



COUNCIL OF EUROPE    CONSEIL DE L'EUROPE

Language Policy Division  
Division des Politiques Linguistiques

## **Språkprøver i forbindelse med sosial integrering og statsborgerskap – en veiledning til de ansvarlige myndigheter**

**(Oversettelse: Rønnaug Katharina Totland)**

**Forfattere:**

**Andrew Balch**

**Shalini Roppe**

**Michael Corrigan**

**Heinrich Rübeling**

**Sara Gysen**

**Steffi Steiner**

**Henk Kuijper**

**Piet Van Avermaet**

**Michaela Perlmann-Balme**

**Beate Zeidler**

**Med mye hjelp fra ALTEs interessegruppe for vurdering av språk for migrasjon og integrering (the LAMI group).**

This study was published by the Council of Europe as part of a collection of "Thematic Studies on the linguistic integration of adult migrants" prepared in view of a seminar organised in Strasbourg in 2008 ([www.coe.int/lang](http://www.coe.int/lang)).

## 1. Innledning

Mange europeiske land innfører for tiden krav til språkferdigheter hos innvandrere som ønsker statsborgerskap. Regjeringer i forskjellige land krever ofte i denne forbindelse at det skal brukes språkprøver eller andre formelle prosedyrer for å måle kandidatenes ferdigheter. Formålet med dette kapitlet er ikke å argumentere for innføring av slike språkprøver, men å gi erfaringsbaserte faglige råd til de som allerede har innført eller vurderer å innføre slike prøver slik at prøvetakernes behov ivaretas, og at de får en riktig og rettferdig bedømming. De ansvarlige myndighetene oppfordres sterkt til å ta de følgende problemstillinger i betraktning dersom de vurderer å innføre denne typen språkprøver:

- Kan det være mer hensiktsmessig å velge en annen vurderingsform enn en prøve?
- Hvis ikke, hvilken type prøve ville være mest formålstjenlig?
- Hvordan skal prøveresultatene brukes?
- Hvilke konsekvenser vil en slik prøve få for samfunnet?
- Hvilke følger vil en slik prøve få for den enkelte innvandrere?

Når man vurderer disse spørsmålene, må man være oppmerksom på at resultatet på en prøve alltid vil inneholde en viss feilmargin og mulig bias<sup>1</sup>. Testutviklere har ansvar for at feilmarginen og bias er innenfor en akseptabel grense i henhold til internasjonale standarder (se avsnitt 3). Men selv om prøven fungerer utmerket for en stor gruppe kandidater, kan den ikke ta hensyn til individuelle forskjeller som personlighetstrekk, læringshistorie og tidligere personlige erfaringer. Ut fra den enkelte prøvedeltakers perspektiv, reflekterer ikke alltid skåren deres virkelige ferdigheter, og det er ikke mulig å utelukke fullstendig at resultatet kan være påvirket av bias. Følgelig må det ønskede utbyttet av en slik prøve ses i forhold til konsekvensene det får å stryke på prøven; avgjørelser som tas på grunnlag av prøveresultatet kan være feilaktige.

Dersom man etter å ha vurdert disse punktene, bestemmer seg for å bruke/innføre en språkprøve, må de ansvarlige myndigheter og andre involverte, sørge for at prøven fungerer slik den er ment, og at den brukes på en hensiktsmessig og rettferdig måte. Rettferdig måling/bedømming er viktig for alle typer prøver og alle målgrupper, men er spesielt viktig når det gjelder språkprøver til bruk ved innvandring, oppholdstillatelse og statsborgerskap på grunn av de vidtrekkende konsekvensene disse prøvene får for prøvetakerne med tanke på borger- og menneskerettigheter. Dette kapitlet gir informasjon til de som er ansvarlige for å tilby en slik prøve når det gjelder å velge hvem som skal utvikle prøvene, tilrettelegge arbeidet for disse testutviklerne og sikre tilstrekkelige ressurser i planleggingsfasen. Bruk av god prøvepraksis skal ikke bare garantere at relevante ferdigheter og kunnskaper blir testet (testvaliditet<sup>2</sup>), men skal også sikre stabilitet i behandlingen av kandidatene og at de ulike versjonene av prøven er ekvivalente (testreliabilitet<sup>3</sup>).

