



Utdanningsdirektoratet

**Skole og velferd: Hvordan vet vi  
hva som gir fremgang?  
Evalueringskonferansen 2014**

An abstract graphic design featuring a light gray background. In the lower right corner, there is a complex pattern of overlapping orange circles and lines. Some circles are solid orange, while others are hollow. Lines of varying thicknesses connect some of the circles, creating a network-like structure. The overall aesthetic is modern and minimalist.

**Utvikling i norsk skole – hvordan måles den?**

## Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels



**NYHETER** - 1. halvår 2013  
Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels

**NYHETER** - 1. halvår 2013  
Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels

**NYHETER** - 1. halvår 2013  
Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels

**NYHETER** - 1. halvår 2013  
Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels

## Blir bare dårligere i matt



**NYHETER** - 1. halvår 2013  
Blir bare dårligere i matt

**DEBATT**

**NETT**

**Et realfagsproblem**

## Et realfagsproblem

**R**ealfagene har blitt...  
Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels

## Seks av ti positive til rushtidsavgift

**NYHETER** - 1. halvår 2013  
Seks av ti positive til rushtidsavgift

## Advarer mot testhysteri

**NYHETER** - 1. halvår 2013  
Advarer mot testhysteri

## Advarer mot testhysteri

**NYHETER** - 1. halvår 2013  
Advarer mot testhysteri

## Advarer mot testhysteri

**NYHETER** - 1. halvår 2013  
Advarer mot testhysteri

## Advarer mot testhysteri

**NYHETER** - 1. halvår 2013  
Advarer mot testhysteri

## Advarer mot testhysteri

**NYHETER** - 1. halvår 2013  
Advarer mot testhysteri

**Finsk refs.** Nordmenn snubler i rettskrivningsreglene: Både det verd å finne ut om opplæring i bare én skriftlig målform gir noen bedring.

## Et folk som ikke kan skrive?



**Kommentar**  
Per Anders Madsen  
Innledning

Norge er en nasjon av dyktigheiter og f...  
Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels

Heller ikke i går ble lærerstreiken løst. Det vil ta lang tid før tilliten og samarbeidsviljen blir gjenopprettet.

## Mistilliten er gjensidig

**INNSIKTS**  
CUNNAR OGGE

I går blev lærerstreiken løst...  
Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels

**Fakta**  
Lærerstreiken

Etter streiken...  
Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels

**Fakta**  
Lærerstreiken

Etter streiken...  
Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels

**Fakta**  
Lærerstreiken

Etter streiken...  
Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels

**Fakta**  
Lærerstreiken

Etter streiken...  
Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels

**Fakta**  
Lærerstreiken

Etter streiken...  
Etter ti år, 66 prosent dyrere skole og 43 tiltak er PISA-resultatene fortsatt middels

# Noen kritiske påstander og kommentarer

---

- PISA og øvrige undersøkelser måler bare enkelte ferdigheter – hvor relevant er det?
- Rangeringene tar oppmerksomheten bort fra viktige pedagogiske spørsmål?
- Resultatene fra undersøkelsene er tilfeldige og ikke pålitelige?
- Norsk utdanningspolitikk skal ikke styres av OECD?

# Innhold

---

- Internasjonale studier; formål, innhold, betydning
- Tilstand og utvikling i norsk skole
- Samlet kunnskapsgrunnlag
- Eksempler på nasjonale tiltak iverksatt på bakgrunn av tilstandsvurderingen
- Evaluering av tiltakene, erfaringer og utfordringer



# Internasjonal kontekst

An abstract graphic design featuring a light gray background. In the lower-left corner, there is a solid orange rectangle containing the text "Internasjonal kontekst" in white. To the right of this rectangle, there is a complex, overlapping pattern of orange lines and circles of various sizes, creating a sense of movement and interconnectedness.

