleksjonsmetodene og faglige prestasjoner er svake og til dels ikke-eksisterende. En økning i skårer på egnethetsvurderingen indikerer derfor ikke en økning i senere faglige prestasjoner på en klar måte.*

### 5.1.1. Videregående skolekarakterer

Korrelasjons- og regresjonsanalysene for begge kull viste at karakterpoeng fra videregående skole predikerte faglige prestasjoner i grunntidningen på en god måte. Spredningen på karakterpoeng for begge kull minimaliserer preseleksjons-problematikk for denne variabelen.

Funnet samsvarer med de to studiene fra befalsskolen og krigskolen i Forsvaret, som også understreker den solide prediktive verdien videregående skolekarakterer har (Kjenstabakk, 2012; Vik, 2013). I meta-analysen til

Schmidt et al. (2016) oppgis det at karakterer fra utdanning har en god samvariasjon med mål på jobbutførelse (0,34). Det ble riktignok ikke oppgitt en verdi når kriteriemålene var knyttet til en opplæringskontekst, og at karakterer fra utdanning i denne metaanalysen også kunne være knyttet til høyere utdanning. En annen meta-analyse oppgir høye korrelasjoner mellom «high school grade point average» og senere akademiske prestasjoner (Westrick, Le, Robbins, Radunzel, & Schmidt, 2015). Det er rimelig å anta at studieveaner fra videregående skole ofte forplantes inn i høyere utdanning, slik at det er mer sannsynlig at gode studieveaner tas med videre, enn at dårlige studieveaner snus om.

### 5.1.2. Psykologiske tester

#### *Evne – og språktester*

For Scales-testene benyttet på kull 2013 hadde alle signifikante korrelasjoner med PHS-karaktersnittet med unntak av språktesten lt-no. Av Scales-testene som ble benyttet for kull 2014 hadde imidlertid kun lt-no en signifikant korrelasjon med PHS-karaktersnittet. Samtlige signifikante korrelasjoner på begge kull var lave, med unntak av Scales sx (a) som hadde en noe høyere korrelasjon enn de andre testene.

Funnene står i kontrast til annen forskning. Det er gjentatte ganger vist i internasjonale studier at evnenivå stort sett er en av de viktigste prediktorene for prestasjoner i både opplæring og i jobbsammenheng, noe metaanalysen til Schmidt et al. (2016) dokumenterer. Kjenstabakk (2012) og Vik (2013) fant at tester på evner var en god prediktor på prestasjoner på henholdsvis befallsskole og krigsskole. Sett opp mot politiseleksjons-funnene til Annell et al. (2015) er den prediktive verdien til Scales sx (a) på omtrentlig samme nivå som den svenske evnetesten, mens den svenske språktesten viser betydelig mer prediktiv verdi enn den norske.

I regresjonsanalysene var det kun Scales sx (a) som bidro med unik prediktiv validitet når det ble kontrollert for karakterpoeng fra videregående. Imidlertid var den prediktive verdien lav, noe som gjør det vanskelig å argumentere for en praktisk nytte av testen i prediksjon av faglige prestasjoner når det tas høyde for karakterpoeng.

Scales testene har altså ikke en prediktiv verdi opp mot karaktersnittet på PHS når det tas høyde for karakterpoeng, med unntak av sx (a). Korrelasjonstabellene i vedlegg 1 og 2 viser dessuten at Scales testene korrelerer svakt med karakterpoeng fra videregående skole; her er igjen sx (a) unntaket med en korrelasjon på 0,20. Til sammenligning oppgir Kjenstabakk (2012) at sesjonstestene på alminnelig evnenivå korrelerer 0,32 med skolepoeng fra videregående skole, samt at Vik (2013) oppgir korrelasjonsverdier på henholdsvis 0,23 og 0,29 mellom to intelligens tester og karaktersnittet som oppnås på krigsskolen. Begge de to forsvarsstudiene har et noe beskjedent utvalg (henholdsvis  $N = 200$  og  $N = 161$ ), samt at fordelingen av testskårer lider av begrenset spredning siden opptaket til Forsvarets skoler gjøres blant preselektert personell (alle har gjennomsnittlig evnenivå eller over). Slik sett er funnene til de to forsvarsstudiene ganske imponerende i favør Forsvarets evnetester.

Siden prediktor – og kriterievariablene i PHS datasettet har god spredning, skulle en forvente en tydeligere sammenheng mellom evne – og språktestene og karaktersnittet, enn det som kommer frem. Funnene samsvarer ikke med cut-e AS sin egen rapportering av testenenes prediktive validitet, som beskrevet i metodekapitlet. Riktignok var disse studiene fra andre bransjer enn en politiutdanning, men det er uansett påfallende at testene ikke viser større samvariasjoner for PHS. Testen sx (a) var eneste test som viste signifikant prediksjonsverdi i denne undersøkelsen, men det er ikke rapportert prediktiv verdi for denne testen i dokumentasjonen mottatt fra cut-e AS.

Funnene setter spørsmålsteget ved Scales-testenes verdi i egnethetsvurderingen. Scales sx (a) viser dog en signifikant brukbar samvariasjon med faglige prestasjoner, men når det tas høyde for skolepoeng fra videregående blir effekten av denne testen lite praktisk nyttig. Funnene retter kritikk mot den kompetanseskåringen som skjer ved hjelp av Scales-testene på *analytisk evne*, og også på *samhandling* hvor språktesten lt-no inngår med en skåre. Det er også meget påfallende at det er et såpass stort avvik mellom hvilke prediktive validitetsverdier cut-e AS rapporterer i sine interne dokumenter for Scales-testene, og hvilke verdier som dokumenteres her i PHS-studien. Forfatteren stiller seg kritisk til begrepsvaliditeten til testene, det vil si om de måler det de er ment å måle.

### *Personlighetstest*

De Shapes-utledede kompetansene *modenhet*, *handlekraft* og *samhandling* viste ingen signifikante korrelasjoner med karaktersnittet for noen av kullene. Dette var kontraintuitive funn når en ser hvilke dimensjoner testen er ment å måle (se Tabell 3.1.). Eksempelvis var det rimelig å tro at nivå av ansvarsbevissthet, selvstendighet og utholdenhet, som lader inn på *handlekraft*, ville gjenspeile seg i prestasjonene på grunnutdanningen. cut-e AS sine egne tall på prediksjonsverdien til Shapes (basic) er gode opp mot intervjukskåre for lærlinger og overordnede vurderinger i en annen (ukjent) bransje, noe som også ga håp om prediksjonsverdi for PHS.

Resultatet er imidlertid i samsvar med Vik (2013) sine manglende funn om prediksjonsverdi for en femfaktor-basert personlighetstest opp mot resultater på Krigsskolen. Anell et al. (2015) viste derimot at de tre trekkene medmenneskelighet, samvittighetsfullhet og emosjonell stabilitet hadde enkelte signifikante sammenhenger opp mot ulike kriterier.

Det er en del forskning som tyder på at grad av ulike personlighetstrekk har prediksjonsverdi for jobbutførelse, spesielt samvittighetsfullhet og nevrotisme (Barrick & Mount, 1991; Furnham, 2008; Schmidt et al., 2016). Det er imidlertid noe uklarhet om enkelte trekk kan tenkes å ha en kurvelineær sammenheng med prestasjoner; altså at for mye av et trekk kan utgjøre en vel så negativ innvirkning, som for lite av et trekk (Le et al., 2011). Det vil i så fall problematisere en lineær matematisk utregning av kompetansenivå ut fra resultat fra en personlighetstest. Et annet problem som nok raskt kan oppstå i en seleksjon er det å svare overdrevent positivt på en personlighetstest. Eksempelvis viste en amerikansk studie at en gruppe politisøkere overdrev sine svar på antatt positive trekk (samvittighetsfullhet og medmenneskelighet) og underrapporterte antatt negative trekk (nevrotisme) på NEO PI-R, en utbredt femfaktor-test (Detrick, Chibnall, & Call, 2010).

