



## Tilleggsdokumentasjon til foreløpige resultater fra FFIs prediksjonsturnering

### Forfattere

Alexander William Beadle

Prosjektnummer 1553

13. januar 2022

### Godkjenner

Alf Christian Hennum, *forskningsleder*

*Dokumentet er elektronisk godkjent og har derfor ikke håndskreven signatur.*

### Emneord

Prediksjon, sikkerhetspolitikk, treffsikkerhet, kognitiv psykologi

### Sammendrag

Dette notatet inneholder histogrammer som viser fordelingen til alle avhengige og uavhengige variabler som ligger til grunn for de foreløpige resultatene fra FFIs prediksjonsturnering og sammenligningene av disse med resultatene fra den amerikanske turneringen *Good Judgment Project* (GJP). Verdiene i FFIs turnering er basert på 833 deltagere som svarte på minst 20 % av de 150 spørsmålene som så langt er avgjort.



---

---

# Innhold

<b>1</b>	<b>Innledning</b>	<b>3</b>
<b>2</b>	<b>Generell treffsikkerhet</b>	<b>4</b>
2.1	Tema	4
2.2	Type	7
2.3	Tidsperspektiv	8
2.4	Prediksjonstidspunkt	9
2.5	Ekspérimentgrupper	10
2.6	Treffprosent	12
<b>3</b>	<b>Ekspertise</b>	<b>13</b>
3.1	Utdanningsnivå	13
3.2	Forsvars- og sikkerhetspolitisk kompetanse	14
3.3	Ansatt i forsvarssektoren	19
3.4	Bruk i media	20
<b>4</b>	<b>Individuelle variasjoner</b>	<b>21</b>
4.1	Individuell treffsikkerhet	21
4.2	Treffsikkerhet over tid	21
4.3	Individuelle egenskaper	22
<b>5</b>	<b>Norske superforecastere</b>	<b>34</b>
5.1	Gjennomsnittlig treffsikkerhet	34
5.2	Individuelle egenskaper	34

---

---

# 1 Innledning

Dette notatet inneholder histogrammer som viser fordelingen til alle avhengige og uavhengige variabler som ligger til grunn for de foreløpige resultatene fra FFIs prediksjonsturnering beskrevet i en egen FFI-rapport.<sup>1</sup> Hensikten er å dokumentere formen på fordelingene, fordi denne har betydning for hvilke statistiske tester som er best egnet til å undersøke om det er signifikante forskjeller mellom deltageres treffsikkerhet og til å måle korrelasjoner med individuelle egenskaper.

I dette notatet er alle verdier basert på de 833 deltagerne som har svart på minst 20 % av de 150 spørsmålene som har blitt avgjort så langt.<sup>2</sup> Notatet er organisert i delkapitler basert på de fire forskningsspørsmålene som ble besvart i resultatrapportens kapittel 5. Disse omhandler:

1) generell treffsikkerhet, basert på snittscorene til alle deltagerne i turneringen, 2) ekspertise, basert på snittscorene til ulike grupper deltagere, 3) individuelle variasjoner, basert på deltageres scores på en rekke kognitive evner og tenkemåter som henger sammen med treffsikkerheten, og 4) superforecastere, basert på snittscorene til de beste, nest beste og alle andre deltagerne og hver av disse tre gruppenes scores på de individuelle egenskapene.

I rapporten med de foreløpige resultatene sammenlignes funnene fra FFIs turnering med resultatene fra den amerikanske turneringen *Good Judgment Project* (GJP).<sup>3</sup> Dette notatet inkluderer derfor også histogrammer som viser fordelingen til alle variabler fra GJP. Her skilles det mellom to datasett som ble brukt i de to mest sammenlignbare studiene fra GJP av hhv. sammenhengene mellom treffsikkerhet og individuelle egenskaper (GJP200) og analysen av superforecastere (GJP350). For å illustrere eventuelle forskjeller i fordelingene av scores på tvers av turneringer og studier inkluderer histogrammene verdiene fra både FFIs turnering (blå), GJP200 (lys grønn) og GJP350 (mørk grønn). Hvis variablene er målt på samme måte, vises verdiene i samme figur. Hvis variablene er målt ulikt, vises verdiene i separate histogrammer.

Antall deltagere og spørsmål som hvert histogram er basert på, oppgis i parentes. Ytterpunktene i figurene er basert på de laveste og høyeste Brier-scorene innenfor hver kategori. Dette gjør det mulig å se hvordan scorene fordeler seg mellom deltagerne innad i og på tvers av turneringene.

---

<sup>1</sup> Beadle, Alexander W. (2021), 'FFIs prediksjonsturnering – datagrunnlag og foreløpige resultater', *FFI-rapport 21/00737* (Kjeller: FFI).

<sup>2</sup> For alle spørsmålene som ble stilt, inkludert svarene på de 150 foreløpig avgjorte, se Beadle, Alexander W. (2021), 'FFIs prediksjonsturnering – spørsmålskatalog', *FFI-rapport 21/00736* (Kjeller: FFI).

<sup>3</sup> For en oppsummering av de viktigste funnene fra GJP, se Beadle, Alexander W. (2021), 'Hvordan forbedre treffsikkerheten til prediksjoner av internasjonal politikk? – en litteraturgjennomgang', *FFI-rapport 21/00735* (Kjeller: FFI).

---

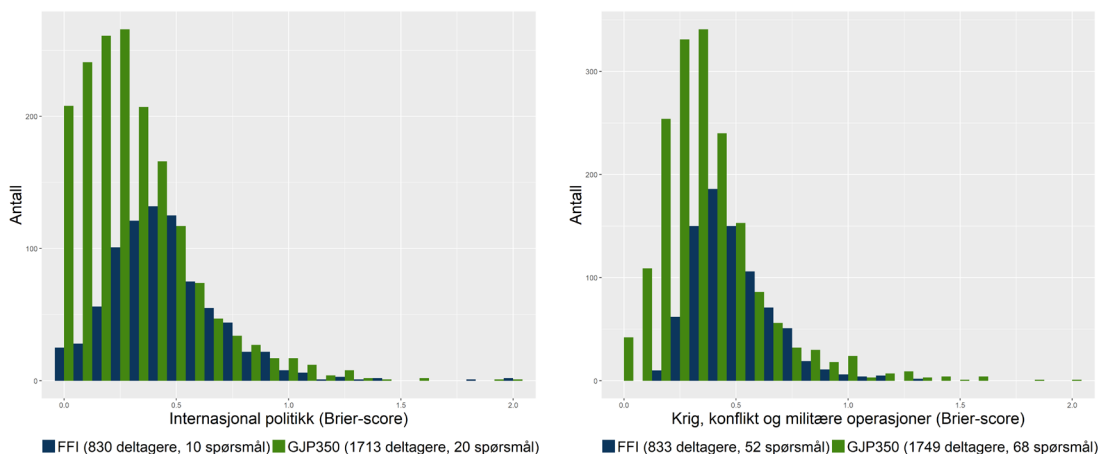
---

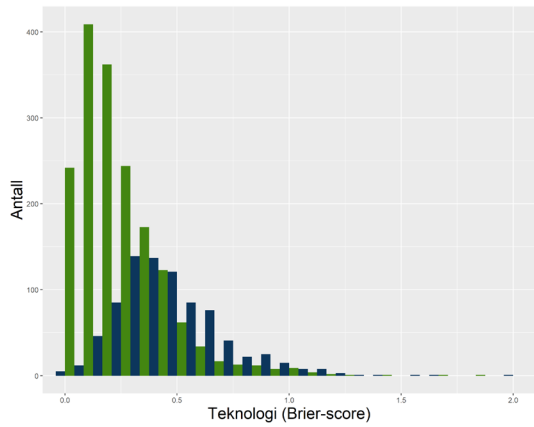
## 2 Generell treffsikkerhet

Dette delkapitlets histogrammer viser fordelingene av gjennomsnittlige Brier-scores i FFIs og GJPs turneringer, inndelt etter spørsmålenes temaer, type og tidsperspektiver, de forskjellige tidspunktene deltagerne kunne predikere på og eksperimentgruppene i GJP. Av disse er prediksjonstidspunktet spesielt viktig, siden dette var en av de største forskjellene mellom FFIs og GJPs turneringer. Mens deltagerne i GJP kunne oppdatere sine prediksjoner helt frem til spørsmålene ble avgjort, kunne deltagerne i FFIs turnering kun predikere i løpet av den første uken etter at spørsmålene ble publisert. Her og i senere analyser gjøres skilles det derfor mellom treffsikkerheten basert på disse to ulike prediksjonstidspunktene. Dette delkapitlet viser også fordelingene av treffprosentene til deltagerne i begge turneringer, som ble brukt som et alternativt mål på treffsikkerheten. Siden det bare er GJP350-studien som resultatene fra FFIs turnering ble sammenlignet med i selve rapporten, er ikke snittscorene fra GJP200 inkludert her.

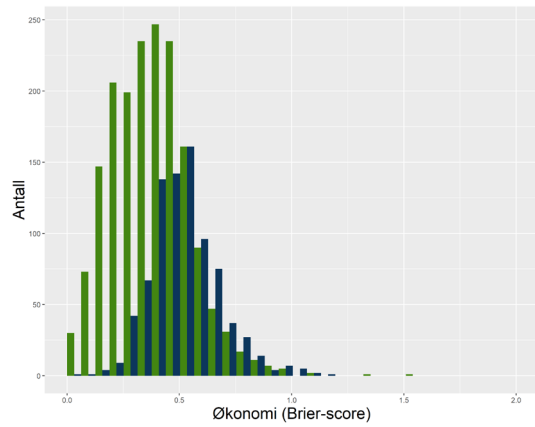
Her er treffsikkerheten målt ved vanlig Brier-score, som går på en skala fra 0 til 2, der lavere score er bedre. Hver deltagers score er basert på snittet av alle Brier-scorene de har fått på alle spørsmål, gitt at de oppfylte de respektive turneringenes minstekrav. Med unntak av figurene om tidsperspektiv, inkluderes det bare kategorier med minst ti spørsmål hver.

### 2.1 Tema

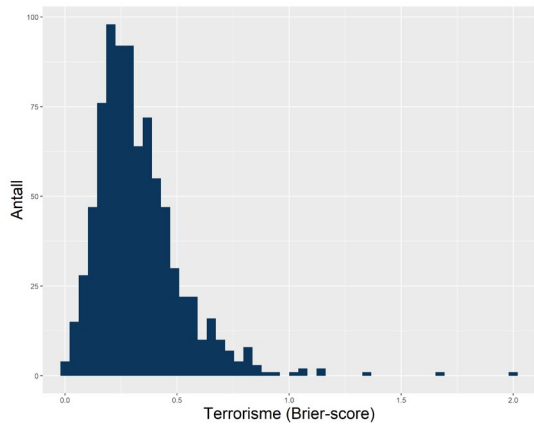




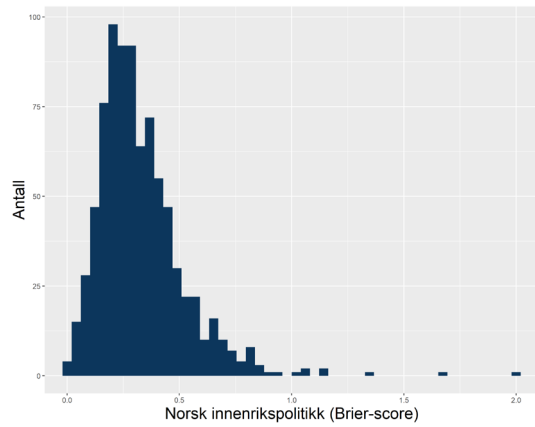
■ FFI (833 deltagere, 17 spørsmål) ■ GJP350 (1718 deltagere, 18 spørsmål)



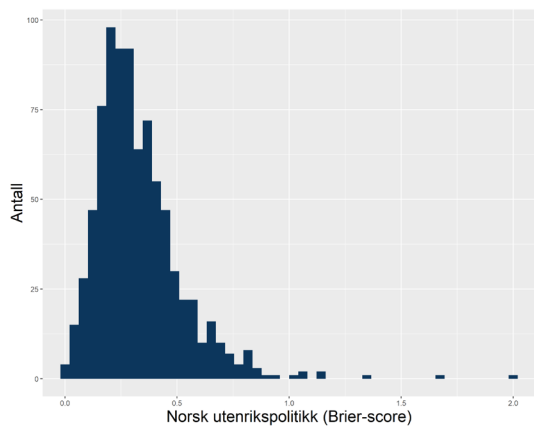
■ FFI (833 deltagere, 24 spørsmål) ■ GJP350 (1745 deltagere, 64 spørsmål)



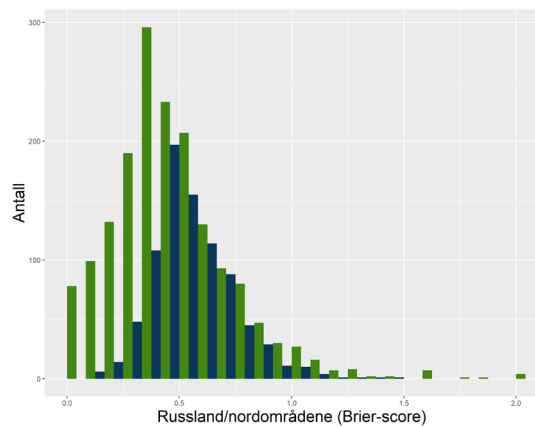
■ FFI (832 deltagere, 12 spørsmål)



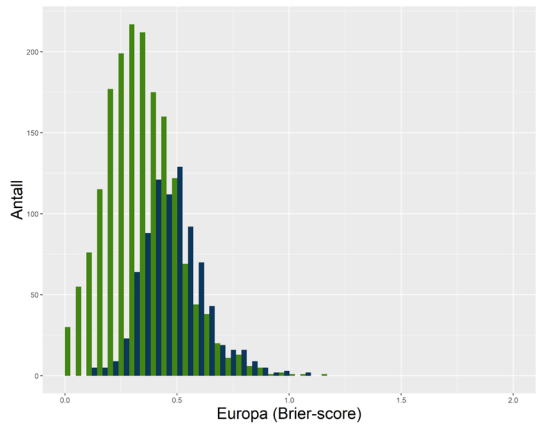
■ FFI (832 deltagere, 12 spørsmål)



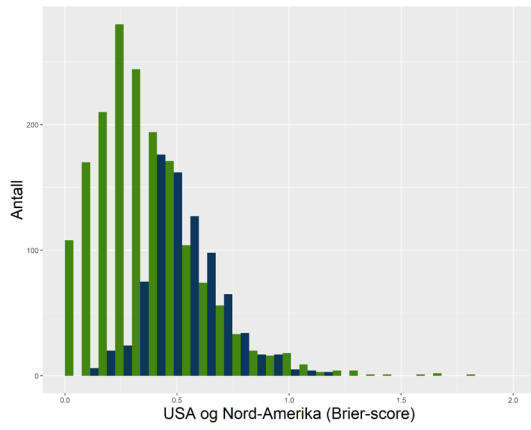
■ FFI (832 deltagere, 12 spørsmål)



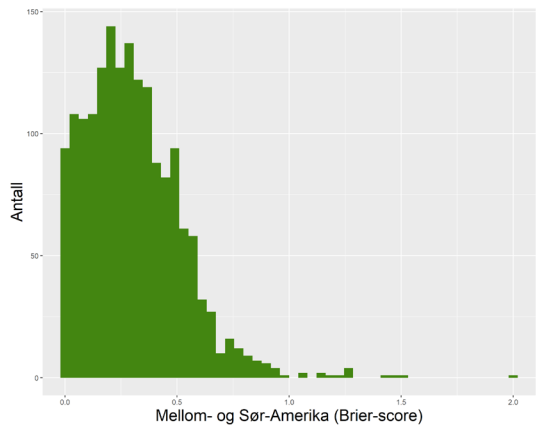
■ FFI (833 deltagere, 28 spørsmål) ■ GJP350 (1690 deltagere, 16 spørsmål)



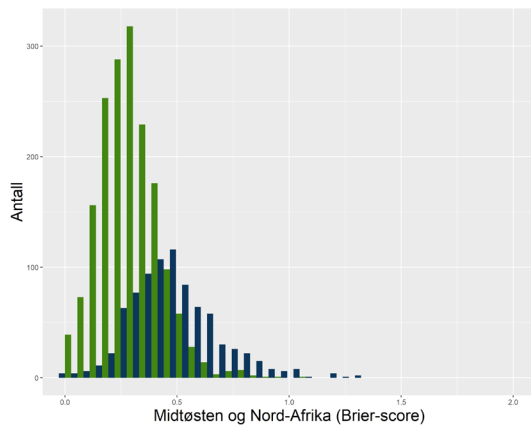
■ FFI (833 deltagere, 33 spørsmål) ■ GJP350 (1750 deltagere, 66 spørsmål)