---

<sup>1</sup> Det forhold at en test eller en oppgave har egenskaper i favør eller ufavør av en bestemt del av kandidatgruppen samtidig som disse egenskapene ikke har relevans for det som skal måles. Kilder til bias kan være kjønn, alder, kultur osv.

<sup>2</sup> Det vil si at prøven måler det den pretenderer å måle. Dersom en prøve skal måle leseforståelse, må den, for å gi valide resultater, måle denne typen språkferdighet isolert fra andre ferdigheter som lytteforståelse eller skriftlig produksjon. En prøve i leseforståelse der instruksjonene blir gitt muntlig og kandidaten skal skrive et langt svar

## 2. Å avgjøre hva som skal testes.

### 2.1 Oversikt

I denne delen blir de ulike tiltakene som skal til for å sikre at prøven blir formålstjenlig, gjennomgått. Det første tiltak i denne prosessen er å formulere nøyaktig og entydig formålet med prøven. Når dette er gjort, må innhold og vanskelighetsgrad fastsettes. Til slutt må en prøvebeskrivelse (testspesifikasjoner) utarbeides, et uunnværlig dokument for videre arbeid med å utvikle og revidere prøven.

### 2.2 Å fastsette formålet med prøven og samfunnsmessige krav til prøvetakerne

Før man utvikler en språkprøve, er det nødvendig å definere den nøyaktige hensikten med prøven. Det holder ikke å si at prøven skal brukes i forbindelse med oppholdstillatelse og statsborgerskap. Selv innen dette området fins det en hel rekke grunner for å teste innvandrere; alt fra å motivere innlærere (hjelp dem til å bruke og forbedre sin kompetanse i målspåket), finne ut om de har gode nok ferdigheter til å delta i en definert sosial situasjon (for eksempel studier eller arbeid) til å ta avgjørelser som har følger for deres juridiske rettigheter, for eksempel retten til å bli i landet og få innvilget statsborgerskap.

Først når prøvens formål er klart definert, er det mulig å identifisere kravene som deltakerne vil møte i det virkelige liv og som må gjenspeiles i prøven. For eksempel å kunne ta aktivt del i samfunnet og utøve statsborgerlige retter og plikter. Et klart og tydelig definert bruksformål for prøven er viktig av flere grunner: for det første er det viktig i utviklingen av prøven, for det andre bidrar det til å sikre en rettferdig prøve. Dessuten vil prøvetakerne vite hva som venter dem, og det hjelper andre samfunnsaktører til å tolke og bruke prøveresultatene på en riktig måte. Denne prosessen med å identifisere behovene for prøven kalles en behovsanalyse.

Når man gjennomfører en slik behovsanalyse, er det nødvendig å ta hensyn til at ulike grupper innvandrere kan ha ulike behov. De som for eksempel vil komme seg inn på arbeidsmarkedet raskest mulig, har ofte andre behov enn de som har tenkt å være hjemme med små barn. Når språktester skal gjennomføre en behovsanalyse, bør de først definere de relevante sammenhenger og situasjoner målgruppen befinner seg i. Det er også viktig at myndighetene setter av nok ressurser, og at representanter for ulike deler av samfunnet involveres.