Funnene omkring Shapes er dermed noe overraskende, men ikke nødvendigvis i strid med forskning på området, som heller ikke er entydig. Shapes er bygget opp etter en annen modell enn femfaktor-teorien, noe som kan gjøre sammenligningen med «main stream» personlighetsforskning problematisk.

Dimensjonene i testen er tuftet på antagelser om arbeidslivskompetanser og jobbatferd, og ikke generiske personlighetstrekk. Slik sett kan Shapes tenkes å ha mer prediksjonsverdi i det senere politiryket, enn overfor faglige prestasjoner i grunnutdanningen.

På grunn av svarsystemet i Shapes kan testen virke delvis ipsativ, noe som vil si at testtakeren måles opp mot sine egne svar i tillegg til en ytre norm. Ipsative personlighetstester byr på særegne utfordringer i valideringsstudier, eksempelvis ved at dimensjonene som utregnes kan komme til å representere kombinerte, fremfor «rene», personlighetstrekk (Martinussen, Richardsen, & Vårum, 2001). Dette kan også være med på å forklare de manglende positive funnene omkring Shapes (basic).

En kan imidlertid stille seg spørsmålet om hvorfor Shapes (basic) brukes til å predikere «utdanningsatferd» når fundamentet til testen er rettet mot arbeidslivet. Forfatteren kjenner som tidligere nevnt imidlertid ikke til bakgrunnen for valget (og bruken) av testene, men det antas at Shapes (basic) ble en treffende test å benytte siden kompetansene til PHS blant annet er utledet av Shapes-modellen til cut-e AS.

Den gangen AgendaKaupang (2012) gjorde sin vurdering ble det påpekt at Shapes (basic) skapte usikkerhet blant opptaksnemdene, og det ble anbefalt at testen fjernes (eventuelt at det trengtes en mer omfattende testopplæring om testen fortsatt skulle brukes). Funnene om Shapes (basic) sin manglende prediktive validitet for faglige prestasjoner støtter heller ikke den videre bruken av testen, og det er spesielt problematisk at resultatene brukes mekanisk for kompetanseskåring.

I likhet med Scales-testene er det også for Shapes meget påfallende å se avviket i hvilke prediktive verdier som oppgis i cut-e AS sine interne dokumenter, og hvilke verdier som dokumenteres her i PHS-studien. Forfatteren synes også det er underlig at ikke personlighetstrekket nevrotisme/emosjonell stabilitet blir forsøkt målt, all den tid forskning viser relevansen av dette trekket for jobbutførelse (Furnham, 2008). Annel et al. (2015) fant at trekket har en sammenheng (om enn noe svak) med prestasjoner i praksisperiode og for den generelle trivselen blant svenske politistudenter. Ikke minst har trekket en hel del å si sett opp mot mental robusthet (også kalt resiliens), som

blant annet vist gjennom en studie på søkere til militær utdanning i Norge (Friborg, Barlaug, Martinussen, Rosenvinge, & Hjemdal, 2005).

### 5.1.3. Intervjuet

Kompetanseskåringen som ble gjort på bakgrunn av intervjudelen om åpenhet og inkludering korrelerte signifikant med karaktersnittet for begge kull, og oppnådde også en prediktiv verdi når det ble kontrollert for karakterpoeng. Korrelasjonene kan tyde på at aspekter ved kompetansen ble målt når karakterene ble satt i grunnutdanningen. Imidlertid kom kompetansemålingen ut med en svak prediktiv forklaringskraft i regresjonsanalysene, slik at det er vanskelig å få øye på den praktiske nytteverdien når det tas høyde for karakterpoeng.

De tre andre kompetansene som skåres i intervjuet, *modenhet, samhandling* og *handlekraft*, hadde ingen prediktiv verdi opp mot karaktersnittet. Disse tre kompetanseskårene korrelerte svært høyt med de Shapes-utledede kompetansene, noe som ikke var overraskende siden intervjuer har testskårene fra Shapes tilgjengelig og gir hypoteser til søkerne basert på testen. Dette er problematisk siden Shapes-utledede kompetanseskårer ikke hadde prediksjonsverdi opp faglige prestasjoner, og slik sett kan redusere den prediktive validiteten til intervjuet. Hadde intervjuer *ikke* hatt Shapes (basic) å støtte seg på, slik som tilfellet var ved vurdering av åpenhet og inkludering, kunne kanskje den prediktive validiteten av de tre andre intervjuutledede kompetanseskårene økt.

Intervjuet beskrives av og til som en sårbar seleksjonsmetode på grunn av en rekke mentale fallgruver som forringer objektiviteten i vurderingene (Skorstad, 2015). Både strukturerte og ustrukturerte intervjuer får imidlertid ganske solide prediksjonsverdier i følge metaanalysen til Schmidt et al. (2016), også når det tas høyde for evnetester. De to studiene fra Forsvaret som er nevnt flere ganger ovenfor fant imidlertid ingen god prediktiv verdi på intervjuene som ble gjennomført. Annell (2015) rapporterer at det semi-strukturerte intervjuet for politistudent-opptaket i Sverige hadde en svak prediktiv verdi opp mot enkelte kriterier. Det hersker dermed en del ulike funn omkring

den prediktive verdien av et opptaks – eller jobbintervju, og hvorvidt grad av struktur vil bedre intervjuets evne til å forutse fremtidige prestasjoner.

Et opptaks – eller jobbintervju har imidlertid flere aspekter ved seg enn å kun være en arena for å predikere fremtidige prestasjoner. Eksempelvis er intervjuet en arena hvor kandidaten møter organisasjonen han/hun søker til, hvor det typisk er mulighet til å stille spørsmål og høste mer informasjon. Graden av struktur på intervjuet kan selvfølgelig begrense en slik type kommunikasjon mellom kandidat og intervjuer.

Oppsummert ser det ut til at opptaksintervjuet ved PHS har den samme svake prediktive validiteten som typisk sees i andre sammenlignbare lokale validitetsstudier. Det er likevel av verdi at intervjuet kan representere flere formål enn det rent prediksjonsmessige. Det også tenkes at intervjuet er en egnet arena for negativ seleksjon, i og med dette er eneste seleksjonsmetode hvor opptaksnemda møter søkerne i en samtale én-til-én. Den høye graden av struktur på intervjuet kan nok være både en fordel og ulempe sett opp mot negativ seleksjon; det er lett å skrive søkere ut om de svarer svakt etter intervjuguiden, men ikke fullt så lett om det er påfallende trekk ved søkeren som ikke dekkes av guiden.

Når det gjelder påfallende trekk, er det trolig også begrensninger i den medisinske undersøkelsen opp mot det som står om «psykisk helsekrav og sosial tilpasning». Utredning av en symptomlidelse, for ikke å snakke om en personlighetsforstyrrelse, er et ressurskrevende klinisk arbeid (Rønnestad & von der Lippe, 2009). Arbeidet bak en «komplett utfylt legeerklæring» vil neppe være dekkende for en grundig klinisk evaluering av de negative seleksjonskriteriene som nevnes omkring psykisk helsekrav og sosial tilpasning. Enkelte søkere med psykopatologisk fungering kan også være preget av manglende selv-innsikt, samt at enkelte kan holde tilbake informasjon om sine psykiske helseplager.

Annell (2015) sitt forslag var for øvrig å gjøre det svenske intervjuet mer fokusert på negativ seleksjon, men den svenske modellen er annerledes med to semi-strukturerte intervjuer på samme søker fra henholdsvis politipersonell og psykolog. Intervjuvurderingene blir deretter sammenlignet. En slik modell brukes også ved opptak til enkelte avdelinger i Forsvaret, hvor hensikten er



så vel negativ som positiv seleksjon. Det er dog ingen systematiske studier av validiteten til denne praksisen i Forsvaret ut fra forfatterens kjennskap. AgendaKaupang (2012) anbefalte forøvrig å bygge ut intervjuet, noe som ville vært mulig om Shapes ble fjernet. Tilbakelesningen av hypoteser fra Shapes (og Squares fra kull 2014 av), tar en del tid i intervjuet.