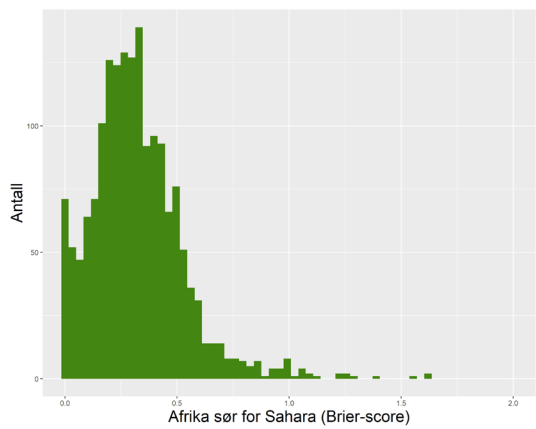
■ FFI (833 deltagere, 29 spørsmål) ■ GJP350 (1724 deltagere, 20 spørsmål)



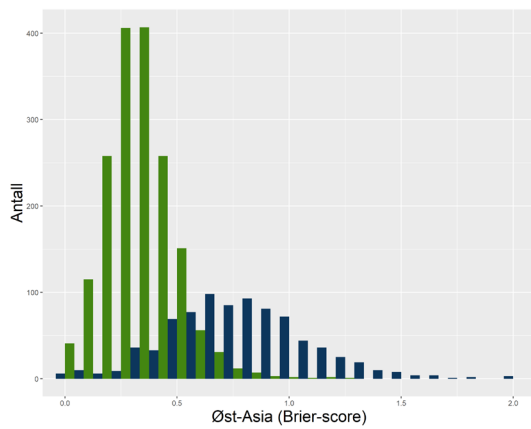
■ GJP350 (1713 deltagere, 24 spørsmål)



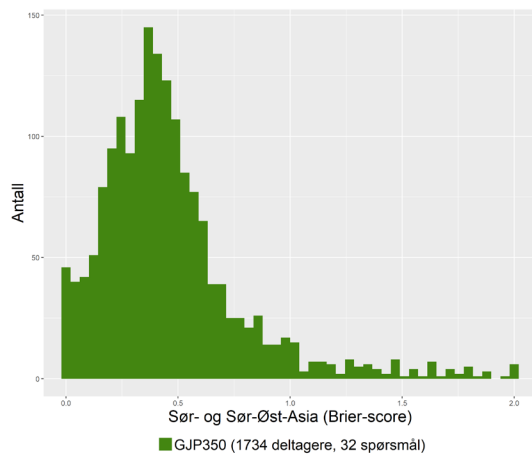
■ FFI (833 deltagere, 23 spørsmål) ■ GJP350 (1751 deltagere, 115 spørsmål)



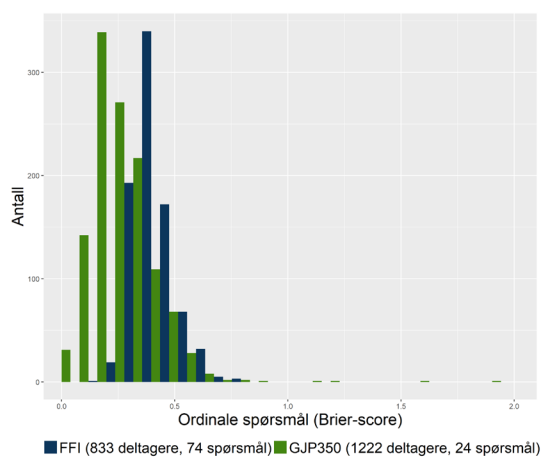
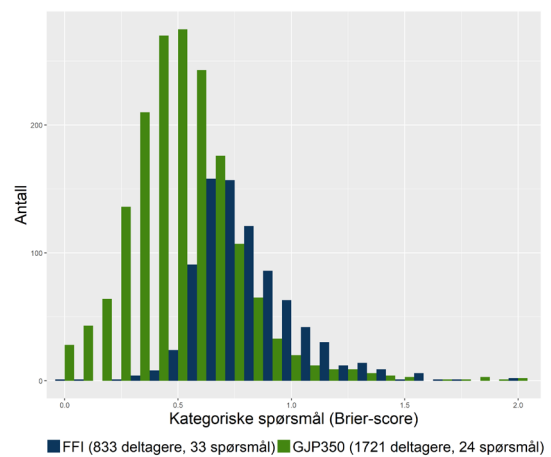
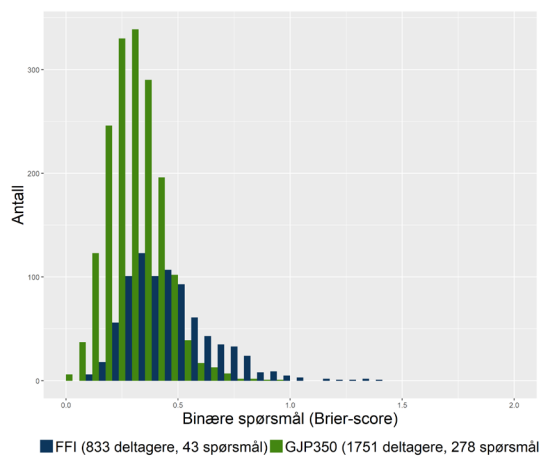
■ GJP350 (1703 deltagere, 32 spørsmål)



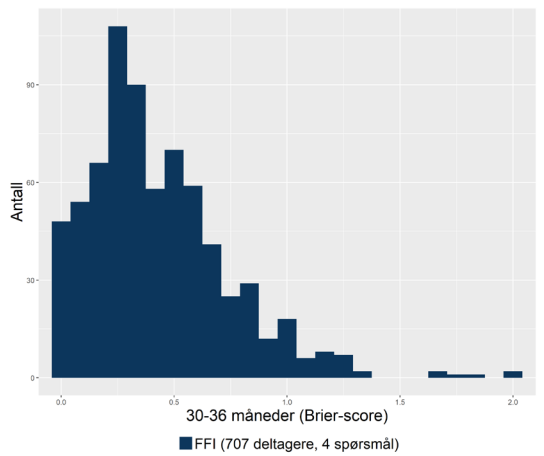
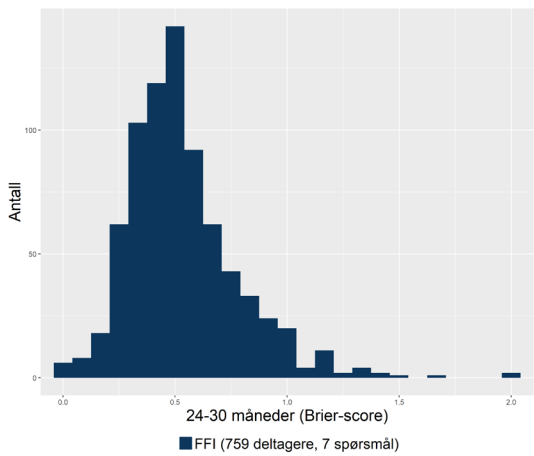
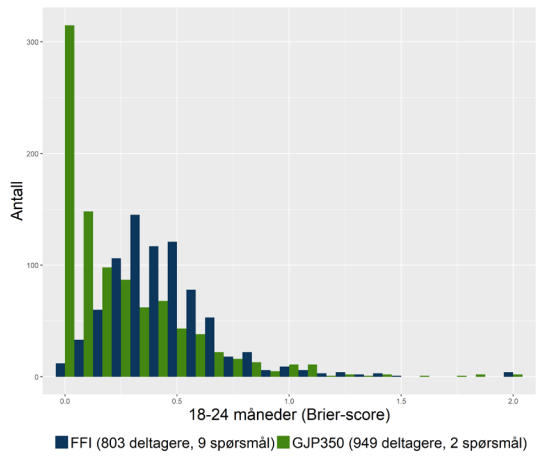
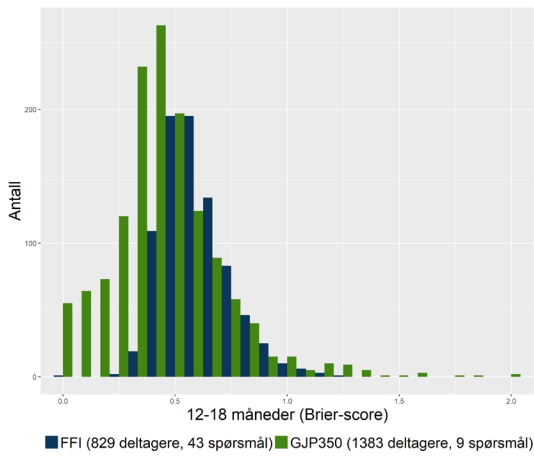
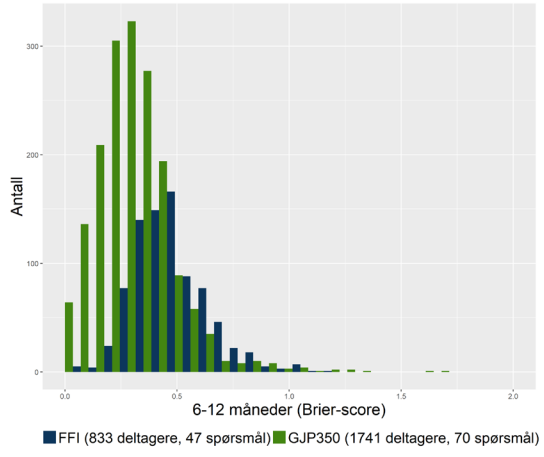
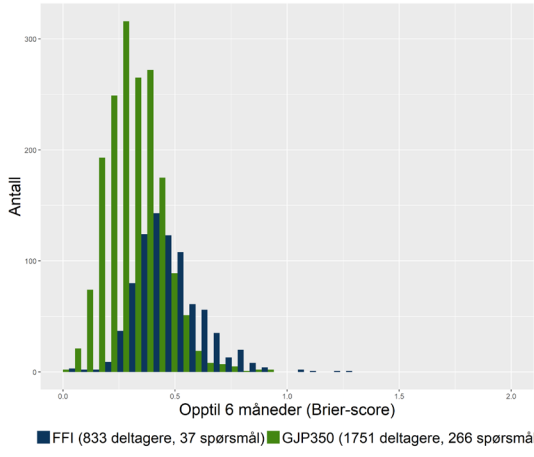
■ FFI (831 deltagere, 10 spørsmål) ■ GJP350 (1751 deltagere, 64 spørsmål)



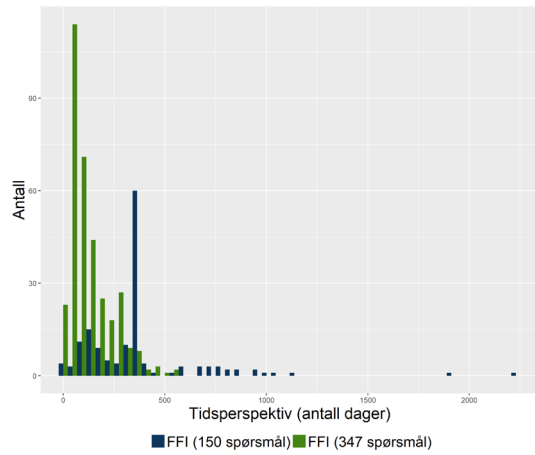
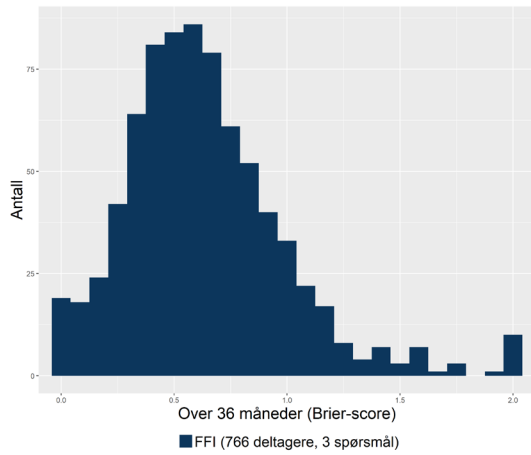
## 2.2 Type



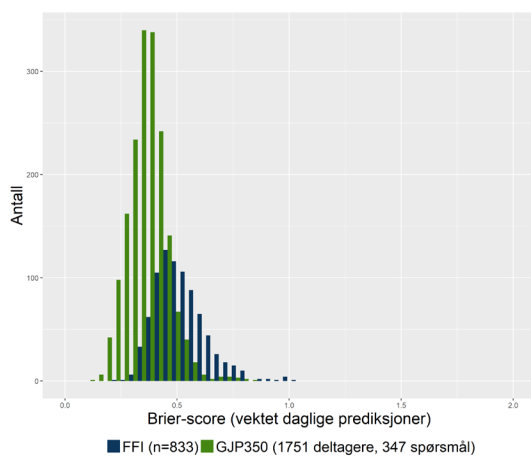
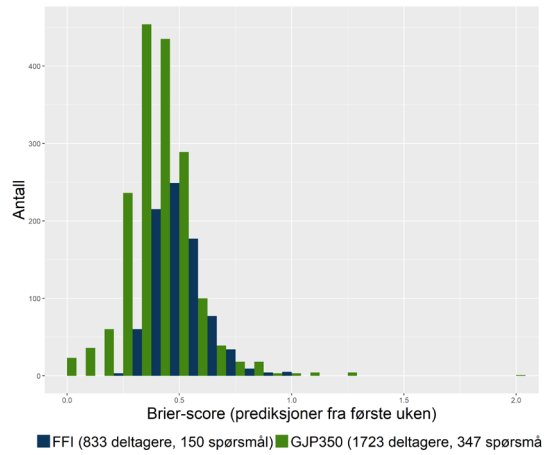
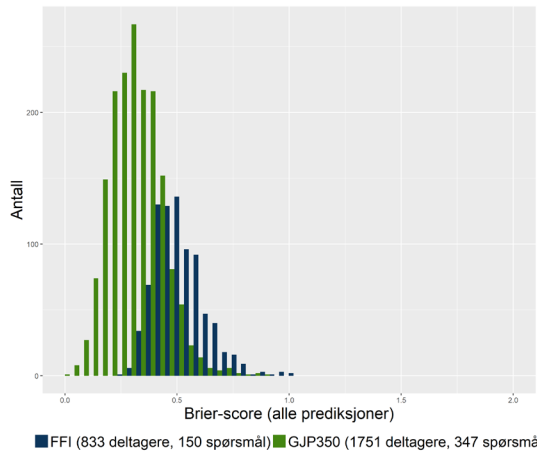
### 2.3 Tidsperspektiv



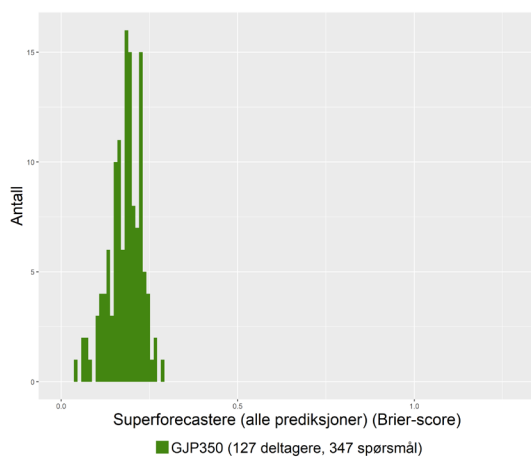
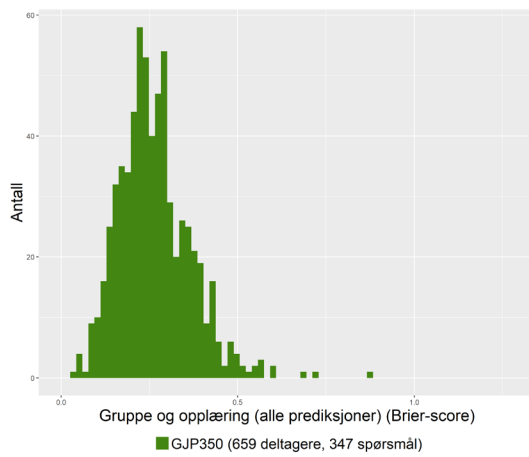
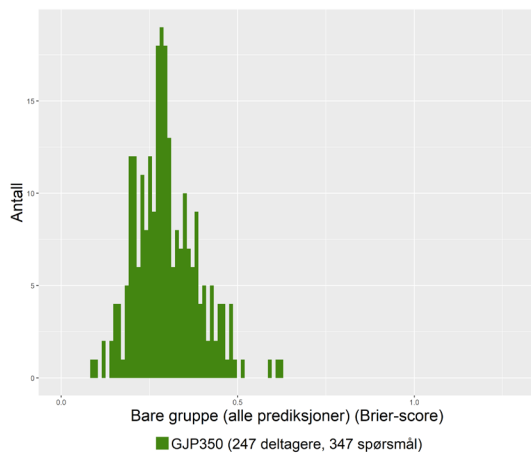
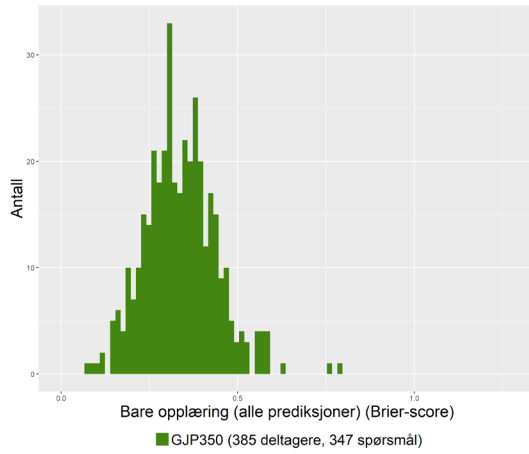
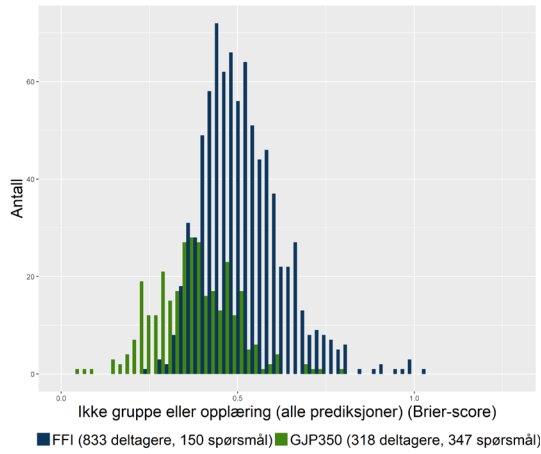