---

som uttrykk for hva han har forstått, vil ikke lenger være en valid prøve på kandidatens leseforståelse. Validitet henger derfor nøye sammen med hva det er man ønsker å måle med testen. I testteorien blir det understreket at en prøve ikke kan være reliabel eller valid *per se*, men at dette må ses i sammenheng med formålet med testen. Det er årsaken til at man ikke sier at en test er valid, men at den gir valide resultater eller skårer. (*Den objektive språktesten: mål eller minne? Cecilie Toledo Carlsen i Brodersen, R. & Kinn, T. 2000 Språkvitenskap og vitenskapsteori: ti nye vitenskapsteoretiske innlegg. Ariadne*)

<sup>3</sup> Reliabilitet har med måleinstrumentets grad av stabilitet å gjøre. Høy grad av reliabilitet betyr at testen er lite påvirket av målingsfeil. Det vil si at man skal kunne stole på at kandidatene ville ha fått samme resultat på prøven ved en annen testavvikling og med andre sensorer. (*Den objektive språktesten: mål eller minne? Cecilie Toledo Carlsen i Brodersen, R. & Kinn, T. 2000 Språkvitenskap og vitenskapsteori: ti nye vitenskapsteoretiske innlegg. Ariadne*)

### **2.3 Å fastsette de språklige kravene**

Når det er avklart hvilke krav samfunnet stiller til innvandrere, må de omsettes til språklige krav, som ikke bare gjelder kunnskaper og ferdigheter, men også hvilket språklig nivå prøvetakerne har behov for. Hvis man for eksempel skal utvikle en prøve som skal måle om prøvetakeren har gode nok språkferdigheter til å klare en yrkesutdanning, er det naturlig at prøven måler om prøvetakeren er i stand til å følge med på undervisningen, kommunisere med lærere og medelever, kan lese relevant faglitteratur og skrive oppgaver/innleveringer etc.

Denne analysen kan også være til hjelp når man skal avgjøre hvilket språklig nivå det er rimelig å kreve i for eksempel lesing eller skriving. Hvis prøven derimot var laget for å sertifisere tolker og oversettere, ville en tilsvarende behovsanalyse vise at det var nødvendig med et langt høyere ferdighetsnivå. En oversetter trenger høyest kompetanse innen lesing og skriving, mens for en tolk vil de muntlige ferdighetene være de viktigste.

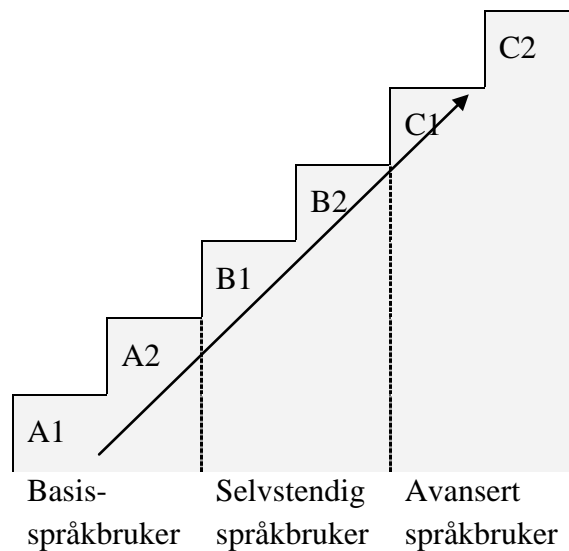
I motsetning til eksemplene over, er det mye vanskeligere å avdekke hva som er nødvendige språkferdigheter for innvandrere eller de som søker statsborgerskap. Forholdet mellom en innlærers ferdighetsnivå i det offisielle språket (/språkene) og hans/hennes evne til å integrere seg i samfunnet og beherske retter og plikter som et statsborgerskap fører med seg, er løsere og derfor mye vanskeligere å kartlegge. Dersom språkferdigheter hadde vært den eneste faktoren som spilte inn, ville alle morsmålsbrukere ha vært like godt integrert i samfunnet. Siden dette ikke er tilfelle, kan vi gå ut fra at også andre faktorer spiller en rolle. Ikke desto mindre er oppgaven til de som utvikler språkprøver, å identifisere de nødvendige språklige behov. Etter en grundig behovsanalyse er det også viktig å forsikre seg om at ingen feiloppfatninger av kandidatens kultur- eller utdanningsbakgrunn påvirker prøven: Mens språkprøver for studie- og arbeidsformål i de fleste tilfeller blir tatt av en homogen kandidatgruppe når det gjelder utdanningsbakgrunn og kognitive evner, må prøver for integrasjon og statsborgerskap passe for en rekke ulike kandidater, både de med lite skolebakgrunn og de med lang akademisk utdanning.