#### 5.1.4. Gruppeøvelsen

For det ene kullet viste gruppeøvelsen signifikant (men svak) samvariasjon med karaktersnittet, samt en lav prediktiv verdi når det ble kontrollert for videregående karakterpoeng. De tre kompetansene som måles, *modenhet*, *handlekraft* og *samhandling*, hadde en høy samvariasjon med hverandre. Det kan derfor problematiseres om denne metoden er finmasket nok til å skille på de tre kompetansene, og om de tilhørende observasjonsskjemaene med atferdsindikatorer er treffende nok med tanke på hvilken atferd som vises i denne øvelsen. Gruppeøvelsens prediktive verdi fra regresjonsanalysen er basert på gjennomsnittsskåre på de tre kompetansene, og det vil også gjelde her at denne verdien er såpass svak at en praktisk nytteverdi er tvilsom når det kontrolleres for karakterpoeng.

Sammenlignbare seleksjonsmetoder og deres validitetsverdier er i følge Schmidt et al. (2016) 0,36 for «assessment centers» og 0,33 for «work sample tests», men tilleggsforklaringen til slike metoder når det tas høyde for evnetester er svært beskjeden. Det ble anbefalt fra AgendaKaupang (2012) sin side å bygge ut gruppeøvelsen, i tillegg til intervjuet. Både med bakgrunn i undersøkelsens funn, og i forskningen generelt, kan verdien av gruppeøvelsen i sin nåværende form trekkes i tvil. I Forsvaret fant riktignok Kjenstabakk (2012) at feltperioden som brukes i opptaket til befalsskolene hadde en god prediktiv verdi. Imidlertid gikk denne praktiske delen av seleksjonen over et lengre tidsrom enn det gruppeøvelsen ved PHS gjør.

## 5.2 Egnethetsvurderingen og bedømmelsen av integritet

For begge kull sammenlagt var det 88 søkere med uakseptabel integritet, som vurdert ut fra skikkethetsvurderingen. Siden skikkethetsvurderingen inne-

bærer søk i politiets registre og tolkning av det som fremkommer der av spesielt egnet politipersonell, vil integritetsbedømmelsen være tuftet på verdifull informasjon av sikkerhetsmessig karakter. Slik informasjon er normalt ikke tilgjengelig i utvelgelsesprosesser. Det ble derfor interessant å se om egnethetsvurderingen hadde samvariasjon opp mot et slikt kriterium. Spesielt interessant var det at PHS benyttet en integritetstest (Squares) for kull 2014; denne testen er nettopp ment å predikere kontraproduktiv jobbatferd (cut-e, 2013).

For begge kull var det enkelte signifikante samvariasjoner mellom kompetansemålene i intervjuet og i gruppeøvelsen på den ene siden, og integritetsbedømmingen på den andre siden. Korrelasjonene var imidlertid svake. Eksempelvis vil den høyeste korrelasjonen (med intervju åpenhet og inkludering,  $r = -.14$ ) kun indikere (.14x.14= 0,02) 2 % felles variasjon mellom de to variablene. Det er derfor vanskelig å se at kompetanseskåringene fra intervjuet og gruppeøvelsen er egnet i praksis til å predikere hvem som hadde akseptabel eller uakseptabel integritet.

I metaanalysen til Schmidt et al. (2016) kommer integritetstester høyt opp i sin evne til å predikere både jobbutførelse (0,43) og opplæring (0,46), og tilleggsforklaringen av integritetstester er solid også når det tas høyde for evnetester. Integritet fra disse forskernes side er først og fremst forstått som et fenomen som samsvarer høyt med personlighetstrekket samvittighetsfullhet (bredt definert), noe også en tidligere omfattende meta-analyse av ulike integritetstester understreker (Ones, Viswesvaran, & Schmidt, 1993). I og med lav grad av trekket blant annet indikerer impulsivitet og lav planmessighet er et slikt samsvar rimelig. Squares korrelerer signifikant med Shapessutledede kompetanser (.14 – .21), hvorav den høyeste samvariasjonen er å finne med *handlekraft* (som ligner en del på samvittighetsfullhet). Korrelasjonene er ikke høye, men kan peke på et lite slektskap mellom egenkapene som Squares måler (etisk bevissthet og impuls kontroll) og personlighetsdimensjoner knyttet til samvittighetsfullhet.

Dersom integritet kan måles på en god måte er dette naturligvis en viktig prediktor i en seleksjonsprosess, både med tanke på negativ og positiv seleksjon. Siden Squares ikke viste sammenheng med hvem som gjennom skikkethetsvurdering ble funnet akseptable og ikke, betviles det at testen er i stand

til å identifisere søkere med påviste kontraproduktive tilbøyeligheter. Testen viste for øvrig heller ikke sammenheng med faglige prestasjoner på utdanningen, noe den kunne tenkes å gjøre på grunn av de krav til etisk bevissthet og impuls kontroll som kreves i enkelte av fagemnene. cut-e AS oppgir gode prediksjonsverdier for et lite utvalg i en norsk varehusbedrift, men et slikt positivt funn ble ikke reproduisert her. I likhet med Scales-testene og Shapes (basic), er det også for Squares et påfallende avvik mellom oppgitte prediktive validitetsverdier fra cut-e AS og det som dokumenteres her i PHS-studien.

Det er dermed grunn til å være kritisk til bruk av denne testen i egnethetsvurderingen. Nå er riktignok hypotesene omkring impuls kontroll og etisk bevissthet ved lave skårer formulert forsiktig og rundt til søkerne, noe som gjør at formuleringene i seg selv ikke er påfallende stigmatiserende (personlig kommunikasjon, cut-e AS, oktober 2017). Det kan imidlertid bli problematisk om en søker som skårer lavt søker opp testen på internett, og ser at den er ment å si noe om sannsynlighet for «kontraproduktiv atferd i jobbsammenheng». Det alvorlige er imidlertid at testen ikke ser ut til å fange opp sannsynlighet for kontraproduktiv jobbatferd, som vist ved at PHS søkere som har vandelsmerknader har like skårer som de som ikke har vandelsmerknader. Begrepsvaliditeten til testen trekkes derfor i tvil.

*Andre hovedfunn i undersøkelsen blir dermed at egnethetsvurderingen ikke har en klar sammenheng med integritetsbedømmelsen. Dette til tross for at det inngår en integritetstest/sårbarhetstest i egnethetsvurderingen.*

### **5.3 Egnethetsvurderingen og fullføring av studiet på normert tid**

Det var 94 studenter på kull 2014 som ikke hadde fullført studiet på normert tid. For denne gruppen var det ulike fag som ikke ble bestått/avlagt eksamen i, hvor noen manglet flere fag mens andre bare ett. Tross variasjonen i gruppen som ikke fullførte på normert tid, er det interessant å bruke dette som et kriterium å måle egnethetsvurderingen opp mot. Det er selvfølgelig et mål å velge ut søkere som tilpasser seg grunnutdanningen på en slik måte at et

normert og ordinært studieforløp gjennomføres.

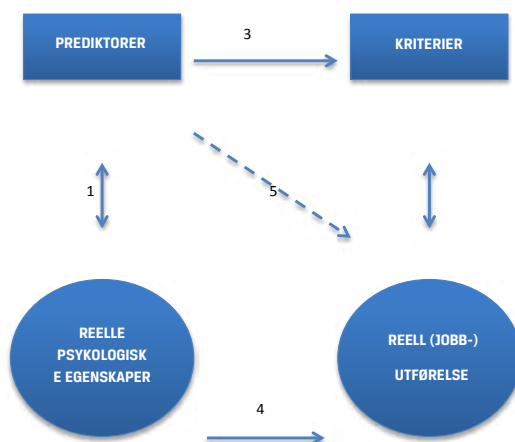
For kull 2014 ble det funnet signifikante korrelasjoner mellom sårbarhetstesten Squares og intervju om åpenhet og inkludering på den ene siden, og om studiet ble fullført på normert tid eller ikke på den andre siden. Sammenhengene var også her såpass svake at en praktisk nyttig effekt nok ikke var til stede. Det er likevel interessant at retningen på korrelasjonen med Squares var i «feil» retning. Dersom sammenhengen var sterkere i positiv retning enn det den faktisk var, ville det betydd at de som ikke fullfører studiet på normert tid skårer bedre enn de som fullfører på normert tid. Dette funnet er også med på å rette et kritisk blikk mot Squares.