# Prøver og internasjonale undersøkelser i kontekst

---

## Utgangspunkt:

- OECD-studie fra 80 tallet om at Norge har dårlig system for å vurdere kvaliteten på opplæringen
- «PISA-sjokk 2001» (resultater fra 2000)
- Etablering av nasjonalt kvalitetsvurderingssystem fra 2004
- Arbeid med Kunnskapsløftet fra 2004 (innført 2006)
- Flere studier som viser svak vurderingskultur i norsk skole



# Hvorfor internasjonale studier?

---

- Mulighet for å relatere norsk utdanning til andre land
- Måling av trender og utviklingstrekk
- Innspill til nasjonal politikkutforming
- Kompetansebyggende
- Stort datagrunnlag for videre forskning og analyser

# Aktuelle studier og aktører

---

- **OECD**
  - PISA: 15-åringers kompetanse og ferdigheter i lesing, matematikk og naturfag
  - TALIS: Læreres og skolelederes oppfatning av læring og undervisning
- **IEA (International Association for the Evaluation of Educational Achievement)**
  - TIMSS: Matematikk og naturfag på 4(5) og 8(9) trinn
  - PIRLS: Leseferdigheter på 4(5) trinn
  - ICCS: Demokrati og medborgerskap
  - ICILS: Tverrfaglig IKT kompetanse

# Hva vi måler er hva vi får?

## – Har vi de riktige måleinstrumentene?

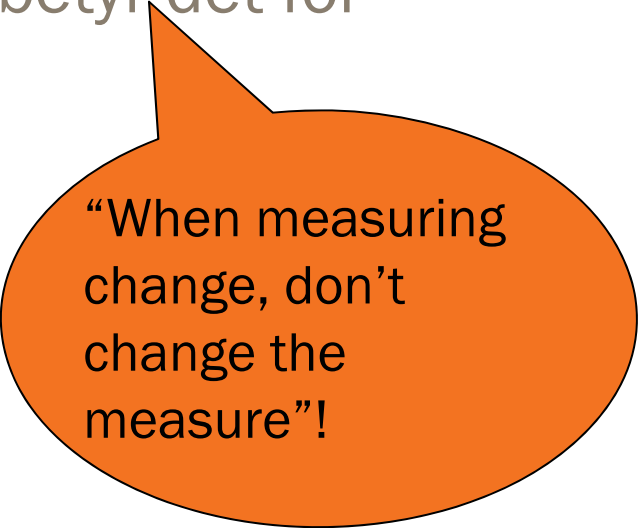
---

- Reading literacy, mathematics and science (PISA, TIMSS, PIRLS)
- ICT (PISA, ICILS)
- Democracy and Citizenship (ICCS)
- Collaboration and problem solving?
- Ability to communicate, reason and analyse?

# Problemstillinger og utviklingstrekk

---

- Testene dekker bare utvalgte kompetanser og ferdigheter – økende etterspørsel etter ”bredere” studier
- Forandring vs trend; testene kan virke ”konserverende”
- Overgang til computerbasert testing – hva betyr det for kostnader, trend, pålitelighet
- Behov for bedre forklaringsgrunnlag
  - Korrelasjoner vs. Årsakssammenhenger
- Representativitet og test-tretthet
- Hva med adaptiv testing?
- Det du tester er det du måler .....



“When measuring change, don’t change the measure”!

# Noen nye initiativ

---

- OECD: Education and Social Progress
  - Planlegger longitudinell studie for å undersøke sammenhenger mellom viktige ikke-kognitive ferdigheter og læring
- TALIS Video:
  - Planlegger komparativ videostudie for få mer kunnskap om sammenhengen mellom undervisning og læring.



An abstract graphic design featuring a light blue background. In the lower right corner, there is a complex pattern of overlapping orange circles and lines. Some circles are solid orange, while others are hollow with orange outlines. The lines are thin and orange, creating a network-like structure. A solid orange rectangular box is positioned in the lower left, containing white text.