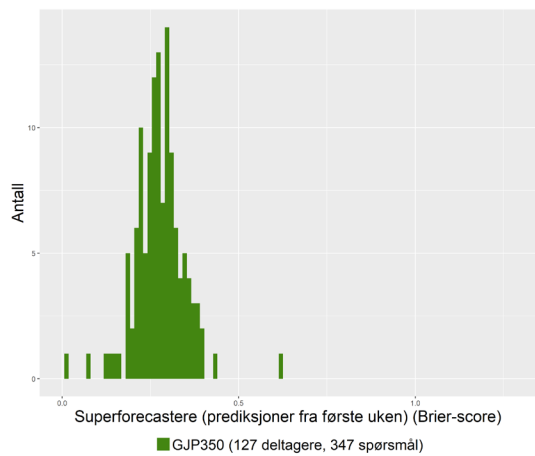
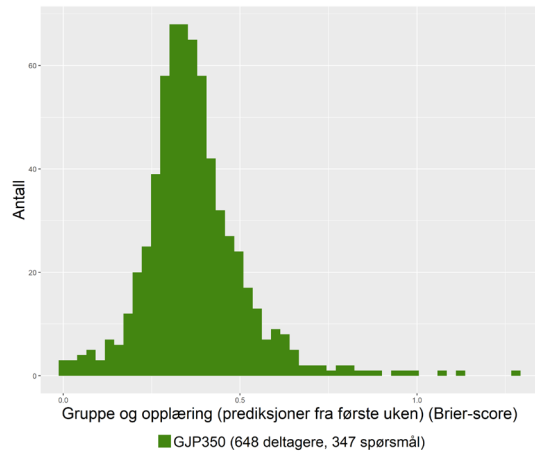
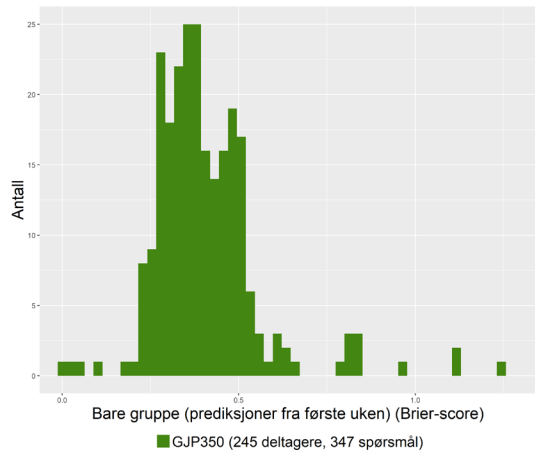
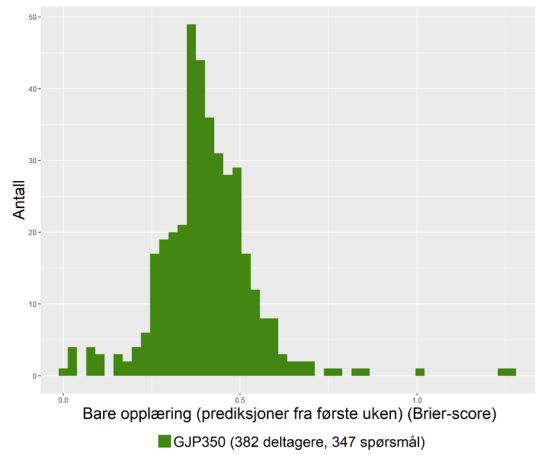
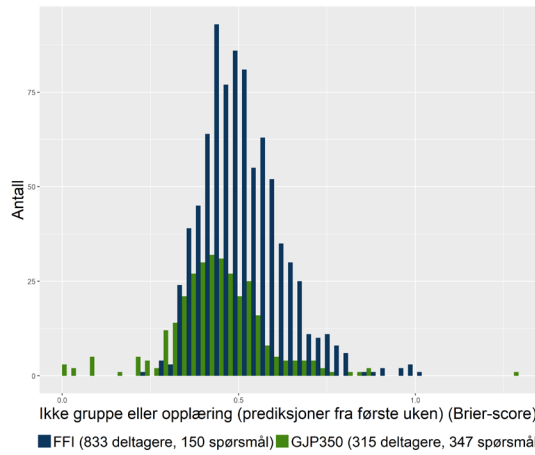


## 2.4 Prediksjonstidspunkt

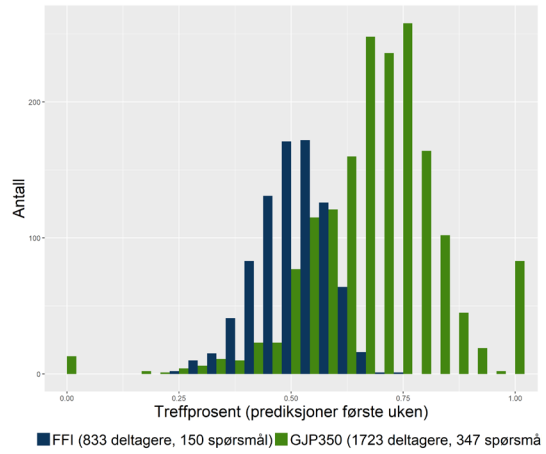
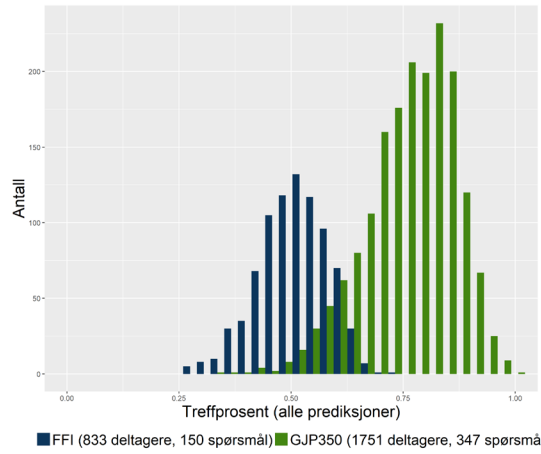


## 2.5 Eksperimentgrupper





## 2.6 Treffprosent

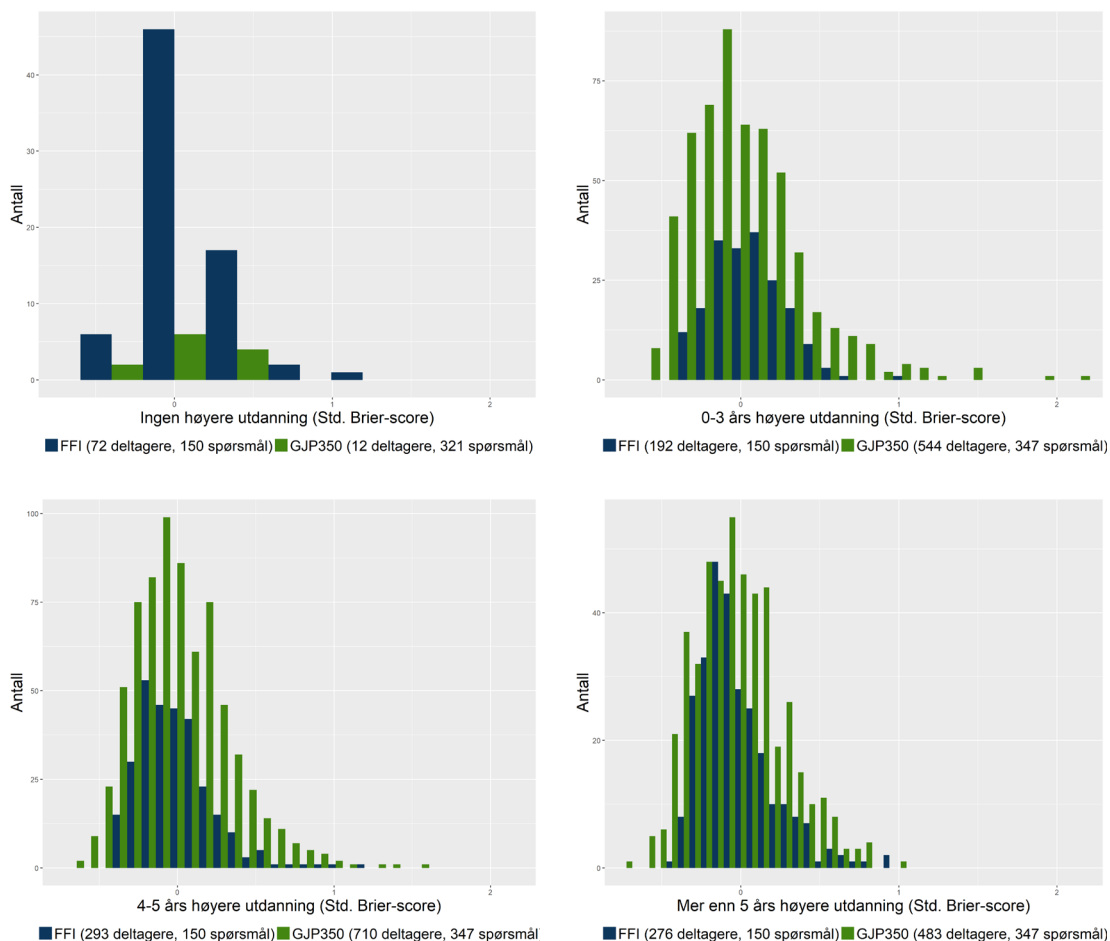


### 3 Ekspertise

Dette delkapitlets histogrammer viser fordelingene av Brier-scorene til deltagerne i FFIs turnering, basert på de ulike målene på ekspertise som ble analysert i rapporten.

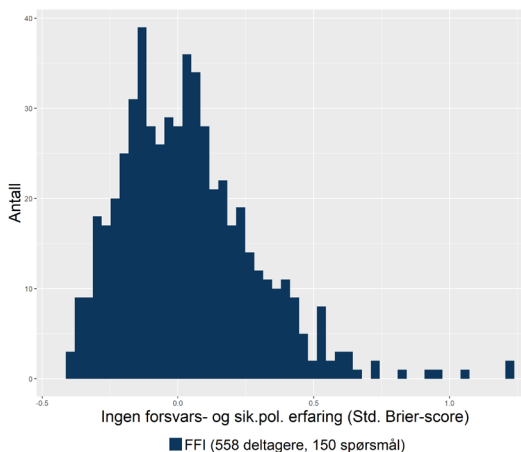
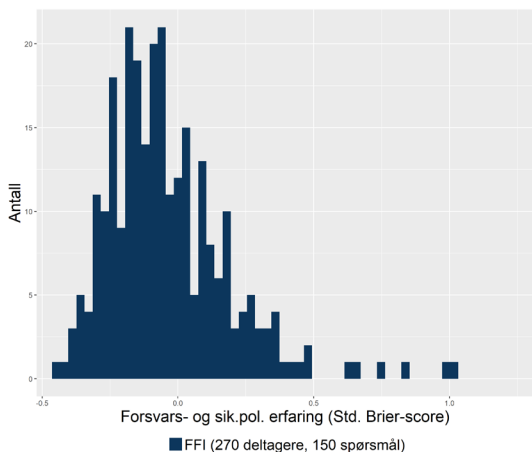
Her er treffsikkerheten målt ved standardisert Brier-score, som sier noe om hvor mye bedre eller dårligere en deltager traff sammenlignet med andre som svarte på samme spørsmål. Deltagernes standardiserte score er basert på snittet av alle scorene de har fått på alle spørsmål, gitt at de oppfylte de respektive turneringenes minstekrav. Her inkluderes bare grupper med minst fem deltagere hver. Siden det bare er utdanningsnivå som det finnes informasjon om i GJPs datasett, er det bare på denne variabelene at deltageres scores sammenlignes på tvers av turneringene.

#### 3.1 Utdanningsnivå

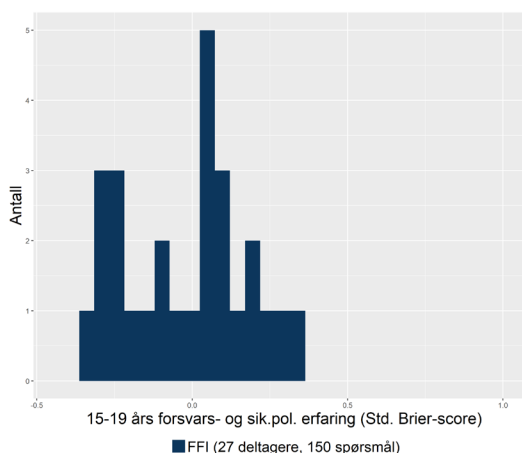
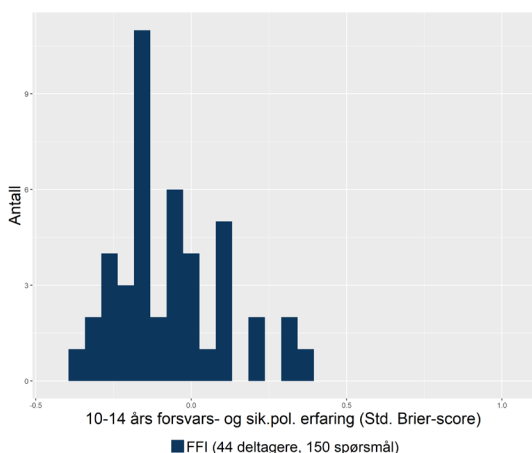
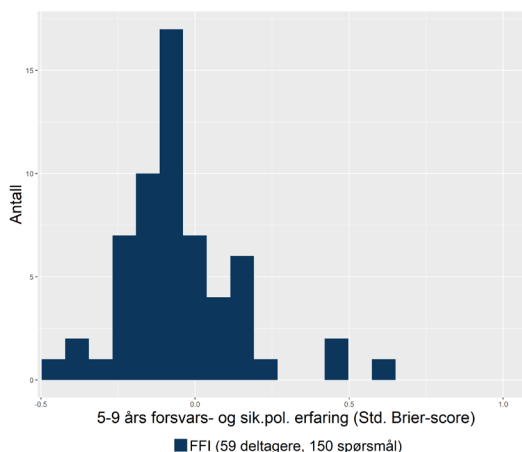
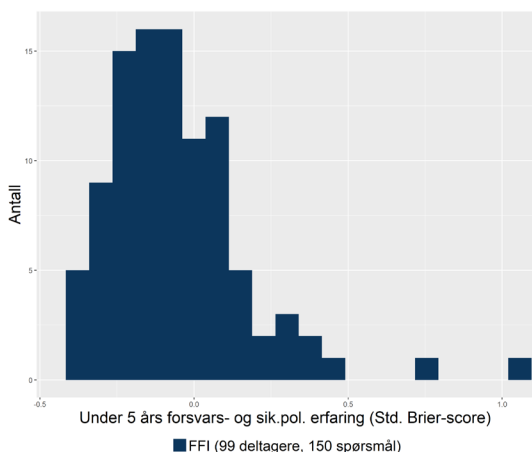