*Tredje hovedfunn i undersøkelsen blir dermed at egnethetsvurderingen ikke har en klar sammenheng med om studiet fullføres på normert tid eller ikke.*

## 5.4 Begrensninger ved undersøkelsen

Enhver undersøkelse av hvorvidt nåtidige målinger kan predikere senere atferd og prestasjoner møter på mange begrensninger. Figur 5.1. kan være til hjelp i å beskrive hovedbegrensningene som er aktuelle i denne undersøkelsen.

Figur 5.1: Sammenheng mellom prediktorer og kriterier og deres underliggende fenomener (bearbejdet fra Ployhart (2012))



Den første utfordringen vises i pil 1, hvor de målbare prediktorene i mer eller mindre grad gjenspeiler de egenskapene de er ment å måle. Det vil alltid være forhold som vil forurense påliteligheten og gyldigheten (reliabiliteten og validiteten) til målemetoder på psykologiske egenskaper (og for så vidt også andre egenskaper, som fysiske og medisinske). Undersøkelsen har ikke på en systematisk måte undersøkt eller diskutert prediktorreliabilitet –og validitet spesifikt, men det vil være knyttet feilkilder til de psykologiske testene, intervjuet og gruppeøvelsen som brukes. Det er usikkert i hvor stor grad disse målemetodene faktisk måler kompetansene de er ment å måle. Dokumentasjonen til de psykologiske testene er utelukkende interne dokumenter fra leverandøren cut-e AS, noe som er en svakhet siden fagfellevurderte artikler om psykometriske egenskaper ved testene mangler. DNV har sertifiseringsordning for tester<sup>5</sup>; i listen som ligger åpent på internett kan ikke forfatteren se at testene som benyttes på PHS er sertifisert (pr. desember 2017). Cut-e AS har imidlertid sertifisering på flere andre tester. Det er også en begrensning til studien at data fra fysiske tester og medisinske undersøkelser ikke var tilgjengelig. Den prediktive validiteten til fysiske tester hadde vært svært aktuell å undersøke på grunn av Annell (2015) sine tydelige positive funn fra Sverige.

Hovedbegrensningen til undersøkelsen er imidlertid kriterienes form, illustrert i pil 2. I hvilken grad gjenspeiler karakterprestasjoner på grunnutdanningen egnethet som politistudent og senere som ansatt? I et mer optimalt forskningsdesign skulle også vurderinger fra praksisåret B2 inngått som kriterier. Siden kompetansene også er ment å gjelde for politipersonell i en yrkeshverdag, skulle det også ideelt sett inngått arbeidsprestasjonsvurderinger som kriterier. Slik ville det ha ligget flere kriterier til grunn for å sammenligne egnethetsvurderingen opp mot, noe som hadde gitt en økt gyldighet av kriteriene opp mot den reelle utførelsen som student og ansatt. Det var imidlertid knyttet utfordringer til å samle inn data fra B2. I perioden undersøkelsen pågikk hadde ikke praksisåret B2 etablerte systemer for graderte vurderinger av studentene, og kompetansene det selekteres etter ble ikke brukt systematisk i praksisåret.

---

5 <https://www.dnvgl.no/services/testsertifisering-52540>

Det er usikkert om den sertifiserte testen Scales verbal er samme test som Scales verbal admin. Shapes (management) er sertifisert, men Shapes (basic) står ikke oppført.

I skrivende stund (desember 2017) er B2 vurderingene heller ikke samsvarende med kompetansene studentene selekteres etter. Det var heller ikke gjennomførbart å samle inn data fra politiansatte studenter, da politidirektoratet heller ikke hadde systemer for ansattvurderinger ut fra kompetansene<sup>6</sup>.

Pil 3 peker på hvordan sammenhengen mellom de tilgjengelige prediktorer og kriteriene undersøkes. I denne undersøkelsen ble korrelasjons- og regresjonsanalyser valgt som statistiske metoder. En potensiell begrensning i slike kvantitative metoder er det som ofte kalles tredjevariabelproblemet. Dette betyr at det kan være ikke-observerte forhold som smitter inn på de påviste sammenhengene mellom prediktorer og kriterier. En påvist samvariasjon eller prediktiv verdi kan være befengt med påvirkning fra andre ukjente fenomener enn de to variabelsettene som sammenlignes. Det kan for eksempel tenkes at et godt klassemiljø og solid lærerkvalitet vil være betydningsfullt for prestasjonene til en del studenter, noe som eksempelvis kan gjøre at de med svake skårer på tester likevel kan gjøre det bra. Motsatt kan et dårlig klassemiljø og suboptimale lærere innvirke negativt på en begavet students prestasjoner. Slike forhold kan påvirke sammenhengen mellom seleksjonstester og utdanningsresultater, enten ved at korrelasjonene blir overdrevet eller underdrevet. I menneskelige sammenhenger er det naturligvis ikke alle fenomener som lar seg tallfeste, og metodene som ble valgt for å undersøke pil 3 er nødvendigvis en forenkling av den reelle sammenhengen.

Pil 4 illustrerer hva som er sammenhengen mellom de reelle psykologiske egenskapene seleksjonsmetodene i mer eller mindre grad er treffsikre i å måle, og den reelle utførelsen av studiet eller arbeidet som kriteriene i varierende grad gjenspeiler. Vi vil aldri kunne gjøre en kvantitativ undersøkelse som innehar feilfrie mål på abstrakte fenomener som psykologiske egenskaper, eller komplekse atferdsmønstre som jobbutførelse. Til det er mennesker for komplekse, men det vil alltid være et mål for kvantitativ forskning å tilnærme seg mest mulig optimale målinger av de fenomenene som skal undersøkes. Undersøkelsen har heller ikke tatt for seg å diskutere om nettopp de

---

<sup>6</sup> Et samarbeid med POD HR ble igangsatt, men ble ikke fulgt opp på grunn av deltidsansettelsen knyttet til undersøkelsen, med de begrensninger i tid det innebar. Ved første møte med POD HR i januar 2016 var ikke personellet der klar over hvilke kompetanser PHS selekterer etter.

seks kompetansene som er valgt er de som best mulig beskriver egnetheten til en politistudent og en politigeneralist. Forfatteren stiller seg imidlertid undrende til at cut-e AS sin modell med 18 såkalte «Shapes-kompetanser» var sentral i arbeidet med å utlede kompetansene (Abraham, 2012). Forfatteren kjenner ikke jobbanalyseprosessen i detalj, men det kunne tilhørt en annen studie å gå den mer i sømmene for å vurdere om det er godt nok samsvar mellom cut-e AS sin kompetansemodell og personellkrav i en sikkerhetsorganisasjon med de operative krav som gjør seg gjeldende der. Det er imidlertid liten tvil om at jobbanalysen i sin tid var et omfattende arbeid, slik det er beskrevet av Abraham (2012).

Endelig er pil 5 selve hovedformålet med seleksjon. Jo mer treffsikkert de målbare prediktorene kan forutsi den reelle jobbutførelsen, desto mer valid og praktisk nyttig er seleksjonssystemet. Som illustrert ved hjelp av de andre pilene, vil denne sammenhengen avhenge av en rekke forhold som har med målingsfeil og en kompleks menneskelig virkelighet å gjøre.