# Elevenes læringsutbytte - status og utvikling

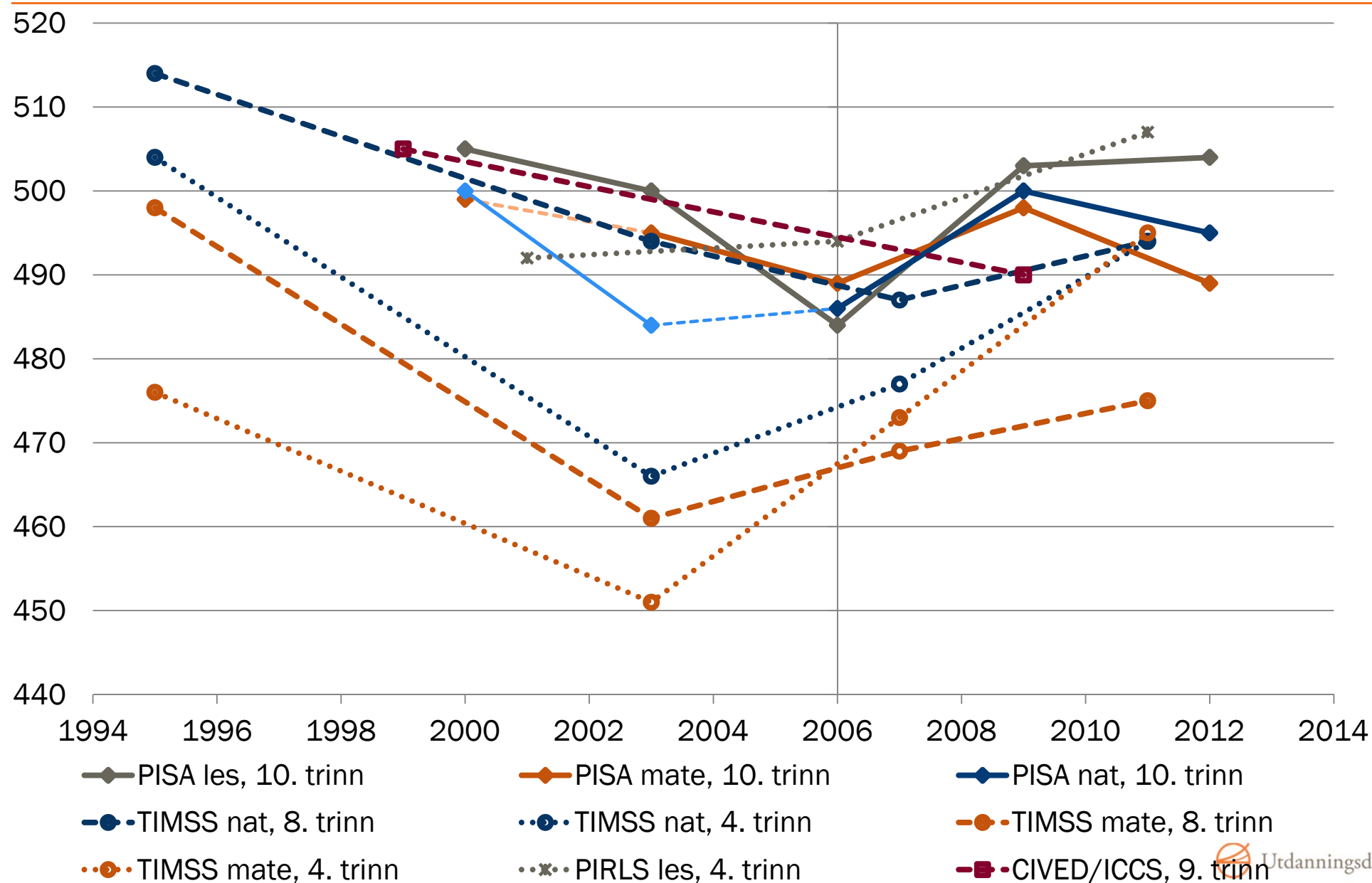
# Den norske utviklingen

---

- "PISA-sjokk" i 2001 – norske resultater mye svakere enn forventet
- Nedgangen fortsatte utover 2000-tallet, med et bunn-nivå i PISA i 2006
- Etter 2006 har vi sett en positiv utvikling både i TIMSS, PIRLS og PISA
- Den positive utviklingen i PISA stagnerte med resultatene fra 2012