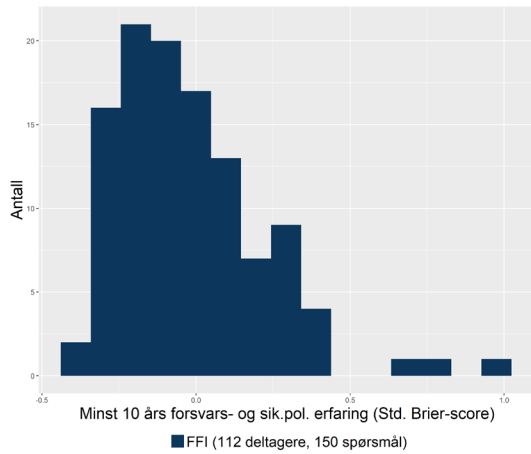
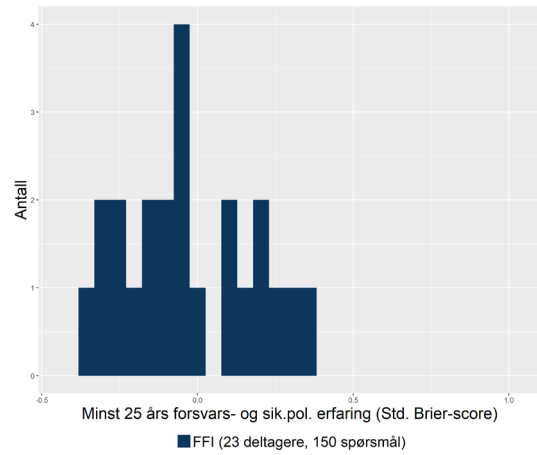
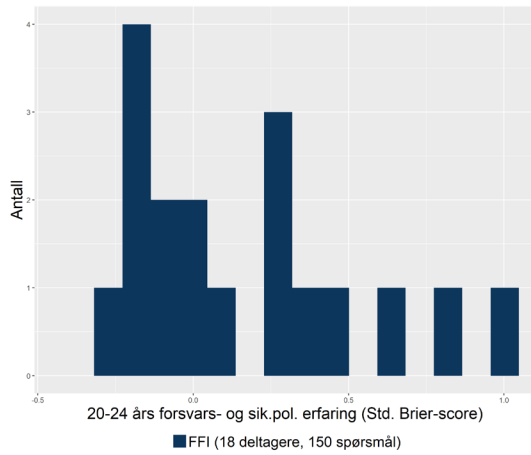
## 3.2 Forsvars- og sikkerhetspolitisk kompetanse

### 3.2.1 Forsvars- og sikkerhetspolitisk arbeidserfaring

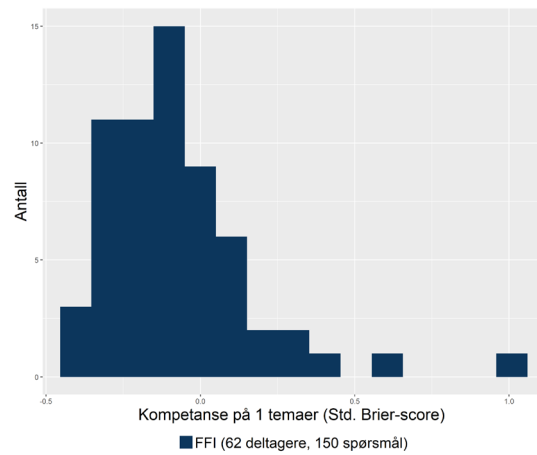
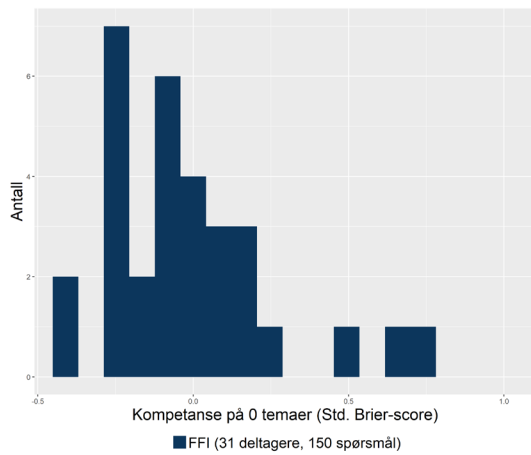


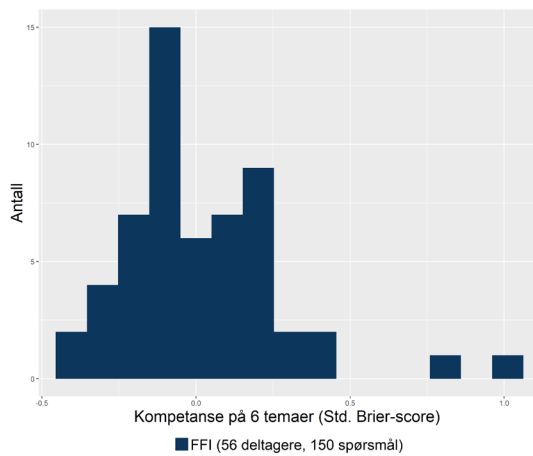
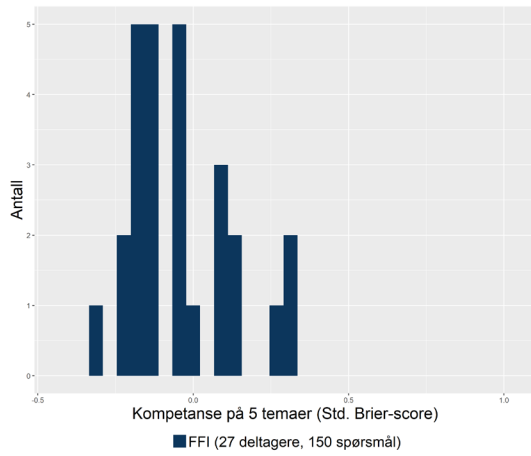
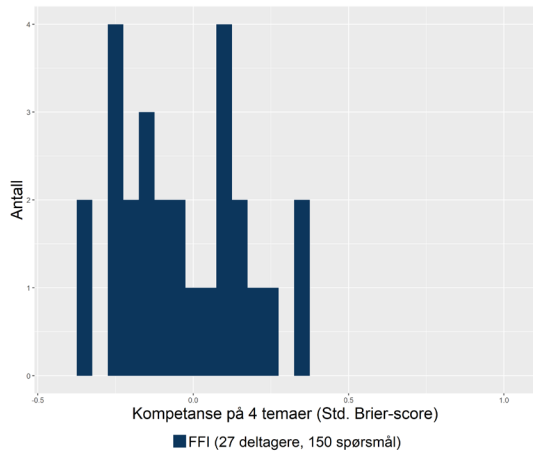
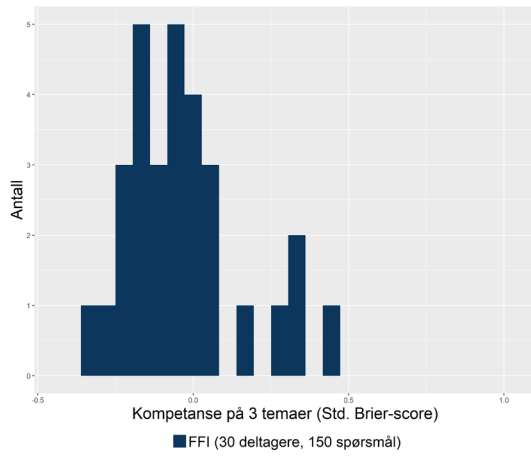
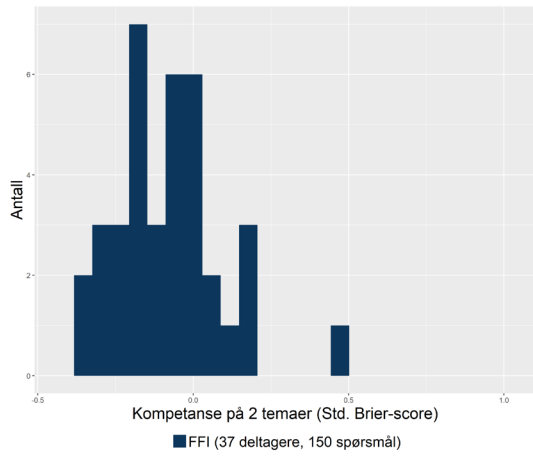
### 3.2.2 Antall års forsvars- og sikkerhetspolitisk arbeidserfaring



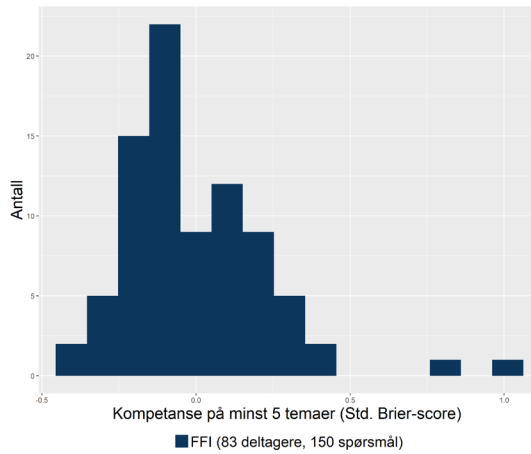
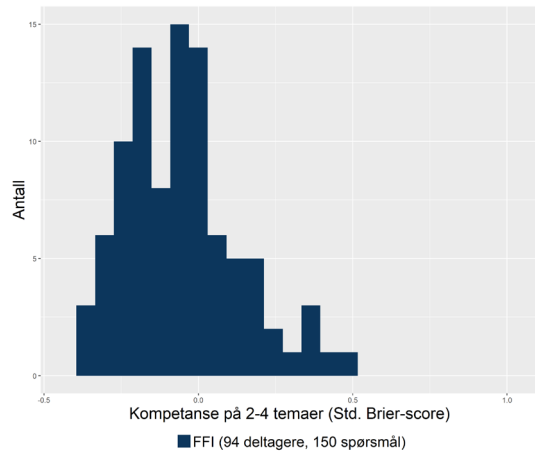
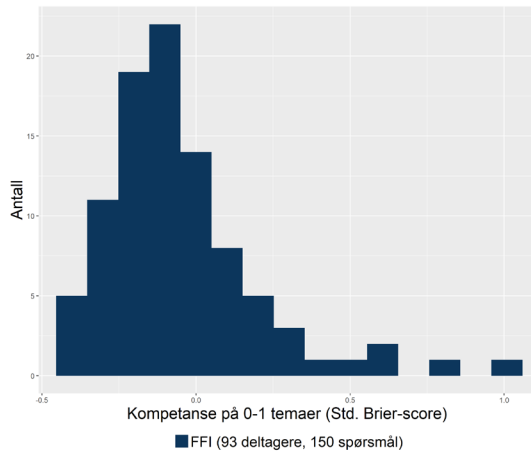


### 3.2.3 Antall kompetansetemaer

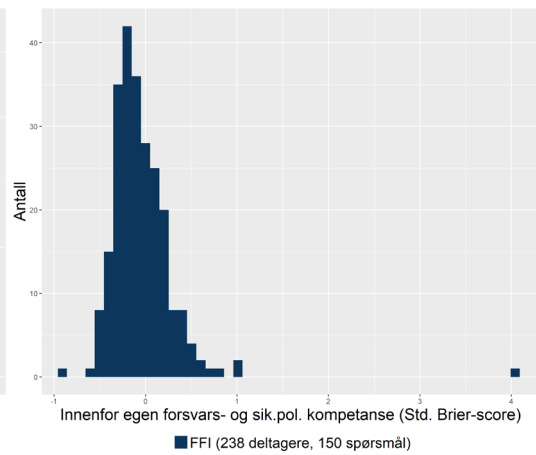
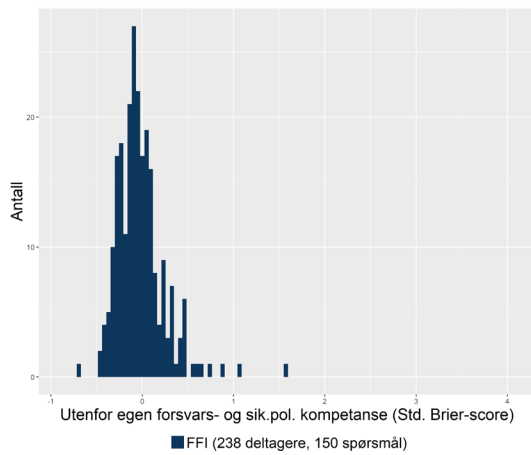




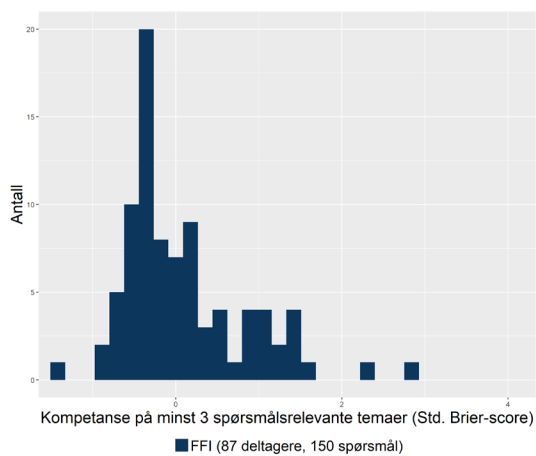
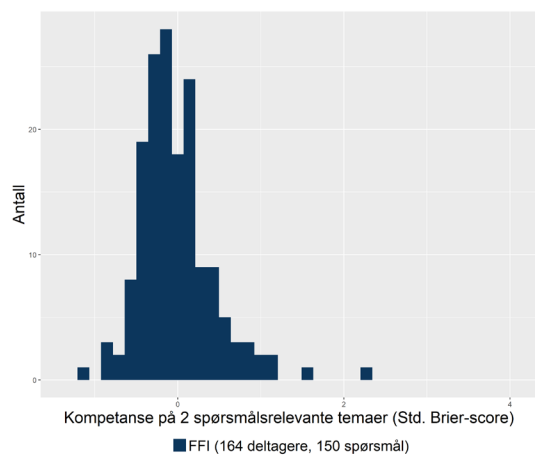
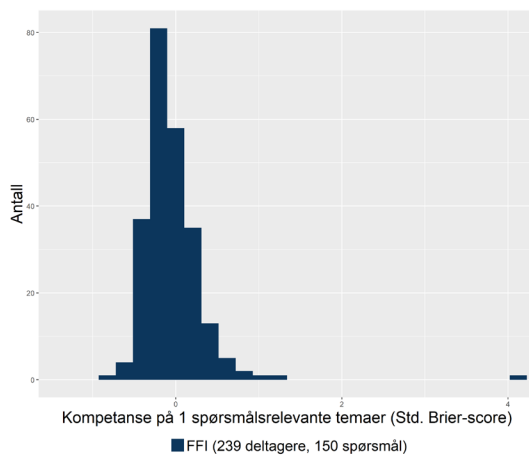
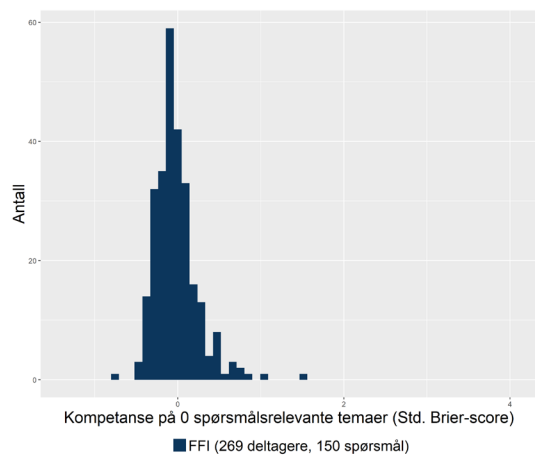




### 3.2.4 Innenfor eller utenfor egen kompetanse

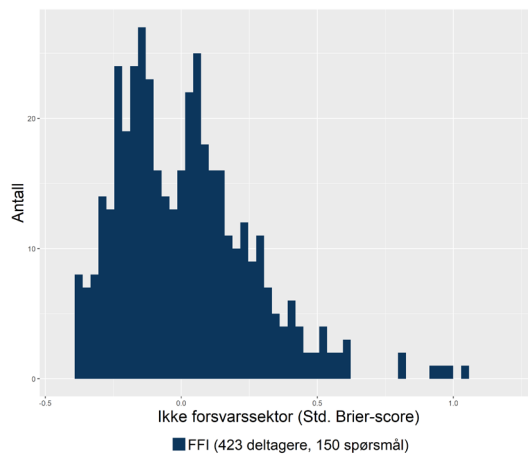
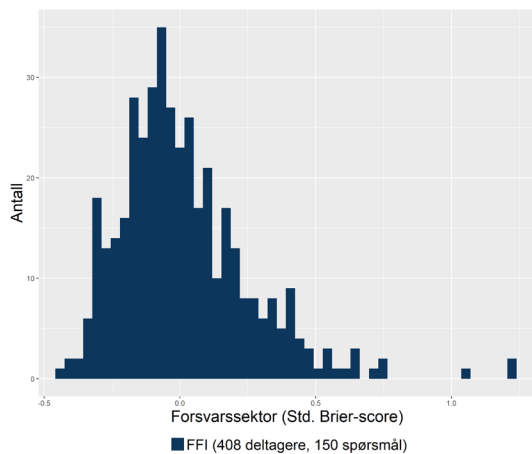