## 5.5 Styrker ved undersøkelsen

En klar styrke ved undersøkelsen er antallet som inngår i datasettet. Til tross for en del mangler ved utregning av det totale karaktersnittet, ble utvalget likevel høyt ( $N=953$ ). Sammenhengene som ble funnet i form av tilstedeværende eller manglende samvariasjoner og prediktive verdier er sikrere med et stort utvalg enn med et lite. Med den høye utvalgsstørrelsen har undersøkelsen en god generell gyldighet omkring egnethetsvurderingen ved PHS, og ikke bare for de to kullene som ble undersøkt. Slik sett er studien egnet til å underbygge funderte konklusjoner om hvordan seleksjonsmetodene fungerer.

En annen styrke ved undersøkelsen er spredningen i prediktor – og kriterieskåringene, som vist i Vedleggene 3 og 4. I en del lokale prediktive valideringsstudier vil det ofte være mer begrenset spredning i datasettet, siden det som regel er de med høyest vurderinger som produserer kriteriedata. Imidlertid finnes det metoder for å korrigere for begrenset spredning, noe som er forholdsvis vanlig (beskrevet i kap. 2).

Egnethetsvurderingen ved PHS er strukturert og standardisert. Dette betyr

at alle søkere er behandlet tilnærmet likt, noe som øker prediktorenes egnethet for statistiske analyser. Det samme vil gjelde de tilgjengelige kriteriene. Karakterfesting av studentenes prestasjoner er en standardisert beslutning som hviler på generelle beskrivelser av karakternivåene, samt fagemnespesifikke sensorveiledninger. Kriteriene ansees dermed mer objektive enn subjektive vurderinger av studentenes prestasjoner, fordi den enkelte faglærer må forholde seg til satte kriterier for prestasjonsbedømmingen.

Det ville vært vanskelig å få samme standardiserte og objektive vurderingsformer i praksisåret B2 eller i yrkeslivet, da vurdering av mer komplekse forhold ved prestasjoner vil øke det tilfeldige subjektive innslaget hos den eller de som skal vurdere prestasjonen. Det vil også være slik at ulikheter på praksissteder med ulik variasjon i konkrete arbeidsoppgaver kan gjøre det vanskelig å sammenligne studenter. Slik sett er det ikke selvsagt at praksisvurderinger vil være egnet til å belyse treffsikkerheten til egnethetsvurderingen, om da ikke kvaliteten på slike kriteriemål er solid (jf. pil 2 i figuren).

Det er også en styrke ved undersøkelsen at kriteriene er tuftet på ulike teoretiske og praktiske vurderingsformer, hvor det stort sett ble stilt krav til en polisier tilnærming til løsning av eksamensoppgaver uavhengig av basalfaget pensumet hviler på. Et eksempel er muntlig eksamen i etterforskning i B3, hvor kandidaten fikk utdelt et tekstcase om en mulig straffbar handling. Kandidaten skulle først gå gjennom sine vurderinger av hva som er relevante etterforskningsskritt med bakgrunn i politifaglige, juridiske og psykologiske (eventuelt kriminaltekniske) fagemner, for deretter å presentere sin løsning til et sensorpanel bestående av politietterforsker, jurist og psykolog (eventuelt kriminaltekniker). Beslutning om karakternivå ble foretatt gjennom en helhetlig tilnærming til kandidatens prestasjon, basert på en sensorveiledning som beskriver en polisier fremgangsmåte og praktisk bruk av pensum fra de enkelte fagemner.



## 6 Konklusjon

Egnettsvurderingen ved PHS har blitt mer fundert, standardisert og forskningsforankret enn det som tidligere gjaldt (Abraham, 2012). Samarbeidet mellom PHS og HR-spisskompetente konsulenter ved cut-e AS har sørget for et gjennomtenkt seleksjonssystem «etter boka». Det er nå klarhet i hva egnettsvurderingen baserer seg på. Seleksjonsmetodene psykologiske tester, kompetansebasert strukturert intervju og atferdsindikatorbasert gruppeøvelse, sammen med opplæring og veiledning av opptaksnemndmedlemmer, gir umiddelbare assosiasjoner til et profesjonelt seleksjonssystem. Det nye systemet har i det store og hele fått gode kvalitetsvurderinger i tidligere evalueringer (AgendaKaupang, 2012). Imidlertid har det inntil nå manglet en prediktiv studie av PHS sin egnettsvurdering: fungerer det strukturerte og gjennomtenkte seleksjonssystemet slik det skal?

Undersøkelsens kvantitative tilnærming i form av en prediktiv valideringsstudie bidrar med ytterligere perspektiver på opptaket til PHS. Funnene fra undersøkelsen med kull 2013 og 2014 viser at egnettsvurderingen treffer i riktig retning som målt opp mot faglige prestasjoner i grunnutdanningen, men at styrken på retningen er lite praktisk nyttig.

Det er signifikante samvariasjoner mellom enkelte seleksjonsmetoder og faglige prestasjoner. De er alle imidlertid forholdsvis lave ( $r = .10 - .18$ ), noe som er påfallende med tanke på at det er en god spredning i datasettet. Det er spesielt overraskende at ikke Scales-testene har en høyere samvariasjon med faglige prestasjoner, siden forskning ofte viser at evne- og språktester er en god prediktor for utdanningsresultater. Når korrelasjonsnivåene er såpass lave blir det vanskelig å få øye på den praktiske nytteverdien opp mot positiv seleksjon. Videre tok regresjonsanalysene hensyn til den forholdsvis høye samvariasjonen karakterpoeng fra videregående skole hadde med faglige prestasjoner. De seleksjonsmetodene som samvarierte signifikant med faglige prestasjoner, oppnådde i disse analysene en svært lav forklaringskraft utover

det karakterpoeng forklarer. Også dette gjør det vanskelig å få øye på den praktiske nytteverdien av seleksjonsmetodene opp mot positiv seleksjon.

Konklusjonen er dermed at en økning i skårer på egnethetsvurderingen ikke indikerer en økning i senere prestasjoner på en klar måte. Seleksjonsmetodene fungerer derfor ikke så godt som de burde.

Egnethetsvurderingen viser også svakheter sett opp mot integritetsbedømmelsen og om studiet fullføres på normert tid. Dette er spesielt overraskende for datasettet til kull 2014, siden testen Squares ble benyttet her. Testen oppgis å være en integritetstest, men evner ikke å skille på de som har vandelmerknader og de som ikke har det. Hypotesene som gis til søkere som har lave skårer på Squares er riktignok forsiktige og lite stigmatiserende, slik at konsekvensene for den enkelte søker nok ikke er alarmerende.

Det er imidlertid grunn til å spørre om integritetstesten måler det den er ment å måle, nemlig sannsynlighet for kontraproduktiv jobbatferd. Forhold som har med personellsikkerhet å gjøre er naturligvis av stor relevans for sikkerhetsorganisasjoner som politiet og Forsvaret. Det må derfor stilles høye krav til tester og metoder som er ment å si noe om slike forhold (personlig kommunikasjon, Forsvarets sikkerhetsavdeling, desember 2017).

For integritetstesten og de andre psykologiske testene på evner, språk og personlighet, er det svært påfallende at det er et såpass stort avvik i hvilke validitetsverdier cut-e AS rapporterer i sine interne dokumenter og det som kommer frem her i denne undersøkelsen. Avvikene reiser flere spørsmål, hvor forfatteren blant annet stiller seg undrende til begrepsvaliditeten til testene. Måler testene det de er ment å måle? Undersøkelsens funn underbygger en forholdsvis sterk kritikk av testbatteriet PHS benytter. PHS burde kunne forvente større prediksjonsverdier av psykologiske tester enn det som dokumenteres i denne undersøkelsen. Ikke minst burde søkerne selv kunne forvente det.