# Trender på internasjonale studier



# Andre funn

---

- PIRLS og TIMSS viser tegn til forbedring i resultatene på barnetrinnene – de elevene som ble testet i 2011 blir testet på nytt på ungdomstrinnet i 2015
- PIRLS og TIMSS bekrefter positiv utvikling når det gjelder læringsmiljø
- Oppfølging av lekser og tilbakemeldinger til elevene ser ut til å være styrket
- Både TIMSS og TALIS bekrefter lite spekter av læringsstrategier og fordypningsstrategier i fagene
- TALIS 2008 viste svak tilbakemeldingskultur og lite bevisst oppfølging av kompetanseutvikling. Dette ser ut til være styrket i TALIS 2013



# Norsk skole – sett utenfra

---

- Et verdigrunnlag med vekt på likhet og likeverd
- Høy ressursinnsats, men resultater rundt gjennomsnittet
  - Tegn til bedring etter flere år med nedgang
- Skolekultur preget av positive relasjoner, tillit og tilhørighet
- Store individuelle forskjeller, men små skoleforskjeller
- Svak ledelseskultur – og svak tilbakemeldingskultur
  - Men positive tegn til bedre utvikling
- Smalt spekter av læringsstrategier, manglende oppfølging og fordypning

# OECD:

## Kjennetegn på systemer med god resultatoppnåelse

---

- Vekt på likhet og likeverd
- Hensiktsmessig ressursallokering med vekt på likeverdig fordeling
- Høy grad av autonomi
- Vekt på evaluering og vurdering (resultatoppfølging)
- Profesjonalisering av lærerrollen
- Satsing på høyt kompetente lærere

An abstract graphic design featuring a light gray background. On the right side, there is a complex pattern of overlapping orange circles and lines. Some circles are solid orange, while others are hollow with orange outlines. The lines are thin and orange, creating a network-like structure. A solid orange rectangular box is positioned on the left side, containing white text.

# Det nasjonale kunnskapsgrunnlaget

# Norsk utdanningspolitikk er basert på et bredt kunnskapsgrunnlag

---

- Nasjonale prøver og eksamen
- Elev- og lærlingeundersøkelsen
- Internasjonale undersøkelser
- Nasjonal forskning, evaluering og rapportering
- Utdanningsstatistikk
- Tilsynsrapporter
- Dialog med sektor

# Hva vet vi nå?

---

- Vi har rikelig med korrelasjonelle data på hvordan utdanningssystemene fungerer. Egentlig så mye at vi har problemer med å tolke motstridende resultater. (f.eks. PISA og TIMSS)
- Vi vet veldig lite om kausalitet.
- Vi vet at sammenhenger mellom variabler er ikke de samme i alle land.
  - Det samme virker sannsynligvis ikke i ulike land og kulturer.
  - Vi må alle gjøre våre egne analyser, i vår egen sammenheng og kontekst.
- Enighet om at en variabel er kanskje den aller viktigste og det er:
  - Læreren og hva som skjer i klasserommet



An abstract graphic design featuring a light blue background. In the lower right corner, there is a complex pattern of overlapping orange circles and lines. Some circles are solid orange, while others are hollow. Lines of varying thicknesses connect some of the circles, creating a network-like structure. The overall style is modern and minimalist.

## Eksempler på nasjonale tiltak

# Ulike nasjonale tiltak

---

- Utviklingstiltak
- Støtte og veiledning
- Kompetanseutvikling
- Tilsyn
- Regelverksendringer

# Atuelle tiltak nå

- Bedre læringsmiljø
- Vurdering for læring
- Fortsatt satsing på ungdomstrinnet
- Veilederkorps
- Styrking av realfagene på alle nivå
- Lærerløft
  - Stor vekt på etter- og videreutdanning for lærere