### **2.4 Å fastsette språklig vanskelighetsgrad**

Når de språklige behovene er kartlagt, må testutviklerne forsøke å knytte dem til ”Jeg-kan-utsagn” (Can-do-statements) i Common European Framework for Reference, (Det felles europeiske rammeverket for språk) fra 2001. Dette rammeverket

gir et felles grunnlag for utforming av lærestoff til språkkurs, retningslinjer for læreplaner, eksamener og læremidler i hele Europa. Rammeverket er en omfattende beskrivelse av hvilke kunnskaper og ferdigheter man må utvikle for å kunne kommunisere effektivt på et språk. (Fra den norske oversettelsen av Det felles europeiske rammeverket for språk)

Rammeverket inneholder en rekke illustrerende skalaer (for muntlig, skriftlig, lesing, lytting og interaksjon), som definerer ferdighetsnivåene og beskriver hva en innlærer kan gjøre på hvert nivå, og som gjør det mulig å måle framskrittene på en sekstrinns nivåskala.



Man må merke seg at selv om avstanden mellom hvert språklige nivå er lik, betyr det ikke at en kandidat trenger like lang tid for å gå fra et nivå til et annet. Fordi bredden av språklige ferdigheter som kreves for hvert nivå, gradvis blir større, tar det lenger tid å gå fra ett nivå til det neste jo høyere opp på skalaen man kommer. Behovsanalysen for en prøve bør derfor ikke fokusere på hvor mange undervisningstimer som trengs for å klare prøven, men derimot på en detaljert beskrivelse av prøvetakerens språkbehov ved hjelp av ”can-do”-utsagn.

Testutviklere må dessuten ha klart for seg at de fleste innlærere ikke har jevn kompetanse i lese-, lytte-, skrive- og snakkeferdigheter. Som regel er kompetansen i muntlig kommunikasjon bedre utviklet enn kompetansen i skriving og lesing. Denne observasjonen sammen med omfattende behovsanalyser, taler for et profilert kompetansebegrep. Prøvebeviset bør derfor ikke bare presentere et samlet ferdighetsnivå, men oppgi ett nivå for hver enkeltferdighet.

Den potensielt ujevne ferdighetsprofilen bør også gjenspeiles i en modulær tilnærming i testing og vurdering. For enkelte innlærere kan det være mer motiverende å ta en prøve i bare én eller to ferdigheter som de behersker godt enn i alle fire ferdighetene på en gang. I praksis betyr dette for eksempel at den muntlige delen av prøven kan tas separat og gi et eget prøvebevis. Denne muligheten er spesielt godt egnet for innvandrere som ikke kan lese eller skrive, eller for deltakere med svært lite skriveerfaring. En logisk konsekvens av denne tilnæringsmåten, er at språkundervisningen kan konsentrere seg om de ferdighetene som trenger ekstra oppmerksomhet. Når testinstitusjonene gir resultat av slike delprøver videre, er det viktig å understreke at resultatet ikke må forveksles med resultatet på prøver som tester alle fire ferdigheter.

## 2.5 Å lage testspesifikasjoner

Når prøvens målgruppe, formål og fokus er bestemt, må dette formuleres i en detaljert testspesifikasjon. Testspesifikasjonen beskriver også prøveformatet, hvilke oppgavetyper som

skal brukes og instruksjoner for gjennomføring av prøven. Disse spesifikasjonene fungerer deretter som referansedokument for senere avgjørelser i forbindelse med prøven.