Undersøkelsen har dog ikke svart på om egnethetsvurderingen evner å identifisere de som er best egnet til praksisåret B2, eller til det senere yrket. Så kanskje er det slik at «halve spørsmålet» er besvart. Det bemerkes imidlertid at karaktergivende eksamener på PHS krever en polisier fremgangsmåte, slik at det tradisjonelle skillet mellom «akademiske» og «operative/praktiske» fag er en forenkling av de faktiske forhold i grunnutdanningen. Etter at grunnutdan-

ningen ble sentrert rundt de fem hovedområdene som gjenspeiler hovedoppgavene til en politigeneralist, er denne todelingen mindre relevant enn det den kanskje var. Uavhengig av dette er det imidlertid en begrensning ved undersøkelsen at vurderinger av studentenes kompetansenivå i B2 ikke var tilgjengelig.

Undersøkelsens design er formet slik at det er forventet at en økning i egnethetsvurderingsskåre skal indikere en økning i senere prestasjoner. I den praktiske seleksjonen er opptaksnemdene avhengig av et så godt som mulig beslutningsgrunnlag å vurdere egnethet ut fra, også for å unngå at ikke-egne søkere kommer inn. Et perspektiv på hva som er fornuftige kriterier på negativ seleksjon og uønskede kandidater kan imidlertid være en undersøkelse i seg selv. Det er naturligvis like viktig å unngå at uønskede studenter kommer inn, som det er å velge de best egnede for utdanning og videre yrke i politiet.

Denne rapporten har ikke tatt sikte på å anbefale konkrete endringer i egnethetsvurderingen, men den kan naturligvis tjene som et beslutningsgrunnlag for eventuelle justeringer og endringer i den fremtidige egnethetsvurderingen ved Politihøgskolen. Formålet har vært å gi et statistisk begrunnet situasjonsbilde av seleksjonsmetodenes treffsikkerhet, hvor enkelte metoder har fått mer kritikk enn andre. Forfatteren er spesielt kritisk til de psykologiske testene, og har stilt spørsmålsteget ved begrepsvaliditeten til Scales-testene, Shapes (basic) og Squares på grunn av de manglende sammenhengene testene viser med kriteriene.

Forfatteren håper avslutningsvis at rapporten også kan bidra til økt bevissthet rundt rekruttering og seleksjon i politi- og lensmannsetaten, spesielt der det leies inn eksterne rekrutteringsbyråer.

## 7 Referanseliste

- Abraham, S. (2012). *Opptak til Politihøgskolen: Utvikling av ny seleksjonsmetode*. (Spesialistoppgave i organisasjonspsykologi, Norsk psykologforening.).
- AgendaKaupang. (2012). *Evaluering av ny arbeidsform for Politihøgskolens opptaksnemnder*. Høvik: Agenda Kaupang AS.
- Annell, S. (2012). *Vilka sökande antas till och påbörjar polisutbildningen?* Rekryteringsmyndigheten.
- Annell, S. (2015). *Hållbar polisrekrytering. Teoretiska, metodologiske och praktiske perspektiv på rekrytering og urval.*, Stockholms universitet.
- Annell, S., Lindfors, P., & Sverke, M. (2015). Police selection - implications during training and early career. *Policing: An International Journal of Police Strategies & Management*, 38(2), 221-238. doi: 10.1108/PIJPSM-11-2014-0119
- Barrick, M. R., & Mount, M. K. (1991). The Big-Five personality dimensions in job-performance: A meta-analysis. *Personell Psychology*, 44(1), 1-26. doi: 10.1111/j.1744-6570.1991.tb00688.x
- cut-e. (2013). *Squares inside, detect who is going to cause problems*. Hamburg.
- cut-e. (2014). *Testdokumentasjon - scales sx-a*. Oslo.
- cut-e. (2015). *Tester og rapporter*. Oslo.
- cut-e. (2016). *Testdokumentasjon - scales ix*. Oslo.
- cut-e. (2017a). *cut-e validation studies - relevant for PHS*. Oslo.
- cut-e. (2017b). *Testdokumentasjon - shapes (basic)*. Oslo.
- Detrick, P., Chibnall, J. T., & Call, C. (2010). Demand effects on positive response distortion by police officer applicants on the Revised NEO Personality Inventory. *Journal of Personality Assessment*, 92(5), 410-415. doi: 10.1080/00223891.2010.497401
- Einarsen, S., & Skogstad, A. (Red.). (2012). *Det gode arbeidsmiljø. Krav og utfordringer*. (2 ed.). Bergen: Fagbokforlaget.

- Friborg, O., Barlaug, D., Martinussen, M., Rosenvinge, J. H., & Hjemdal, O. (2005). Resilience in relation to personality and intelligence. *International Journal of Methods in Psychiatric Research*, 14(1), 29-42. doi: 10.1002/mpr.15
- Furnham, A. (2008). *Personality and Intelligence at work: exploring and explaining individual differences at work*. East Sussex: Routledge.
- Ganzach, Y. (1998). Intelligence and job satisfaction. *Academy of Management Journal*, 41(5), 526-539. doi: 10.2307/256940
- Gonzales-Mulé, E., Mount, K. M., & Oh, I. S. (2014). A Meta-Analysis of the Relationship Between General Mental Ability and Nontask Performance. *Journal of Applied Psychology*, 99(6), 1222-1243. doi: 10.1037/a0037547
- Hansen, I. (2006). *Bidrag til psykologitjenestens historie i Forsvaret fra 1946–2006*. Oslo: Forsvarets skolesenter.
- Harlem, T. (2016). *Seleksjon flygere F-35*. Forsvarets Høgskole.
- Hove, K. (2012). *Politiutdanning i Norge - fra konstabelkurs til bachelorutdanning*. Oslo: Politihøgskolen.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement, and Conceptual Issues. I O. P. John, R. W. Robins & L. A. Pervin (Red.), *Handbook of personality: Theory and research* (s. 114-158). New York: Guilford Press.
- Kaplan, R. M., & Saccuzo, D. P. (2009). *Psychological Assessment and Theory: Creating and Using Psychological Tests*. (8 utg.). Canada: Wadsworth Cengage Learning.
- Kile, S. M. (1990). *Helsefarlige ledere og medarbeidere*. Oslo: Hjemmets bokforlag.
- King, R. E. (2014). Personality (and Psychopathology) Assessment in the Selection of Pilots. *The International Journal of Aviation Psychology*, 24(1), 61-73. doi: 10.1080/10508414.2014.860844
- Kjenstabakk, T. J. (2012). *Seleksjon til befalsskolen: en evaluering av seleksjonssystemets prediktive validitet*. Oslo: Forsvarets Høgskole.
- Laurence, J. H., & Matthews, M. D. (2012). *The Oxford Handbook of Military Psychology*. Oxford University Press.
- Le, H., Robins, S. B., Holland, E., Oh, I. S., Ilies, R., & Westrick, P. (2011). Too Much of a Good Thing: Curvilinear Relationships Between Personality Traits and Job Performance. *Journal of Applied Psychology*, 96(1), 113-133. doi: 10.1037/a0021016