# Hva kjennetegner utviklingstiltakene (eksempler)?

---

Læringsressurser på nett

Bedre læringsmiljø

Vurdering for læring

Ungdomstrinnssatsingen

Nettverksbygging

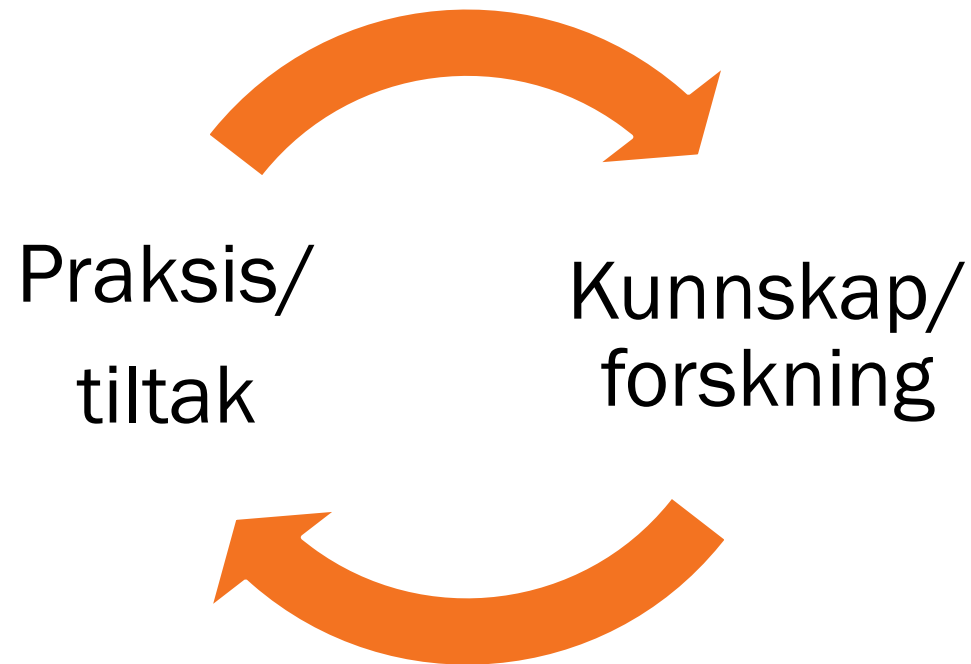
Skolebasert  
kompetanseutvikling

An abstract graphic design featuring a light gray background. In the lower right corner, there is a complex pattern of overlapping orange circles and lines. Some circles are solid orange, while others are hollow with orange outlines. The lines are thin and orange, creating a network-like structure. A solid orange rectangle is positioned in the lower left, containing white text.

**Hvordan skjer evalueringene?**

# En strategi for å lære

---



Hvordan etablere en god interaksjon mellom praksis/tiltak og kunnskap/forskning?

# Evaluering av statlige utviklingstiltak

---

- Utgangspunkt i mål for satsingene
- Metodetriangulering: Dokumentstudier, breddeundersøkelser, bruk av registerdata der det er aktuelt, case-studier
- Utfordringer ved evalueringene:
  - Gir god informasjon om hvilke prosesser som settes i gang med utgangspunkt i tiltakene
  - Kan gi god informasjon om deltakernes **opplevelse** av tilfredshet og måloppnåelse
  - Gir ikke mulighet for måling av effekt på lengre sikt

# Noen eksempler på ulike evalueringer

---

- Evaluering av Kunnskapsløftet
  - Omfattende paraplyprosjekt med eget programstyre. Mange forskningsmiljøer involvert.
- Satsing på ungdomstrinnet
  - Måle samspill mellom en rekke ulike virkemidler
  - Datainnsamling i to omganger på mange nivåer
  - Metodetriangulering
- Vurdering for læring
  - Tre faser: Kartlegging, casestudier av praksis, analyser av resultater og utbytte
- Et lag rundt læreren
  - Tre faser: Kunnskapsoversikt, modellutprøving, drøfting av modeller