### 3.2.5 Antall relevante kompetansetemaer

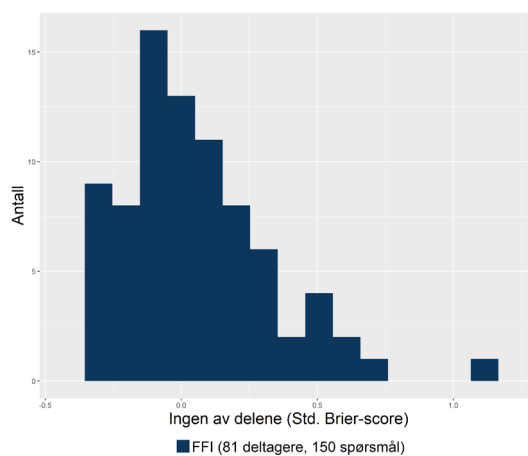
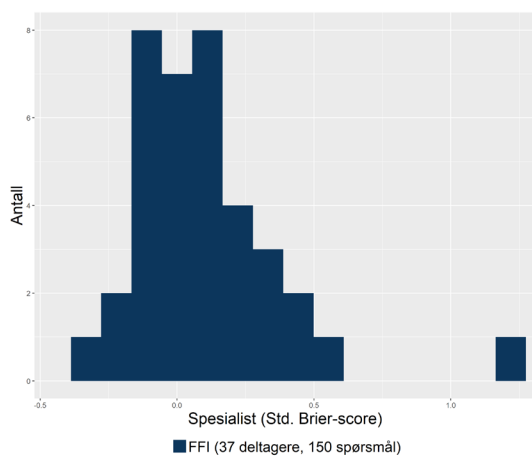
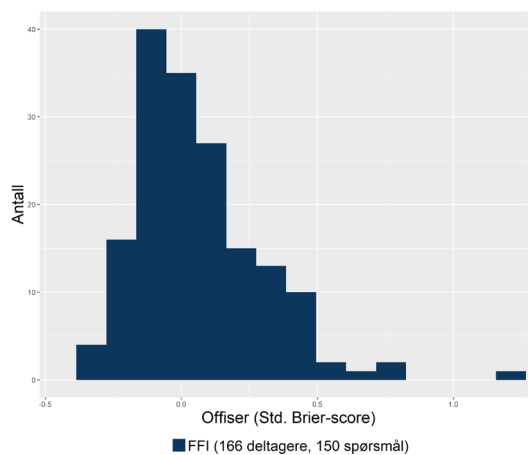
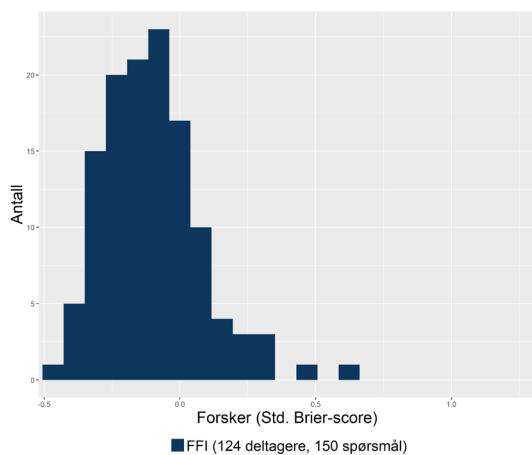


### 3.3 Ansatt i forsvarssektoren

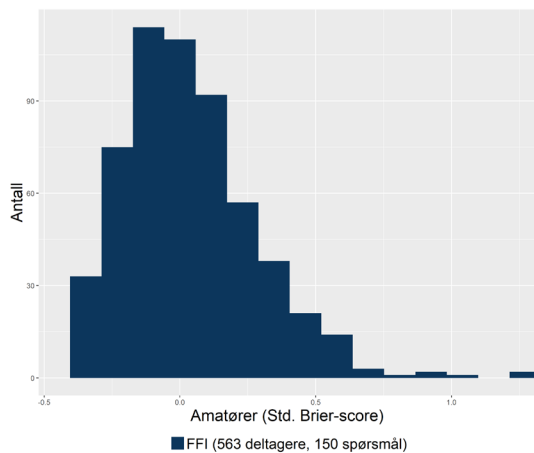
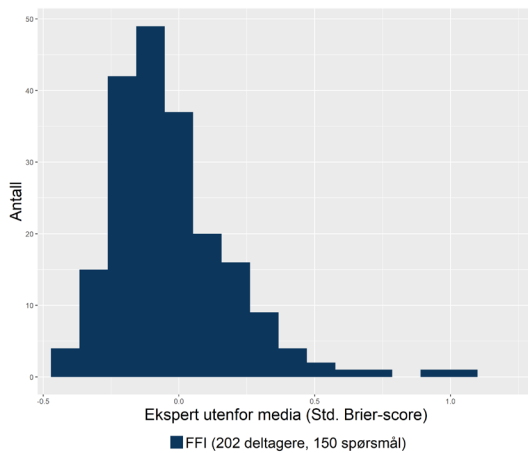
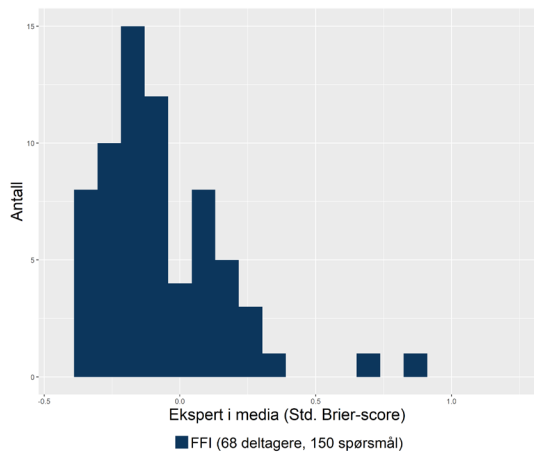
#### 3.3.1 Ansatt innenfor forsvarssektoren eller ikke



#### 3.3.2 Stilling innenfor forsvarssektoren



### 3.4 Bruk i media



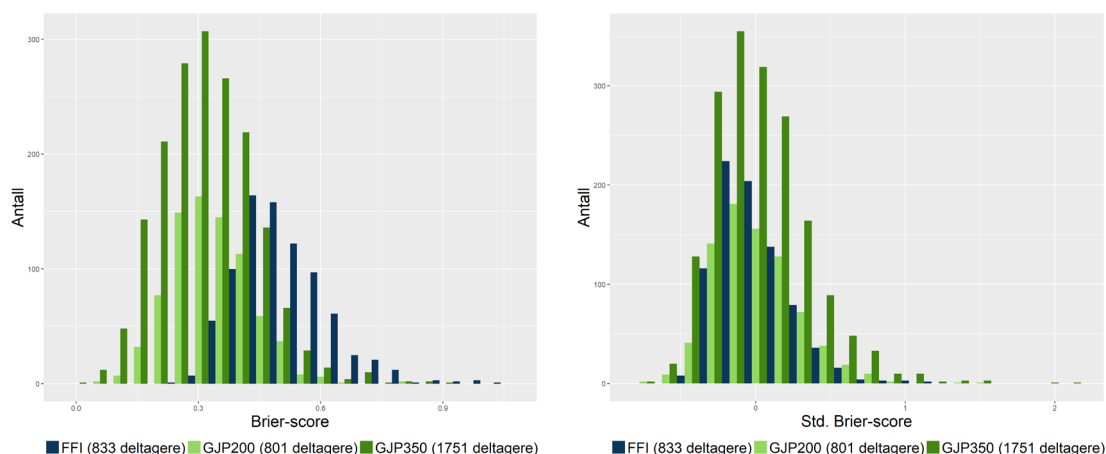
---

---

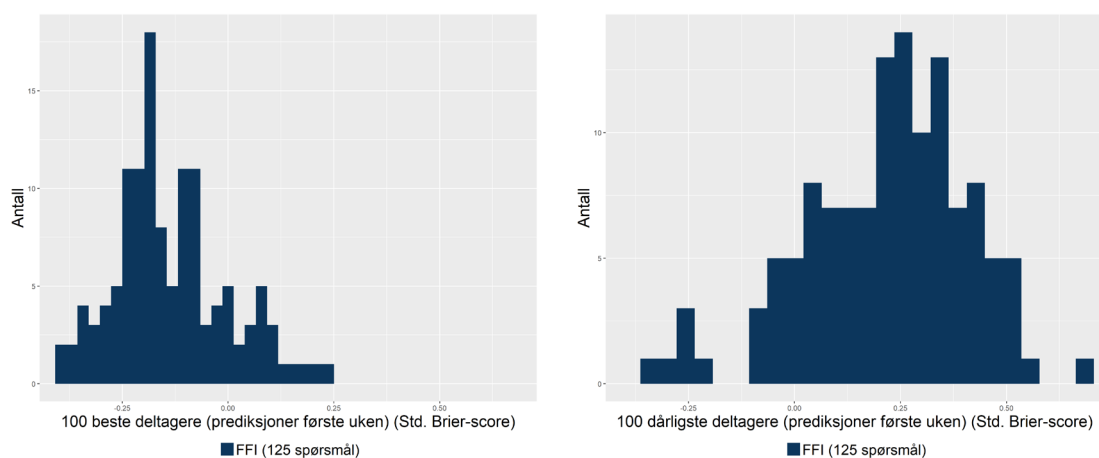
## 4 Individuelle variasjoner

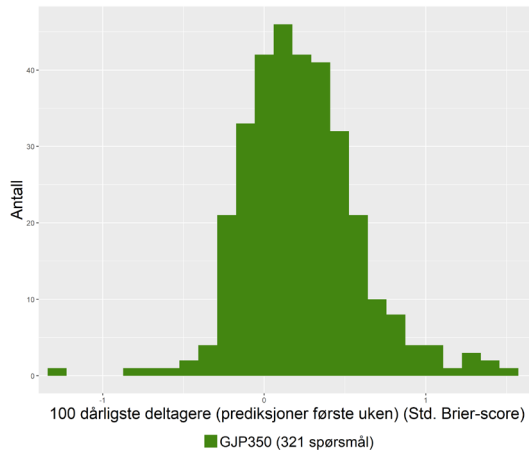
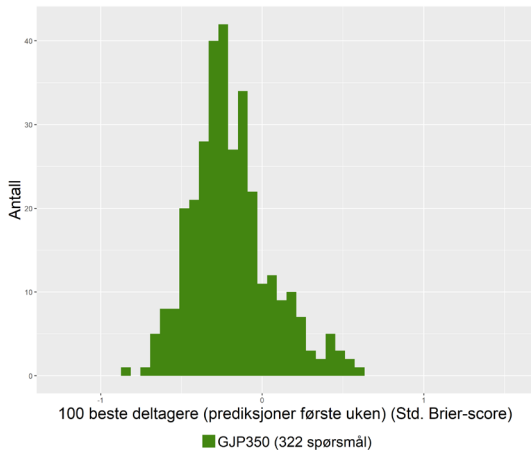
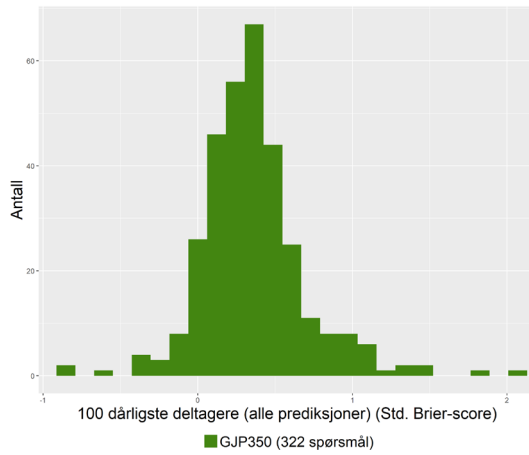
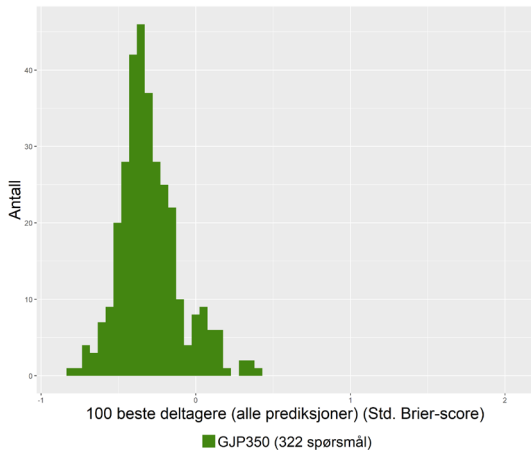
Dette delkapitlets histogrammer viser fordelingene av deltageres Brier-scores og scorene deres på alle disposisjonelle og innsatsrelaterte variabler. Først sammenlignes deltageres individuelle Brier-scores i FFIs og GJPs turneringer. Deretter skilles det mellom Brier-scorene til de 100 beste og 100 dårligste deltagerne som ble brukt til å måle treffsikkerhet over tid i begge turneringer og basert på forskjellige prediksjonstidspunkt. Deretter følger histogrammer som viser scorene på alle uavhengige variabler som ble brukt til å måle sammenhenger mellom deltageres treffsikkerhet og individuelle egenskaper. Til slutt oppgis histogrammer som viser fordelingen av Brier-scores til deltagerne som brukte eller ikke brukte forskjellige prediksjonsspesifikke teknikker og deltagerne som svarte ulikt på spørsmål om hvor godt grunnlag de følte de hadde for å svare, hvor sikre de var på prediksjonene sine og hvor taktisk de tenkte.

### 4.1 Individuell treffsikkerhet



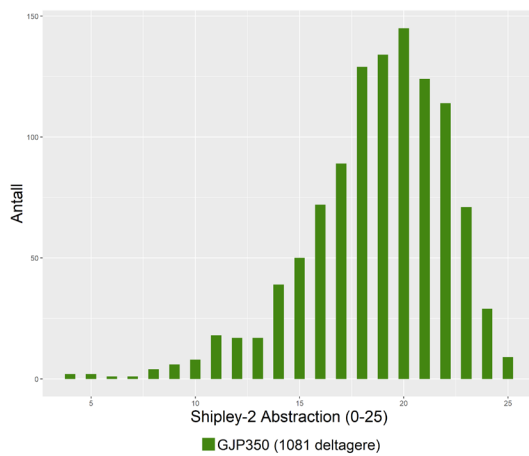
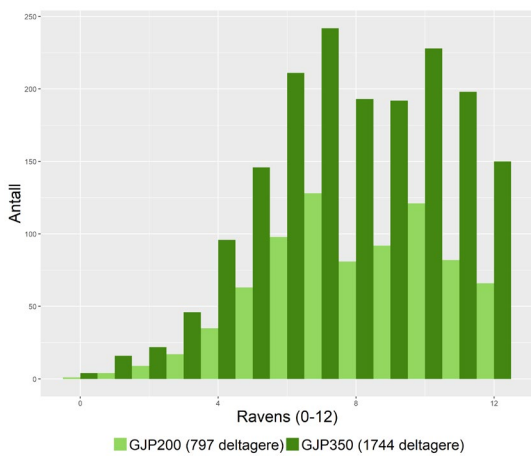
### 4.2 Treffsikkerhet over tid

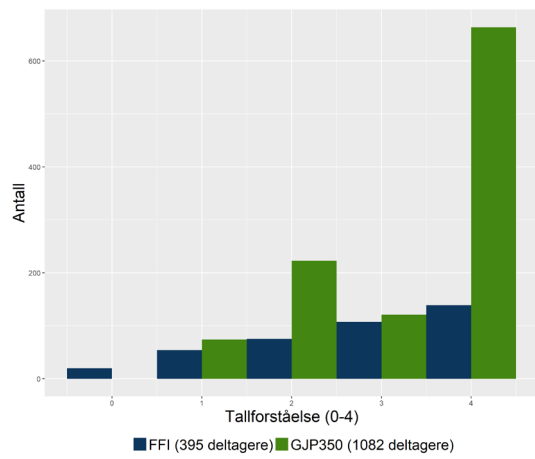
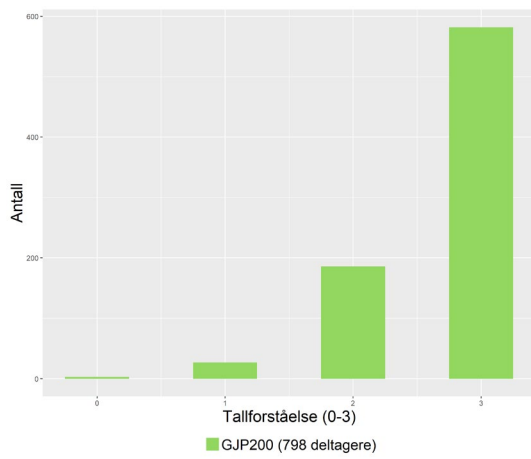
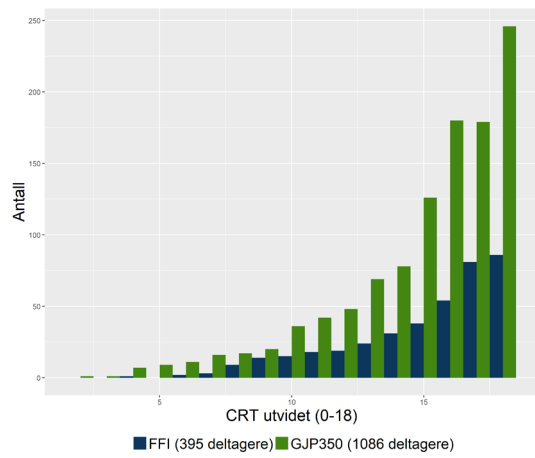
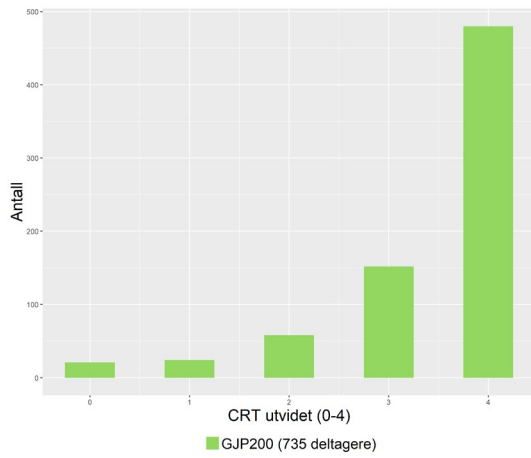
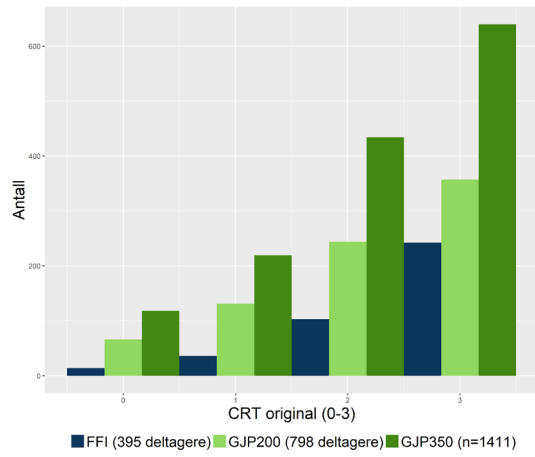
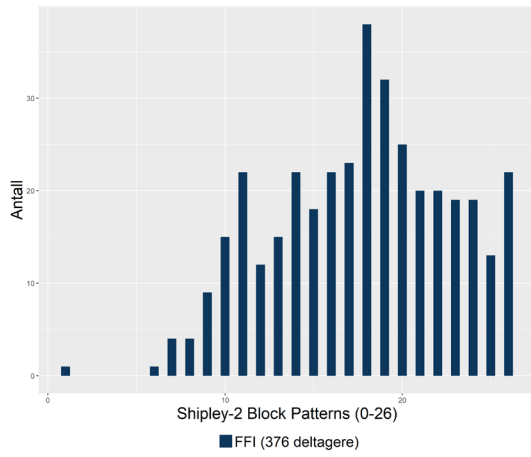