- Mansikka, H., Harris, D., & Virtanen, K. (2017). An Input–Process–Output Model of Pilot Core Competencies. *Aviation Psychology and Applied Human Factors*, 7(2), 78-85. doi: 10.1027/2192-0923/a000120
- Martinussen, M., Richardsen, A. M., & Vårum, H. W. (2001). Validation of an ipsative personality measure (DISCUS). *Scandinavian Journal of Psychology*, 42(411-416). doi: 10.1111/1467-9450.00253
- Ones, D. S., Viswesvaran, C., & Schmidt, F. (1993). Comprehensive Meta-Analysis of Integrity Test Validities: Findings and Implications for Personnel Selection and Theories of Job Performance. *Journal of Applied Psychology*, 78(4), 679-703. doi: 10.1037/0021-9010.78.4.679
- Ployhart, R. E. (2012). Personnel Selection: Ensuring Sustainable Organizational Effectiveness Through the Acquisition of Human Capital. I S. W. J. Kozlowski (Red.), *The Oxford Handbook of Organizational Psychology* (s. 221-246). New York: Oxford University Press, Inc.
- Politidirektoratet. (2005). *Medisinske krav ved opptak til Politihøgskolen*. Politihøgskolen. Oslo.
- Politihøgskolen. (2014). Rammeplan for Bachelor-politiutdanningen. Hentet fra [http://www.phs.no/Documents/2\\_Studietilbud/1\\_Bachelor/Rammeplan\\_av\\_23.\\_januar\\_2014.pdf](http://www.phs.no/Documents/2_Studietilbud/1_Bachelor/Rammeplan_av_23._januar_2014.pdf).
- Politihøgskolen. (2016). Opptakstatistikk. fra <http://www.phs.no/om-phs/fakta-og-tal/opptakstatistikk/>
- Politi-loven. (1995). Lov om politiet. Hentet fra <https://lovdata.no/dokument/NL/lov/1995-08-04-53>.
- Rønnestad, M. H., & von der Lippe, A. (2009). *Det kliniske intervjuet* (2 utg.). Oslo: Gyldendal Akademisk.
- Schmidt, F., & Hunter, J. (1998). The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. *Psychological Bulletin*, 124(2), 262-274. doi: 10.1037/0033-2909.124.2.262
- Schmidt, F., Oh, I. S., & Le, H. (2006). Increasing the accuracy of corrections for range restriction: Implications for selection procedure validities and other research results. *Personnel Psychology*, 59(2), 281-305. doi: 10.1111/j.1744-6570.2006.00065.x

- Schmidt, F., Oh, I. S., & Shaffer, J. A. (2016). *The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 100 Years...* Working paper.
- Skorstad, E. (2015). *Rett person på rett plass. Psykologiske metoder i rekruttering og lederutvikling* (Vol. 2). Oslo: Gyldendal Akademisk.
- Svensson, K. (2013). *Flygarseleksjon i Forsvaret: ein studie av seleksjonssystemet sin prediktive validitet*. Forsvarets Høgskole.
- Vik, J. S. (2013). *Har seleksjon noen betydning? En studie av seleksjonens prediktive validitet*. Universitetet i Tromsø.
- Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J. M. R., & Schmidt, F. L. (2015). College Performance and Retention: A Meta-Analysis of the Predictive Validities of ACT® Scores, High School Grades, and SES. *Educational Assessment*, 20(1), 23-45. doi: 10.1080/10627197.2015.997614
- Aamodt, M. G. (2004a). *Law enforcement selection: Research summaries*. Washington D.C.: Police Executive Research Forum.
- Aamodt, M. G. (2004b). *Research in law enforcement selection*. Florida: Brown Walker Press.

# Vedlegg

Vedlegg 1: Kull 2013. Gjennomsnitt, standardavvik og interkorrelasjoner for alle variabler

Vedlegg 2: Kull 2014. Gjennomsnitt, standardavvik og interkorrelasjoner for alle variabler

Vedlegg 3: Kull 2013. Antall, gjennomsnitt, standardavvik og minimum- og maksimum-verdi for variablene i koblet datasett for prediktorer og kriterier.



**VEDLEGG 1: KULL 2013. Gjennomsnitt (M), standardavvik (SD) og interkorrelasjoner for alle variabler (N = 552-1128; høyest psykologiske tester, lavest på Karaktersnitt Totalt). Alle variabler er kodet på en 5-delt skala, med unntak av karakterpoeng VGS. Kriterier er kodet fra A=5 til E=1.**

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
<b>Prediktører</b>																								
1. Karaktersnitt VGS	40.72	4.90																						
2. Scales verbal admin	2.98	1.00	.08																					
3. Scales ix	3.00	1.01	.17	.20																				
4. Scales sx (a)	3.05	1.20	.20	.16	.16																			
5. Scales ll-no	2.99	1.04	.18	.43	.14	.15																		
6. Shapes Modenhet	2.98	.81	.02	.10	.04	.02	.10																	
7. Shapes Handlekraft	2.98	.89	.04	-.02	-.06	-.05	-.02	.12																
8. Shapes Samhandling	2.99	.86	-.10	-.21	-.12	-.08	-.16	-.28	-.24															
9. Gruppeøvelse Modenhet	3.27	.85	.06	.10	.02	.05	.13	.10	.01	-.11														
10. Intervju Modenhet	3.03	.73	.00	.14	.05	.04	.12	.63	.10	-.23	.25													
11. Gruppeøvelse Handlekraft	3.18	.91	.04	.11	.02	.05	.15	.11	.02	-.11	.78	.25												
12. Intervju Handlekraft	3.02	.75	.04	.04	-.04	-.02	.09	.16	.66	-.20	.21	.32	.20											
13. Gruppeøvelse Samhandling	3.27	.80	.05	.07	.01	-.01	.11	.06	.01	-.06	.77	.20	.73	.17										
14. Intervju Samhandling	3.57	.77	-.06	-.07	-.07	-.04	-.08	-.10	-.08	.54	.08	.07	.04	.07	.12									
15. Intervju Åpenhet og Inkludering	3.10	.80	.12	.07	.03	.08	.14	.07	.01	-.06	.27	.37	.22	.28	.23	.25								
<b>Kriterier</b>																								
16. Karaktersnitt Operative oppgaver	3.40	.49	.31	.06	.06	.12	.00	-.04	-.03	-.03	.14	.04	.10	-.02	.11	-.00	.16							
17. Karaktersnitt Melode	3.16	.95	.24	.09	.08	.12	.03	-.01	-.10	-.05	.06	.01	.05	-.01	.05	-.09	.06	.40						
18. Karaktersnitt Polit og Samfunn	3.36	.57	.35	.13	.06	.17	.08	-.02	-.05	.00	.12	.04	.07	-.02	.08	-.02	.11	.48	.40					
19. Karaktersnitt Forebyggende oppgaver	3.06	.76	.23	.09	-.00	.11	.07	.05	-.01	-.06	.06	.08	.02	.05	.05	-.05	.02	.25	.26	.36				
20. Karaktersnitt Etterforskning	3.43	.78	.32	.13	.07	.13	.05	-.04	-.06	.01	.09	.04	.06	-.04	.04	-.03	.09	.53	.42	.51	.27			
21. Karakter Bacheloroppgaven	3.39	.93	.14	.06	.05	.03	.05	.00	-.04	.01	.04	.07	.02	.01	.00	.04	.10	.22	.19	.22	.22	.21		
22. Karaktersnitt Totalt	3.40	.44	.42**	.12**	.10	.18	.06	-.05	-.06	-.04	.15**	.05	.11	-.02	.10	-.04	.18**	.82	.59**	.80**	.45**	.74**	.39**	

Note. \*p < .05, \*\*p < .01 (to-talet)

**VEDLEGG 2: KULL 2014. Gjennomsnitt (M), standardavvik (SD) og interkorrelasjoner for alle variabler (N = 401-1315; lavest på Karaktersnitt Totalt, høyest på psykologiske tester). Alle variabler er kodet på en 5-delt skala, med unntak av karakterpoeng. Kriterier er kodet fra A=5 til E=1.**