# Hvordan kan evalueringene bli bedre?

---

- God kartlegging av situasjonen før tiltaket iverksettes (kunnskapsoversikt)
- Tydelighet på mål og kjennetegn på måloppnåelse
- Bevissthet om ønsket effekt på kort og lang sikt
  - Valg av tiltak må avhenge av hvilke effekter man ønsker å oppnå
- Hvilke effekter vil vi måle?
  - Utvikling i læringsutbytte (langsiktig)
  - Endring i atferd (på kort og lang sikt)

# Effektmåling

---

- Effekt: forskjell mellom utfall for en gruppe som har fått et tiltak og en som ikke har fått tiltaket
- For å måle effekt må man ha en sammenligningsgruppe (kontrollgruppe)
- Den klassiske effektmålingen: elevene deles i en forsøksgruppe og en kontrollgruppe
  - ..like stor sjanse for å havne i de to gruppene.

# Når bør vi vurdere RCT

---

- Når vi forsøker å finne ut hvorvidt et tiltak fører til endring i utfall/utbytte
- Før et tiltak skaleres opp
- For å avgjøre om et tiltak fortsatt «virker» etter at det er skalert opp.

# Når er RCTs mulig

---

- Lærere/foreldre/elever er villig til å delta
- Tiltaket innebærer en klar kontrast mot eksisterende praksis
- Tiltaket er klart definert med en klar «theory of action» (programteori)
- Utvalget er stort nok til å oppdage små endringer
- Tiltaket er relativt stabilt

# Når er det etisk forsvarlig?

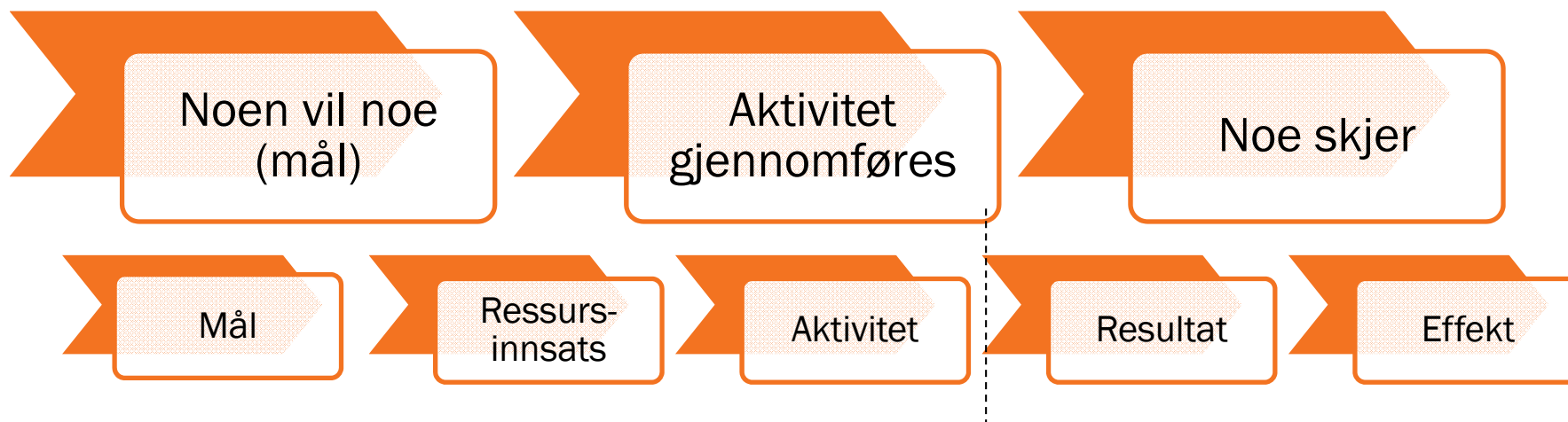
---

- Større etterspørsel etter tiltak enn tilbud – altså må prioritere mellom tiltak og mottakere
- Ingen kjente og uakseptable risikoer for de som deltar. Samtidig ikke kjent hvilke gevinster tiltaket gir.
- Deltagerne gir informert aksept for å delta



# Logikken bak effektmåling

---



- Vil det som skjer oppfylle målene?
- Hvor stor er effekten?
  - hvordan fungerer mekanismene i den gitte konteksten (programteorien)