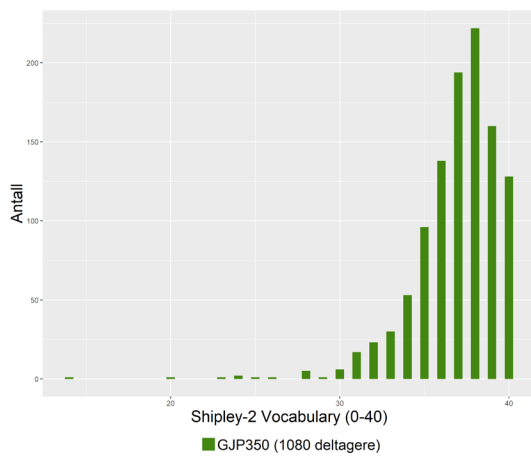
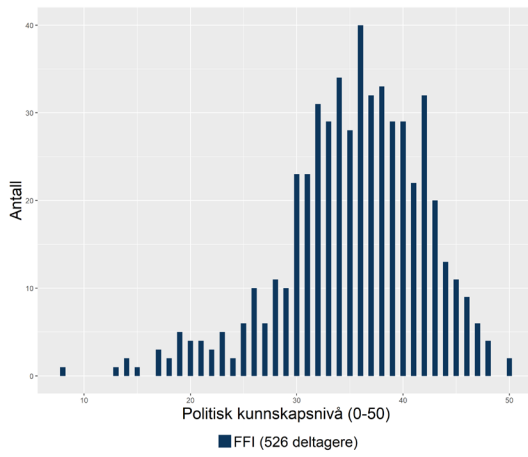
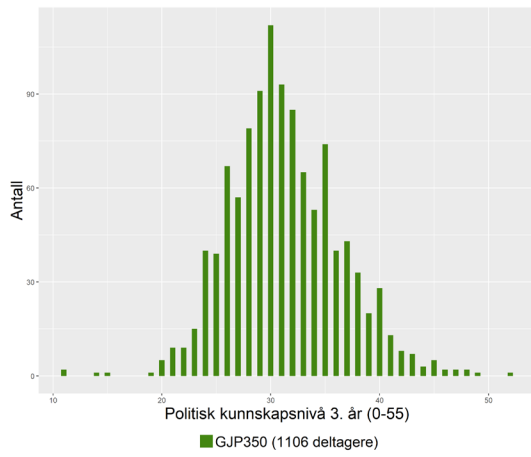
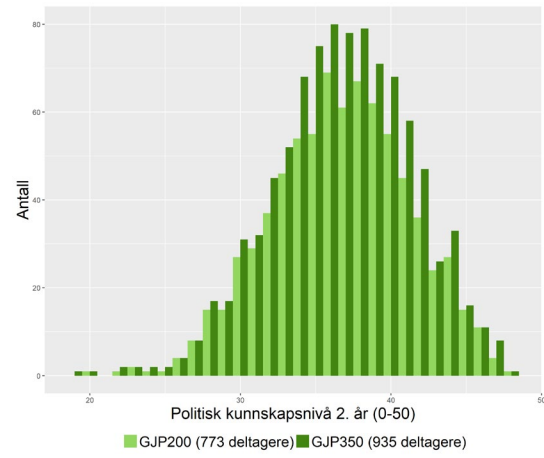
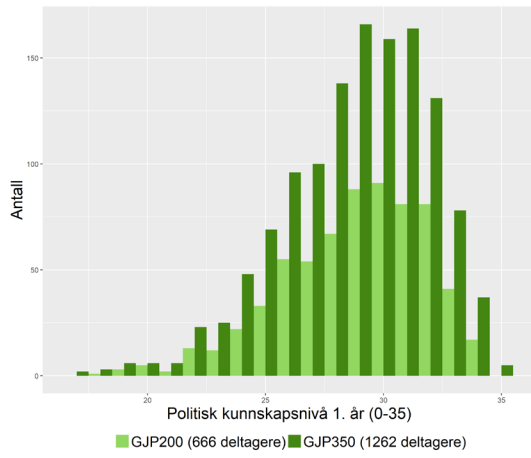
### 4.3 Individuelle egenskaper

#### 4.3.1 Kognitive evner



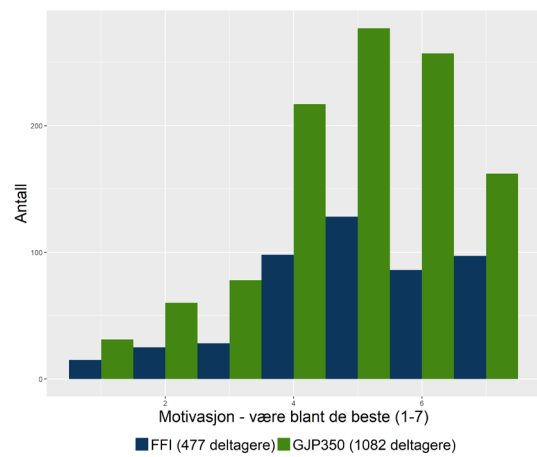
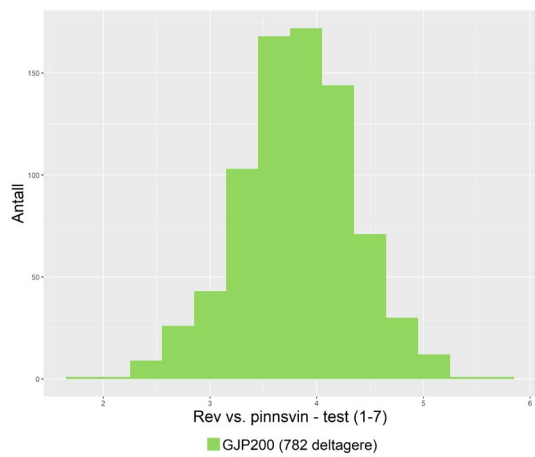
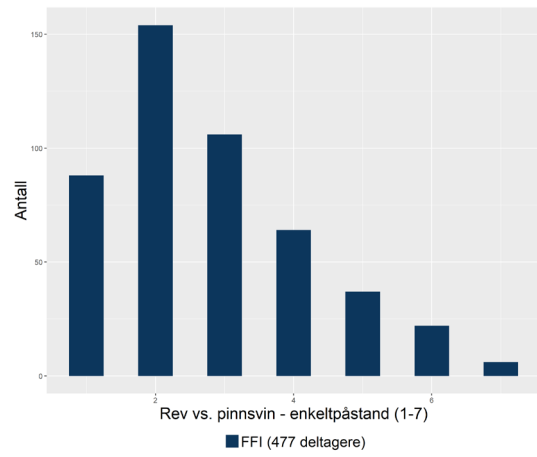
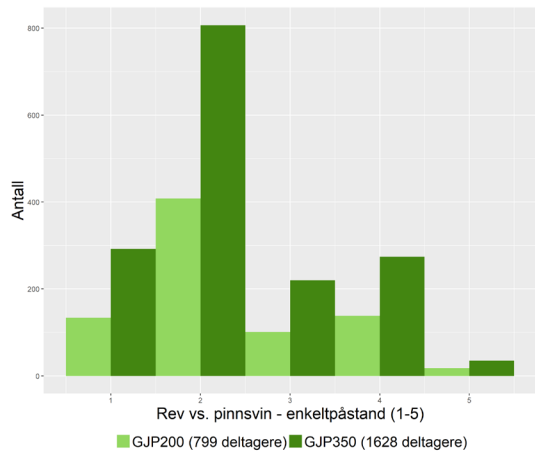
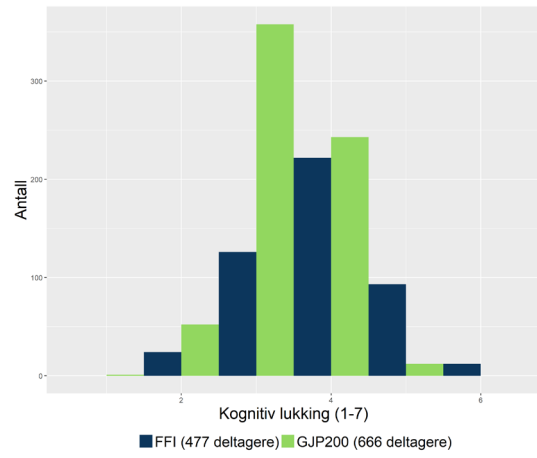
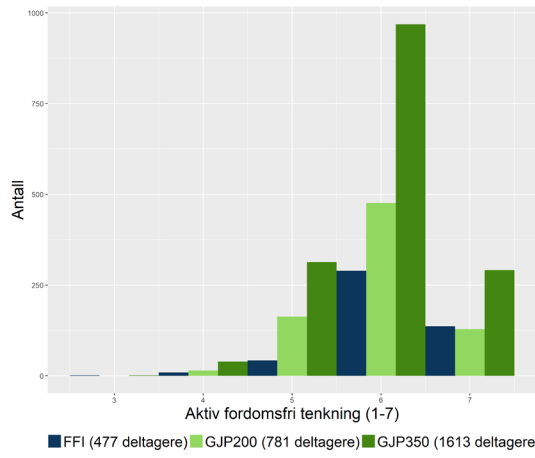


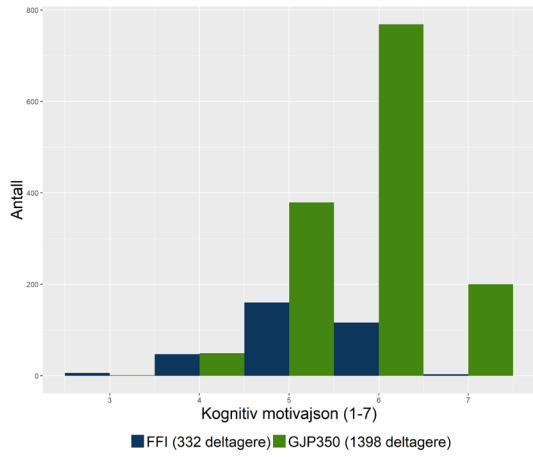
### 4.3.2 Kunnskapsnivå



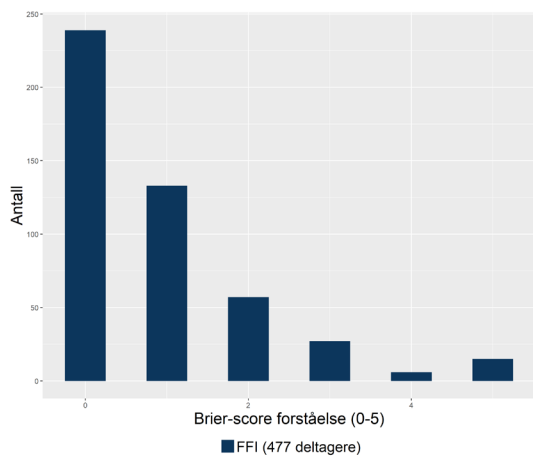
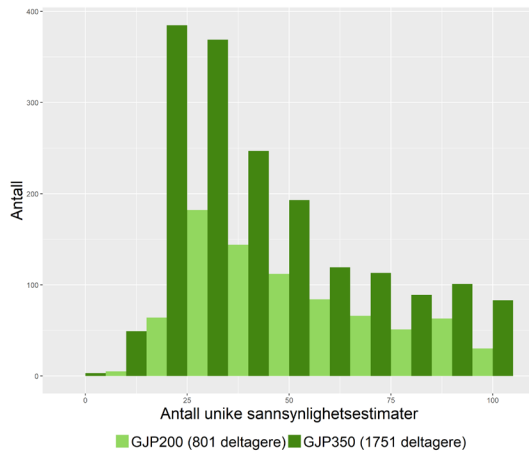
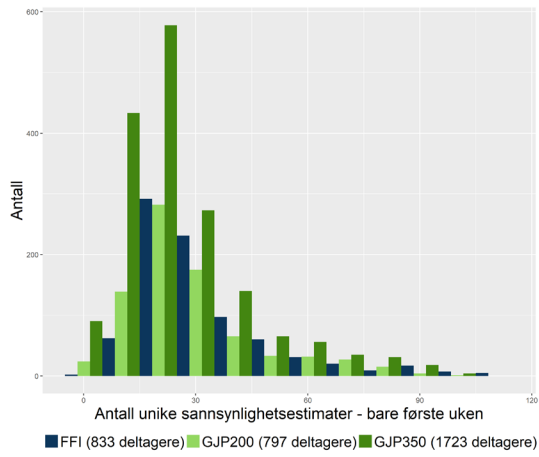


### 4.3.3 Tenkemåter

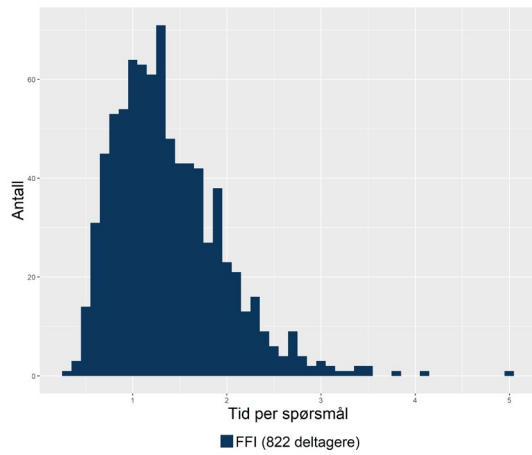
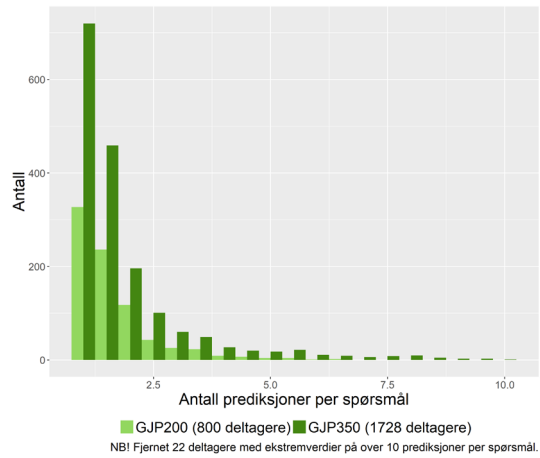
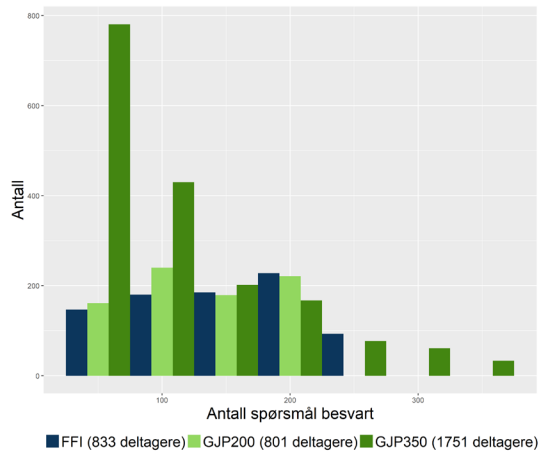