### **3. Å sikre at prøvespesifikasjonene blir fulgt**

#### **3.1 Oversikt**

Når testspesifikasjonene er ferdigstilt, gjenstår ennå mye arbeid for å sikre at prøven fungerer slik som den skal. Vurderingskriterier og prøveformat må utvikles, og oppgaver ("spørsmål") må utarbeides i tråd med testspesifikasjonene. Man må kunne vise at oppgavene stemmer overens med beskrivelsene i testspesifikasjonene, og oppgavene må settes sammen til en prøve på en måte som er i samsvar med spesifikasjonene. Prøven må i neste omgang gjennomføres på en konsekvent og rettferdig måte for de prøvetakerne den er tiltenkt. Til slutt må resultatene fra prøven analyseres for at man skal kunne bekrefte at prøven fungerer som forventet. Dersom dette ikke er tilfelle, må man justere oppgavene det gjelder.

#### **3.2 Vurderingskriterier og prøveformat**

Prøvefokuset som er skissert i testspesifikasjonene (2.5) må deles opp i enkeltstående enheter. Først når det er dannet et nøyaktig bilde av hvordan kandidatene skal testes og vurderes, kan utviklingen av en passende kombinasjon av oppgaver og oppgavetyper begynne. Testutviklerne bør først og fremst sørge for at kandidatene får oppgaver som gir dem anledning til å vise at de oppfyller kravene i vurderingskriteriene.

#### **3.3 Oppgaveskriving**

For å kunne lage egnede oppgaver, må oppgaveutviklerne få klare retningslinjer. Disse gir som regel en oversikt over prøvens målgruppe og hensikt, sammen med generelle råd; egnede og uegnede tema, tekstmengde (hvor mange ord en lesetekst skal inneholde (Input) og hvor mange ord kandidaten skal svare (Output)) og i hvilken grad tekstene skal være autentiske. Etter dette må hver enkelt oppgavetype vurderes for seg. Når oppgavene er ferdiglaget, bør andre eksperter vurdere om retningslinjene er fulgt.

#### **3.4 Utprøving**

De eksterne ekspertene kan vurdere om oppgavene synes egnet for prøvens formål. For å sikre at oppgavene faktisk fungerer etter intensjonen, er det nødvendig å prøve ut materialet under ekte testbetingelser på kandidater som så godt som mulig gjenspeiler den virkelige testpopulasjonen når det gjelder demografisk profil. Gjennom en utprøving forsikrer man seg om at oppgavene fungerer som planlagt – at de tester målspråkskompetansen, skiller sterke fra svake kandidater, ikke diskriminerer enkelte deltakergrupper osv. Utprøvningsresultatet av lukkede oppgaver (for eksempel flervalgsoppgaver) kan bli analysert i en detaljert statistisk analyse, de åpne oppgavene (for eksempel muntlige oppgaver) blir oftest analysert kvalitativt for å se hvordan kandidatens produksjon svarer til testutviklernes forventninger. På grunnlag av slike analyser vil oppgavene enten bli brukt i reelle prøver, omarbeidet og utprøvd på nytt eller forkastet. Dersom det også fins en oppgavebank der oppgavene og analyseresultatene lagres, kan fremtidige prøver settes sammen av oppgaver fra denne banken. Man kan velge oppgaver som innfrir bestemte krav, for eksempel til vanskelighetsgraden.

### **3.5 Prøveavvikling**

De som tilbyr prøven må forsikre seg om at alle prøvedeltakere får ta prøven under like betingelser. Av den grunn anbefales det å gi tydelige og detaljerte instruksjoner for gjennomføring av prøven. Disse instruksene bør inneholde de følgende punktene:

- Alle prøvesteder må være autorisert for å administrere prøven
- Prøvelederne og andre involverte i avviklingen har den nødvendige kompetansen, blir godt informert og tilstrekkelig fulgt opp
- Prøvestedene må garantere god sikkerhet og konfidensiell behandling gjennom hele prøveprosessen; fra kandidatene melder seg på til de får resultatet
- De fysiske betingelsene i prøvelokalene er gode både når det gjelder støynivå, temperatur, avstanden mellom kandidatene osv.
- Det må legges til rette for kandidater med spesielle behov