# Effektmåling – vanlige utfordringer

---

- Uklare mål som effektene skal måles mot
  - Hva vil vi egentlig oppnå?
  - Finnes det en programteori? Holder den?
- Nivåproblematikk
  - Kontekstavhengige resultater
  - Stor mål – lite penger (der det gjelder)

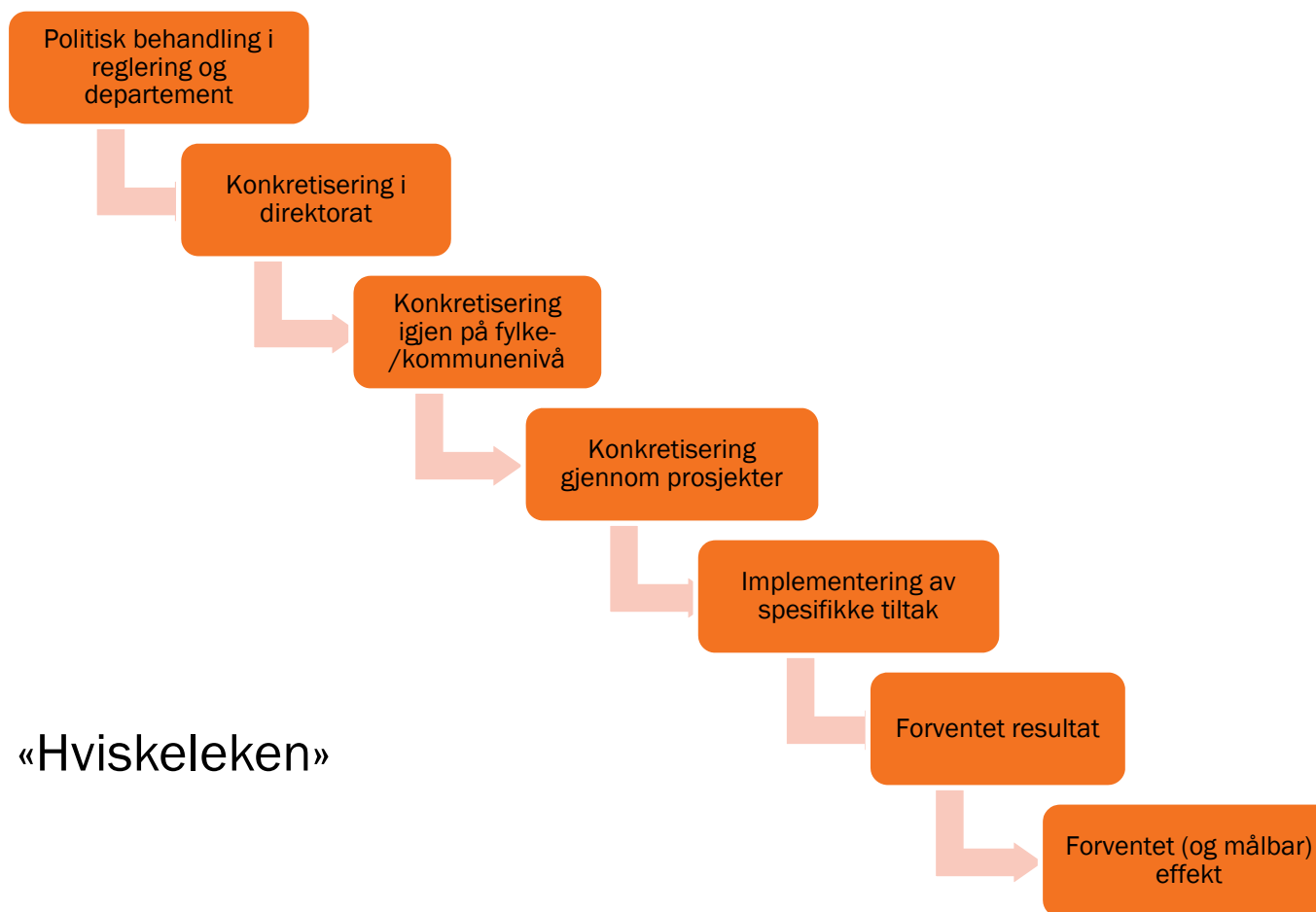
# Utfordringer forts

---

- Det kontrafaktiske dilemma
  - Er det mulig å vite hvordan verden ville være uten tiltaket?
- Tidsforskyvning mellom tiltak og effekt
  - Hvilket tiltak (blant mange) ga hvilken effekt?
- Alternativ ressursbruk
  - Hvor stor skal effekten være for at dette er den optimale bruken av ressursene?