### 4.3.4 Oppgavespesifikke ferdigheter

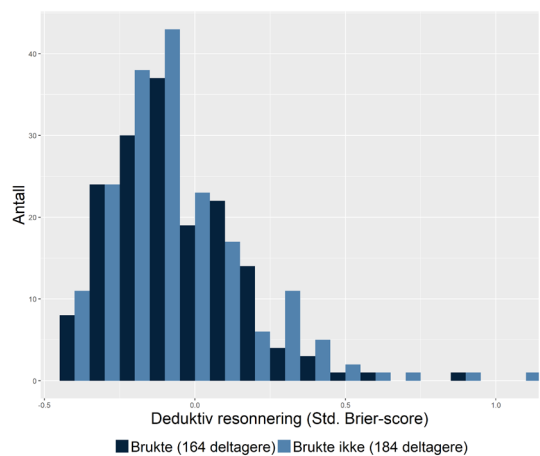
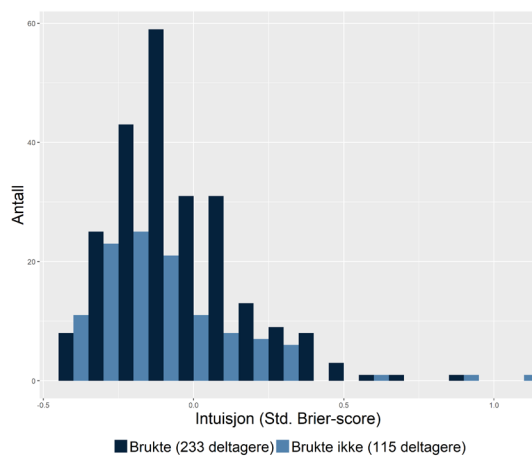


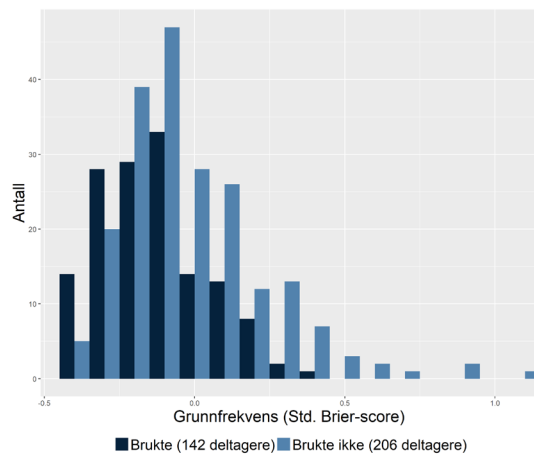
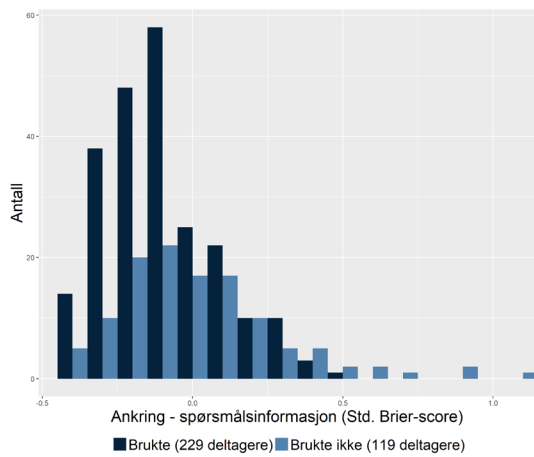
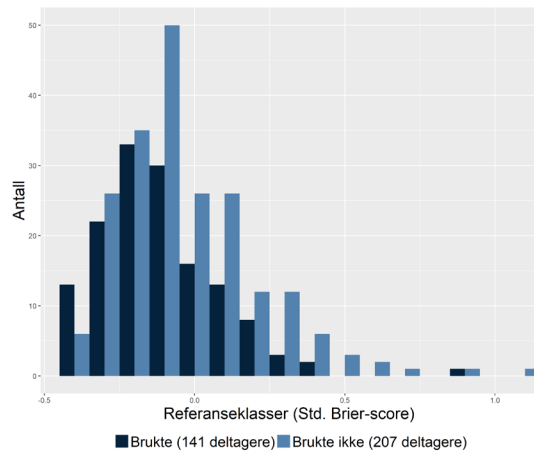
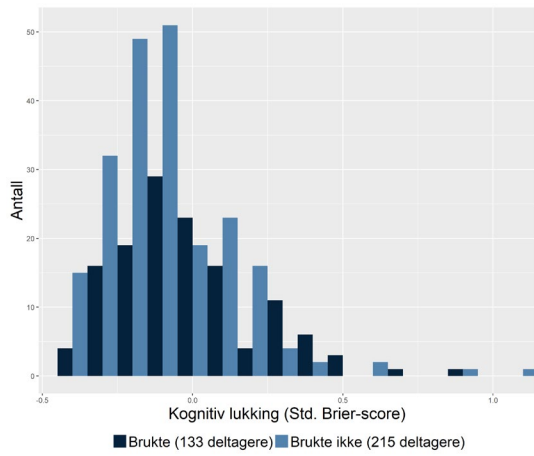
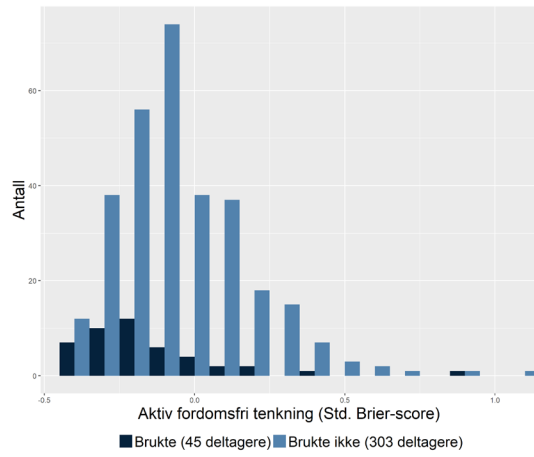
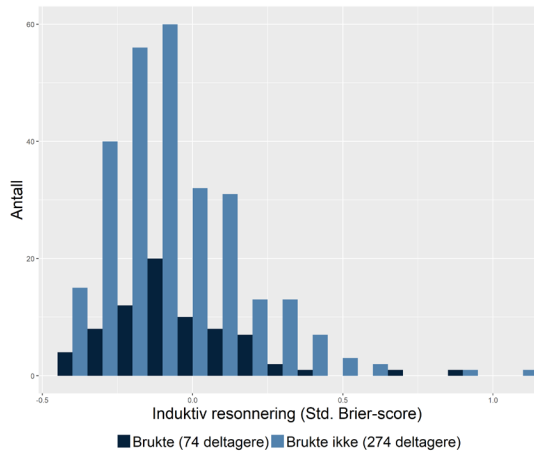
### 4.3.5 Innsats

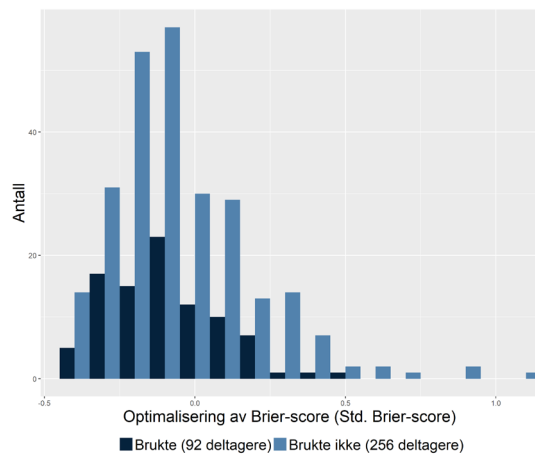
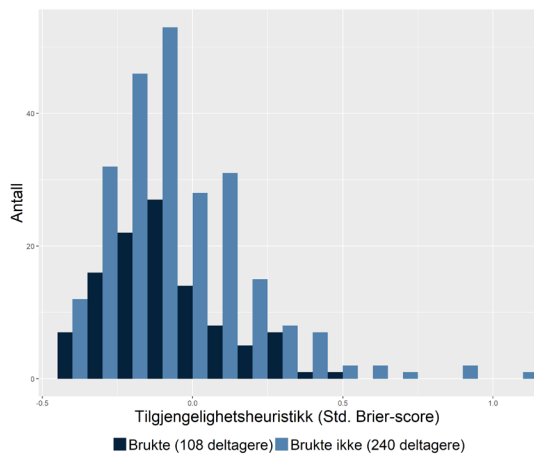
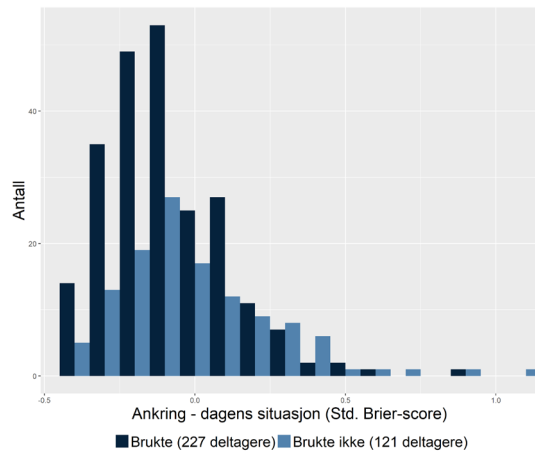
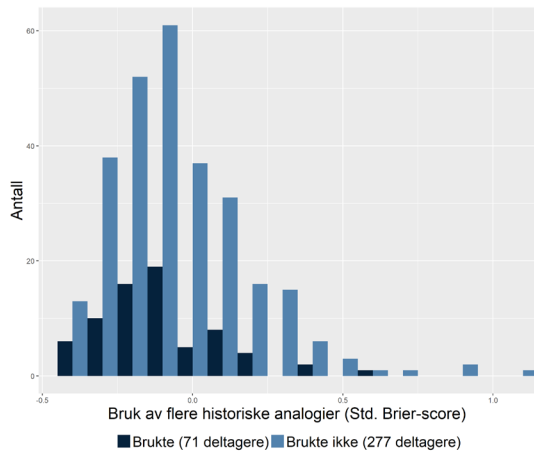
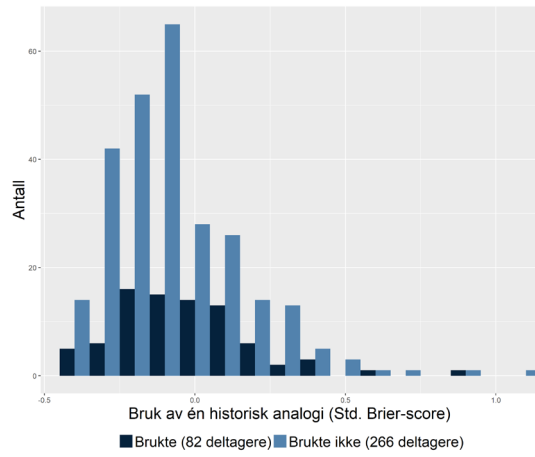
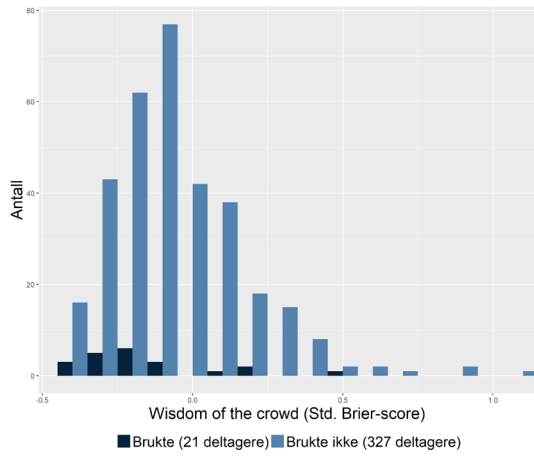


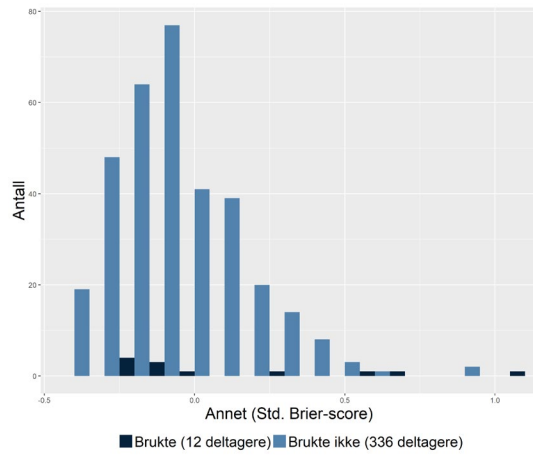
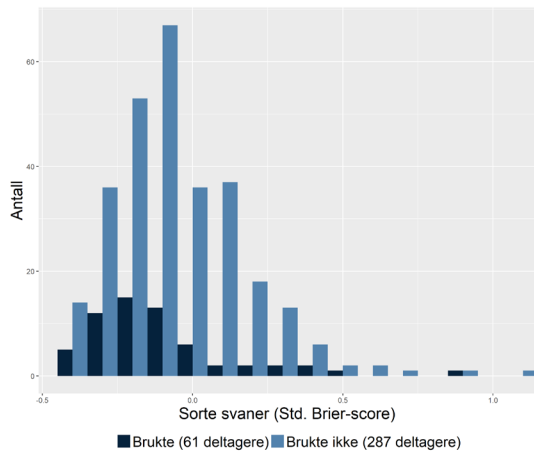
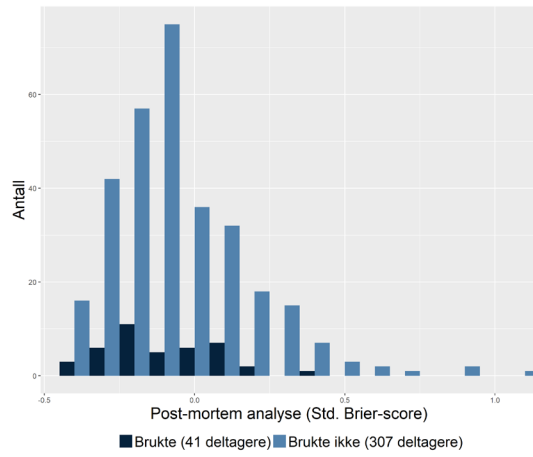
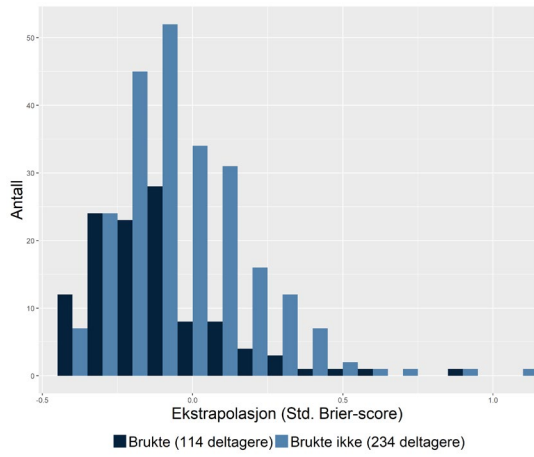
### 4.3.6 Prediksjonsspesifikke tenkemåter

#### 4.3.6.1 Teknikker

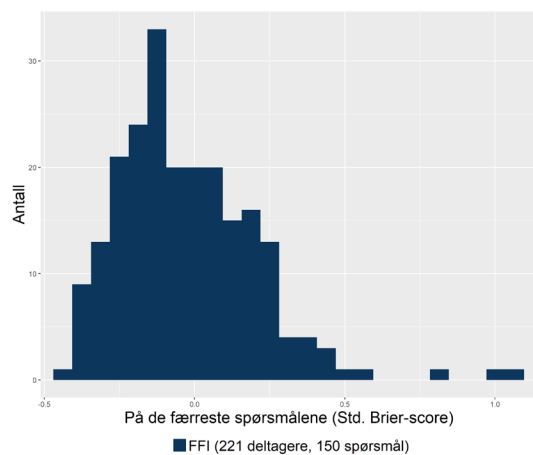
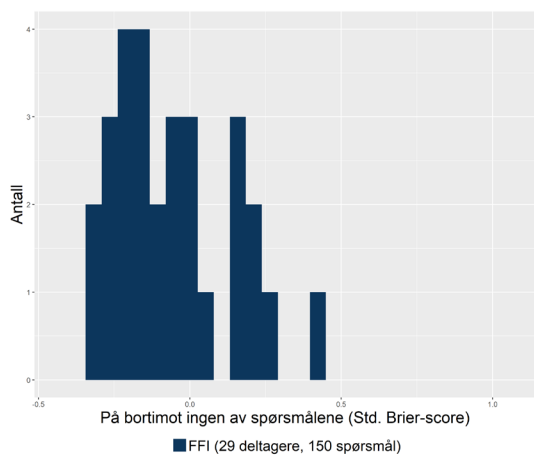