### **3.6 Tilrettelegging for kandidater med spesielle behov**

Prøvesystemet må ikke diskriminere prøvedeltakere med spesielle behov, som for eksempel forbigående eller varige fysiske, psykiske eller emosjonelle problemer/handicap, lærevansker, forbigående eller kroniske sykdommer. Det kan også dreie seg om analfabetisme på morsmålet eller innlærerspråket, religiøse forskrifter, fengselsstraff eller andre begrensninger som gjør det vanskelig eller umulig for en prøvedeltaker å gjennomføre prøven på samme måte som andre.

Det må finnes faste ordninger for

- fritak for hele eller for deler av prøven
- hvilke forholdsregler som må tas for å sikre en rettferdig bedømmelse av personer med behov for spesiell tilrettelegging
- hvilke institusjoner som har myndighet til å gi fritak for deler av eller hele prøven
- hvilken type tilrettelegging som tilbys i særtilfeller (for eksempel prøvemateriell i blindeskrift, stort format, datamaskiner med blindeskrift eller spesialutstyrte datamaskiner til disposisjon, assistent som kan lese høyt eller skrive, ekstra tid til bestemte delprøver, ekstrapauser, tegnspråktolk, annet prøvetidspunkt eller prøvested) for å sikre at kandidaten kan klage på avslag om tilrettelegging, informere om hvordan kandidatene kan klage og forklare hvordan den endelige avgjørelsen tas

All informasjon om disse punktene bør være offentliggjort og lett tilgjengelig for alle prøvetakere.

Anbefalt litteratur:

"Special Educational Needs in Europe. The Teaching & Learning of Languages. Teaching Languages to Learners with Special Needs", European Commission, DG EAC 23 03 LOT 3, January 2005.

### **3.7 Vurdering og karaktersetting**

Oppgaver i en språkprøve kan deles inn i to kategorier etter hvordan de rettes: Objektivt rettede oppgaver har ett rett svar og kan derfor rettes maskinelt eller etter fasit. Subjektivt

vurderte oppgaver (åpne oppgaver) må derimot rettes av øvede sensorer. I denne prosessen kan det oppstå både skrivefeil/overføringsfeil og vurderingsfeil. Vurderingsfeil viser seg i inkonsistens i vurderingen av kandidatprestasjonene og/eller (feil bruk av) prøvens vurderingskriterier. Sensorer kan være inkonsistente enten når man sammenligner dem med andre sensorer eller når man sammenligner deres egne vurderinger (hvordan de bruker kriteriene/rettemalen) over et visst tidsrom. Slik inkonsistens kan bli merkbart redusert ved grundig sensoropplæring og ved kontroll av sensorvurderingene. Bruk av flere sensorer til å vurdere noen eller alle kandidatens delprøver kan også redusere inkonsistensen. Vurderingene kan også justeres for å utjevne virkningen av konsekvent milde og konsekvent strenge sensorer. I tilfeller med store avvik, må sensorene kurses på nytt.

### **3.8 Kontroll med prøvens funksjon**

På samme måte som man må kontrollere sensoradferd (se kapittel 3.7), må man samle inn og analysere både prøvesvar og persondata (for eksempel alder, kjønn og nasjonalitet). Dette er nødvendig for å sikre at:

- prøven måler de ferdighetene den er ment å måle
- disse ferdighetene blir prøvd på samme måte i alle versjoner av prøven
- hver prøveversjon gir rettferdige betingelser til alle kandidater den er tiltenkt uavhengig av bakgrunn

Resultatene fra en slik kontroll kan så brukes for å sikre at prøveresultatet gir et riktig bilde av kandidatens ferdigheter. Dessuten bør utfallet av analysen få følger for videre prøveutvikling, gjennomføring og vurdering, slik at disse prosessene fortløpende kan bli forbedret.