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
<b>Prediktorer</b>																						
1. Karakterpoeng VGS	41.32	4.54																				
2. Scales verbal admin	3.02	.99	.07																			
3. Scales lst	2.74	.98	.09	.16																		
4. Scales lt-no	3.09	1.04	.12	.46	.11																	
5. Shapes Modenhhet	3.04	.73	-0.00	.06	-0.01	.04																
6. Shapes Handlekraft	3.05	.81	.04	.01	-0.02	-0.01	.10															
7. Shapes Samhandling	2.95	.77	-0.04	-.15	-.06	-.13	-.29	-.23														
8. Squares	4.01	.90	-0.00	-.11	-.09	-0.10	.14	.21	.16													
9. Intervju Åpenhet og Inkludering	3.10	.83	.06	.09	.04	.11	.07	-0.01	-0.03	.01												
10. Gruppeøvelse Modenhhet	3.27	.88	-0.01	.10	.02	.08	.05	-0.03	-0.05	-0.01	.29											
11. Intervju Modenhhet	3.10	.71	.05	.08	.02	.09	.64	.08	-.20	.08	.41	.20										
12. Gruppeøvelse Handlekraft	3.25	.89	.02	.09	.01	.11	.05	-0.01	-0.07	-0.00	.27	.79	.19									
13. Intervju Handlekraft	3.13	.76	.02	.01	-0.02	.03	.14	.65	-.17	.18	.30	.11	.33	.12								
14. Gruppeøvelse Samhandling	3.24	.79	.02	.06	.01	.09	.06	-0.02	-0.03	.02	.23	.76	.18	.70	.10							
15. Intervju Samhandling	3.00	.71	-0.01	-.07	-0.01	-0.04	-0.12	-.09	.64	.12	.29	.12	.13	.10	.12	.12						
<b>Kriterier</b>																						
16. Karaktersnitt Orden og Beredskap	3.54	.51	.34	.12	.06	.06	.01	.02	.01	.04	.13	.06	.03	.08	.01	.11	-0.01					
17. Karaktersnitt Politi og Samfunn	3.40	.71	.32	.09	.09	.19	.05	.00	-0.08	.02	.12	.01	.10	.04	.01	.06	.00	.33				
18. Karaktersnitt Forebyggende oppg.	3.17	.82	.37	-0.01	.05	.09	-0.06	-0.00	.05	.03	.15	.04	-0.01	.07	.00	.05	.07	.48	.36			
19. Karaktersnitt Eterforskning	3.44	.60	.35	.09	.12	.10	.01	-0.02	-0.07	.01	.14	.09	.03	.13	-.00	.11	-.06	.64	.37	.50		
20. Karakter Bacheloroppgave/Metode	3.39	.94	.14	-.04	-0.00	.01	-0.04	-0.05	.00	-0.04	.02	-0.01	-0.05	-0.01	-0.02	.02	-0.00	.31	.19	.27	.32	
21. Karaktersnitt Totalt	3.44	.48	.41	.05	.07	.12	-.02	.03	-.06	.07	.14	-.01	-0.01	.03	-0.01	.07	-0.04	.73	.61	.74	.73	.66

Note. \*p < .05; \*\*p < .01 (to-halet)

**VEDLEGG 3: Kull 2013**

**Antall (N), gjennomsnitt (M), standardavvik (SD) og minimum- og maksimum-verdi for variablene i koblet datasett for prediktorer og kriterier.**

**PREDIKTORER: Karakterpoeng VGS og Psykologiske tester**

		Karakterpoeng	Scales verbal admin	Scales ix	Scales sx (a)	Scales It-no	Shapes Modenhet	Shapes Handlekraft	Shapes Samhandling
N	Valid	715	638	638	638	638	638	638	638
	Missing	5	82	82	82	82	82	82	82
M		40,98	3,12	3,11	3,24	3,15	3,01	2,96	2,95
SD		4,77	1,00	0,97	1,14	1,05	0,77	0,90	0,84
Min.		25,70	1	1	1	1	1	1	1
Max.		55,40	5	5	5	5	5	5	5

**PREDIKTORER: Intervju og Gruppeøvelse**

		Gr.øvelse Samhandling	Intervju Samhandling	Intervju Åpenhet og Inkl.	Gr.øvelse Modenhet	Intervju Modenhet	Gr.øvelse Handlekraft	Intervju Handlekraft
N	Valid	628	628	627	628	627	628	626
	Missing	92	92	93	92	93	92	94
M		3,42	3,65	3,30	3,43	3,18	3,36	3,11
SD		0,74	0,72	0,71	0,76	0,64	0,84	0,72
Min.		2	1	2	1	1	1	1
Max.		5	5	5	5	5	5	5

**KRITERIER: Karakterer oppnådd på B1 og B3 (E=1, A=5)**

		AT330	ETF350	FOP310	FORB 340	FYS331	KKH330	KRI150	KRIM 310
N	Valid	692	690	690	694	664	693	714	690
	Missing	28	30	30	26	56	27	6	30
M		3,28	3,34	3,46	3,06	2,83	3,68	3,51	3,27
SD		0,88	1,02	0,86	0,76	0,99	0,95	0,94	0,97
Min.		1	1	1	1	1	1	1	1
Max.		5	5	5	5	5	5	5	5

**KRITERIER: Karakterer oppnådd på B1 og B3 (E=1, A=5)**

		OPPG 300	OPR 30	ORD 130	POL 110	PSY 130	SOS 110	STR 150	VEG 130	YRK 310
N	Valid	683	693	704	717	711	718	704	703	692
	Missing	37	27	16	3	9	2	16	17	28
M		3,39	3,79	4,00	3,10	3,28	3,64	3,16	3,15	3,31
SD		0,93	0,89	0,95	0,96	0,92	1,01	0,95	0,97	0,84
Min.		1	1	1	1	1	1	1	1	1
Max.		5	5	5	5	5	5	5	5	5

**VEDLEGG 4: Kull 2014**

**Antall (N), gjennomsnitt (M), standardavvik (SD) og minimum- og maksimum-verdi for variablene i koblet datasett for prediktorer og kriterier.**

**PREDIKTORER: Karakterpoeng VGS og Psykologiske tester**

		Karakterpoeng	Scales verbal admin	Scales Ist	Scales It-no	Shapes Modenhet	Shapes Handlekraft	Shapes Samhandling	Squares Pålidelighet
N	Valid	660	668	668	668	668	668	668	668
	Missing	11	3	3	3	3	3	3	3
M		41,35	3,16	2,87	3,26	3,05	3,06	2,90	3,98
SD		4,62	0,96	1,00	1,02	0,70	0,79	0,75	0,89
Min.		30,97	1	1	1	1	1	1	2
Max.		55,40	5	5	5	5	5	5	5

**PREDIKTORER: Intervju og Gruppeøvelse**

		Intervju Åpenhet og Inkl.	Gr.øvelse Modenhet	Intervju Modenhet	Gr.øvelse Handlekraft	Intervju Handlekraft	Gr.øvelse Samhandling	Intervju Samhandling
N	Valid	655	655	655	654	655	655	655
	Missing	16	16	16	17	16	16	16
M		3,30	3,48	3,21	3,44	3,24	3,41	3,07
SD		0,75	0,81	0,65	0,83	0,74	0,73	0,69
Min.		2	1	1	1	1	1	1
Max.		5	5	5	5	5	5	5

**KRITERIER: Karakterer oppnådd på B1 og B3 (E=1, A=5)**

		AT30	ETTERF30	KORD30	KRIF10	KRIF30	KRIM30 (50 % av kullet)	KRIT10	OPR30
N	Valid	649	638	619	657	654	270	662	627
	Missing	22	33	52	14	17	401	9	44
M		3,33	3,44	3,46	3,27	3,06	3,36	3,71	3,92
SD		0,83	1,06	0,83	1,02	1,04	0,83	0,87	0,77
Min.		1	1	1	1	1	1	1	1
Max.		5	5	5	5	5	5	5	5

**KRITERIER: Karakterer oppnådd på B1 og B3 (E=1, A=5)**

		ORFK10	PSYK10	REL10	SOPF10	STRA10	VEGT10	YRKE30 (50 % av kullet)
N	Valid	627	664	657	661	659	655	252
	Missing	44	7	14	10	12	16	419
M		3,69	3,29	3,36	3,38	3,17	3,32	3,50
SD		0,89	0,92	0,84	0,98	0,92	0,97	0,76
Min.		1	1	1	1	1	1	1
Max.		5	5	5	5	5	5	5





**Politihøgskolen**  
Slemdalsveien 5  
Postboks 5027, Majorstuen  
0301 Oslo  
Tlf: 23 19 99 00  
[www.phs.no](http://www.phs.no)