- Kvantifiserbare mål
  - Indikatorer som treffer (i forhold til teori)



# Illustrasjon nivåproblematikk

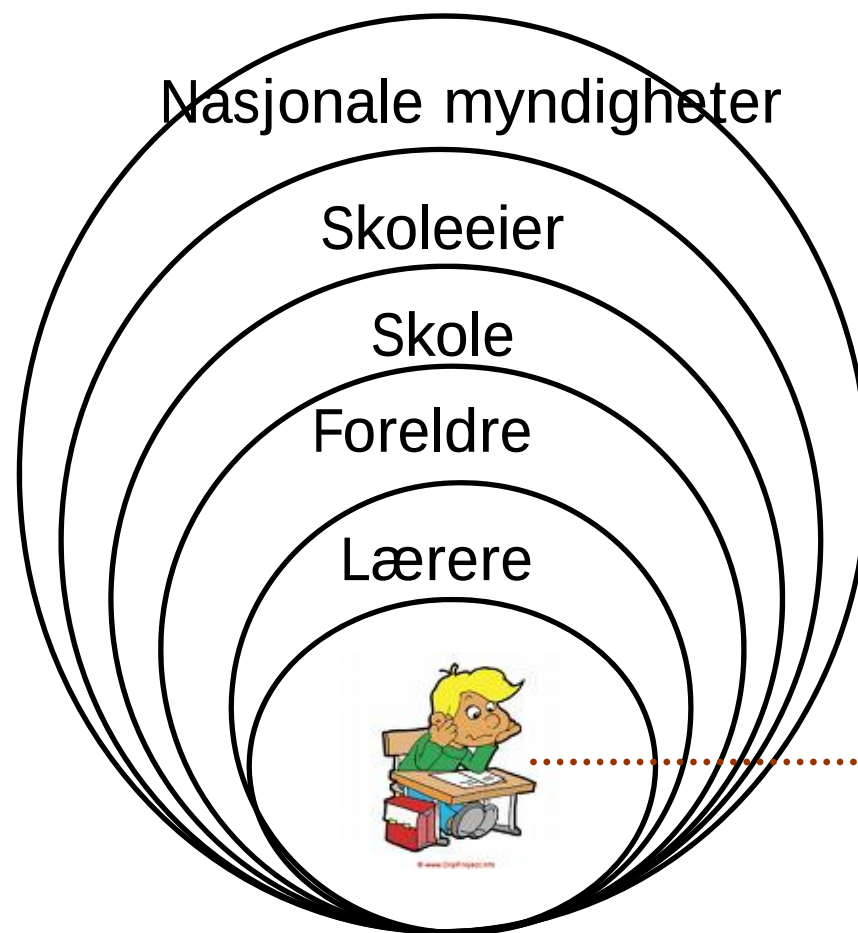


# Oppsummering

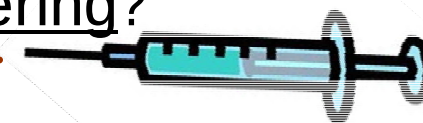
---

- Våre utdanningssystemer har aldri vært så godt dokumentert.
- Men:
  - Vi vet for lite om hva som foregår i klassene.
  - Vi vet også for lite om sammenhengen mellom undervisning og læring
- Vi trenger brukbare forklaringer på hva virker og hvorfor.

# Det grunnleggende spørsmålet



Hva gjør vi for å  
stimulere til god  
læring?





Utdanningsdirektoratet