#### **3.8.1 Analyse av prøvesvarene**

Svar på oppgaver fra reelle prøveavviklinger analyseres for å se både hvordan råskåren (testskåre som ikke er transformert gjennom for eksempel vekting eller reskalering) på prøven er, og for å undersøke hvor godt de enkelte oppgavene måler kandidatens ferdigheter (se 3.4). Man må analysere både oppgavens vanskelighetsgrad og i hvor stor grad oppgaven skiller mellom sterke og svake kandidater. Disse og andre statistiske beregninger må lagres slik at man kan sammenlikne de ulike prøveversjonene. Dette skal sikre at resultatene fra forskjellige versjoner av prøven er ekvivalente. Når en prøve baserer seg på utprøvde oppgaver, kan man gå ut fra at spørsmål og oppgaver langt på vei fungerer som forventet på den reelle prøven. Likevel er det viktig å analysere prøveresultatene for å sammenlikne disse med utprøvningsanalysene. Dersom det er stort avvik mellom disse to analysene, kan det være nødvendig å justere oppgavene. Det er selvsagt også viktig at hver versjon av prøven måler kandidatens ferdigheter på samme måte som tidligere og kommende versjoner. Av denne grunn blir ofte resultatene fra de reelle prøvene brukt for å bestemme grensen mellom ulike karakter eller mellom bestått og ikke bestått. Dersom en kandidat hadde tatt to ulike versjoner av en prøve, er det usannsynlig at han/hun hadde fått akkurat like mange poeng. Derimot kan man sikre at kandidaten vil få samme karakter på to prøveversjoner, eller ville havne på samme side at bestått-/ikke beståttgrensen. Dette blir sikret gjennom ekvivalering av prøvene; å avgjøre hvilken skåre på prøve B som tilsvarer skåren som kreves



for å bestå prøve A. Dersom oppgavene har vært prøvd ut, kan en lettere og mer presist avgjøre dette.

### **3.8.2 Analyse av bias**

Liknende prosedyrer som ble beskrevet under 3.8.1 bør også brukes på resultatene fra kandidater som er på samme ferdighetsnivå. Hvis for eksempel kandidater fra et spesielt land gjør det påfallende bedre eller dårligere på én eller flere oppgaver enn andre kandidater, kan dette ha ikke-språklige årsaker som er til ulempe eller fordel for bestemte grupper. Dette vil falle urettferdig ut for enkelte deltakergrupper. En annen årsak kan imidlertid være språklige likheter/ulikheter mellom morsmål og målspråk, noe som ikke kan sies å være urettferdig. I begge tilfeller er det nødvendig med en kvalitativ undersøkelse. Dersom det viser seg at en oppgave er systematisk urettferdig på gruppenivå, må den vurderes fjernet og prosedyrene for oppgavelaging må forandres slik at man unngår dette for fremtidige oppgaver.

## **4. Konklusjon**

Som vi har beskrevet i andre deler av dette dokumentet, kan et prøveresultat få viktige følger både for kandidaten selv og for større grupper mennesker, som samfunnet i sin helhet. Man må derfor vurdere nøye om en prøve er nødvendig i en bestemt sammenheng og hva prøven bør inneholde. Dersom det er avgjort at en språkprøve skal brukes i forbindelse med oppholdstillatelse/statsborgerskap, er det viktig at prøven tilfredsstillende kravene som er skissert i dette dokumentet. Prøven må kontrolleres kontinuerlig for å sikre at funksjons- og kvalitetskravene er innfridd. For å sikre at en språkprøve fungerer godt som et politisk virkemiddel, må de ansvarlige myndigheter arbeide sammen med testutviklere. Dette innebærer en presis definisjon av formålet med prøven og et overslag av hvilke ressurser arbeidet vil kreve. Hensynet til at alle kandidatene skal få en riktig og rettferdig behandling må til enhver tid komme i første rekke.