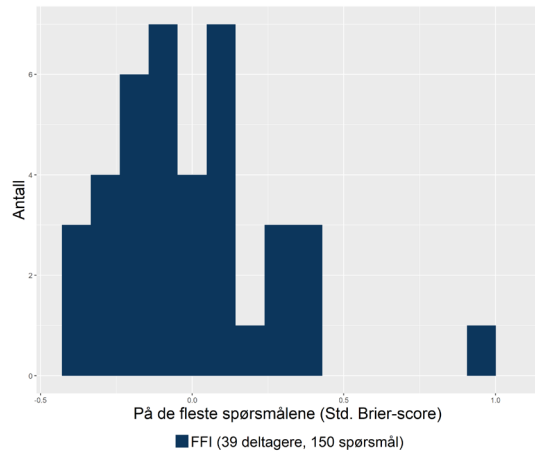
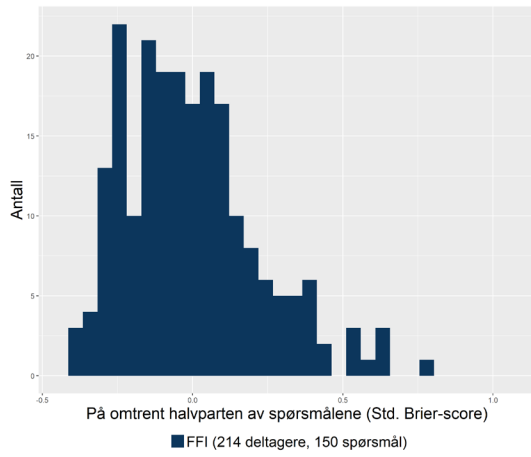




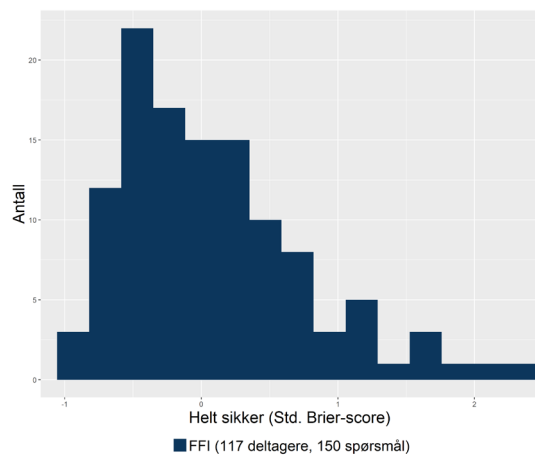
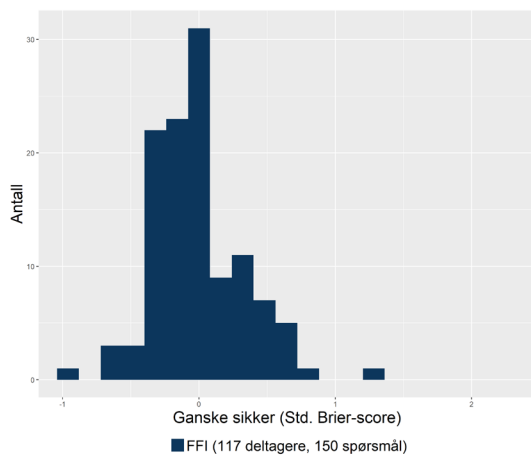
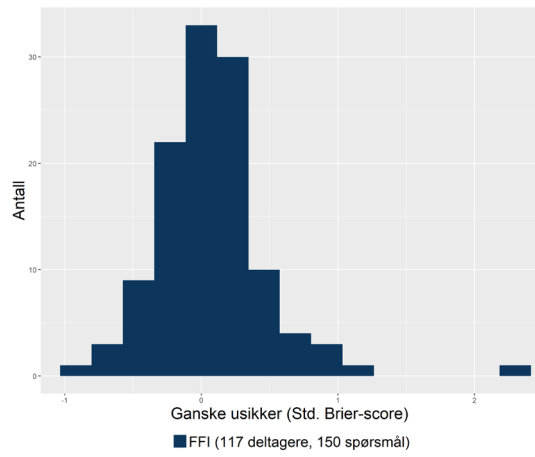
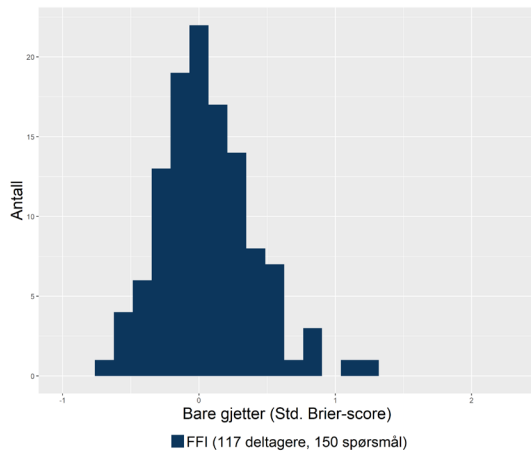


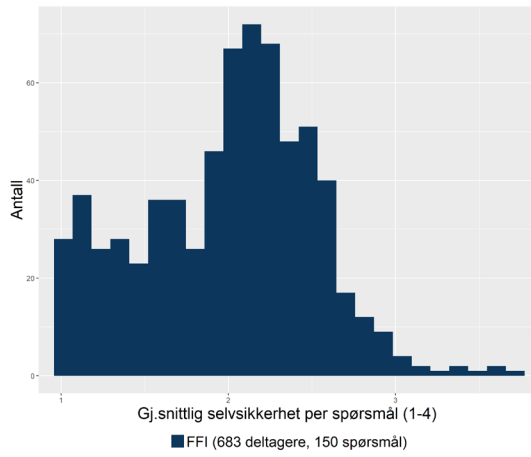
#### 4.3.6.2 Grunnlag for å svare



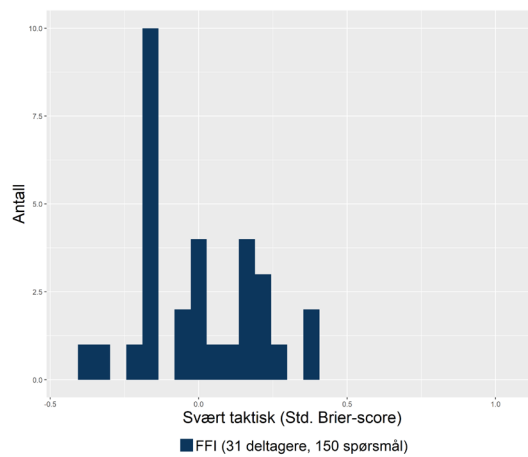
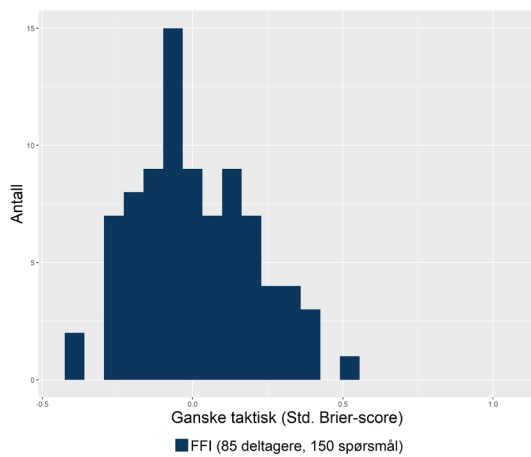
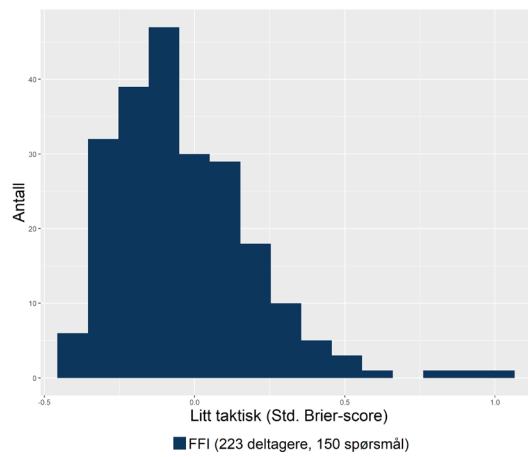
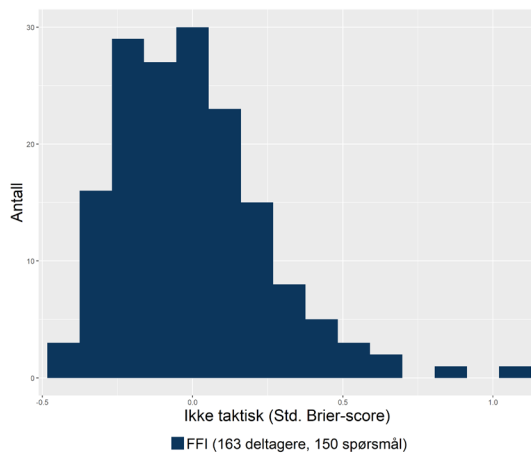


#### 4.3.6.3 *Selsikkerhet per spørsmål*



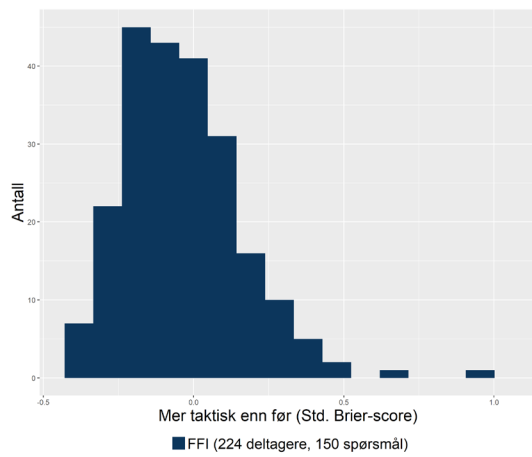
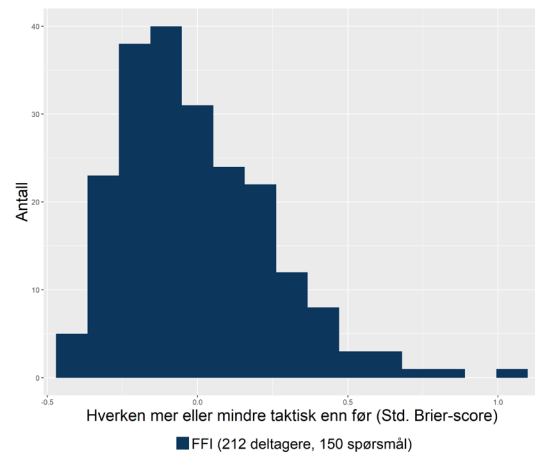
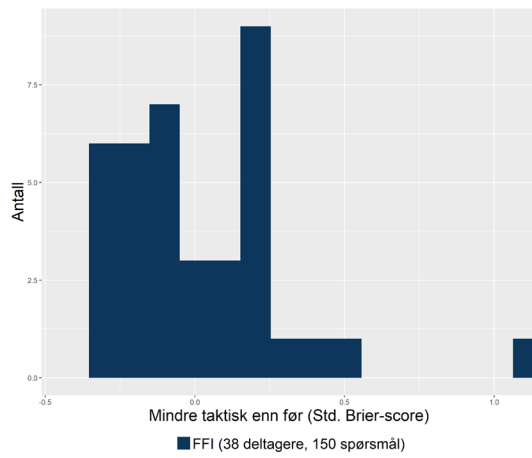


#### 4.3.6.4 Taktisk tenkning





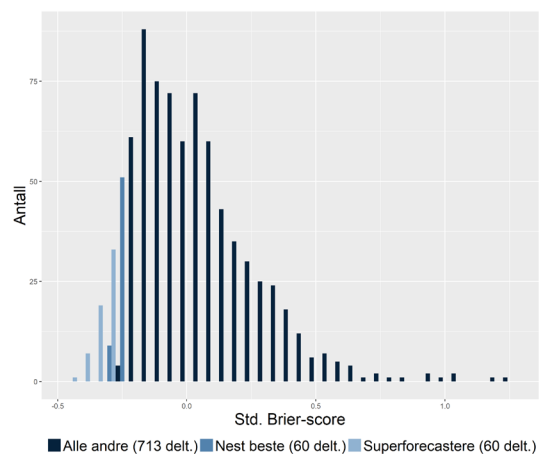
### 4.3.6.5 *Endring i taktisk tenkning underveis*



## 5 Norske superforecastere

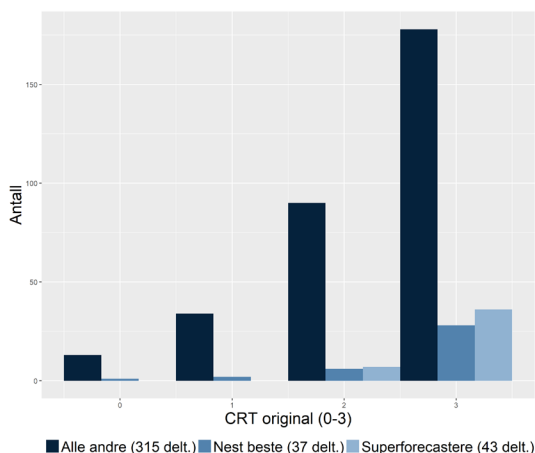
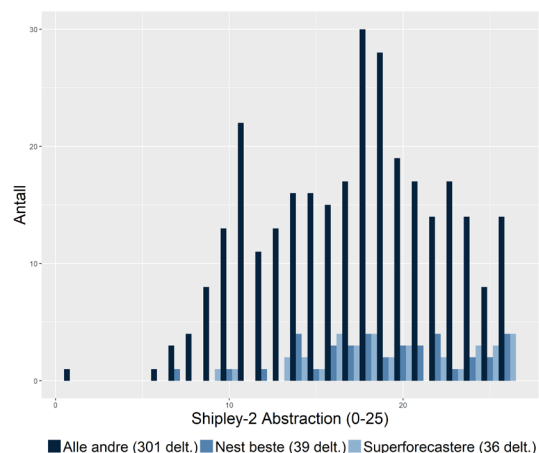
Dette delkapitlets histogrammer viser fordelingene til alle variablene som ble brukt til å analysere norske superforecastere. Først sammenlignes Brier-scorene til de 60 beste deltagerne som utgjorde superforecasterne (lys blå), de 60 nest beste deltagerne (blå) og alle andre (mørk blå). Disse scorene ble brukt til å måle om det var signifikante forskjeller mellom treffsikkerheten til de tre deltagergruppene. Deretter sammenlignes de tre deltagergruppens scores på alle disposisjonelle og innsatsrelaterte variabler, som ble brukt til å undersøke om superforecasterne skilte seg signifikant fra andre deltagere.

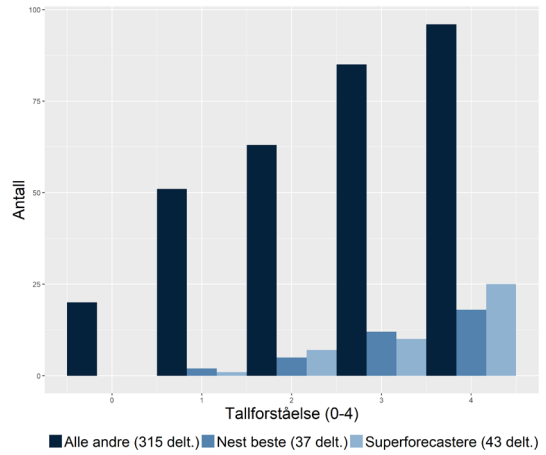
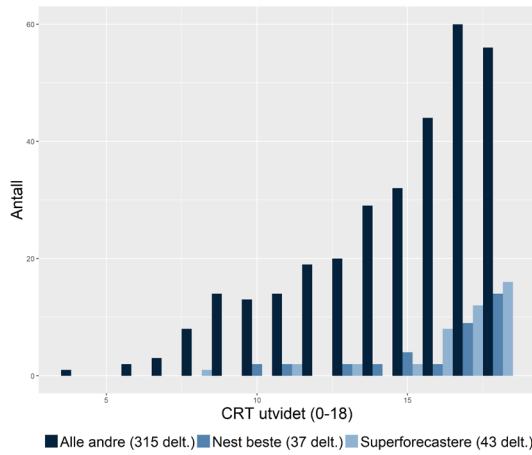
### 5.1 Gjennomsnittlig treffsikkerhet



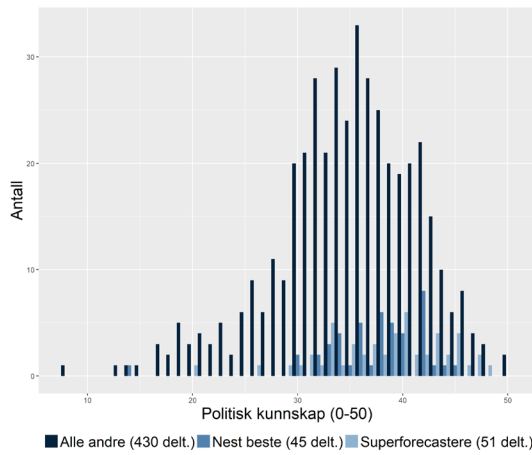
### 5.2 Individuelle egenskaper

#### 5.2.1 Kognitive evner

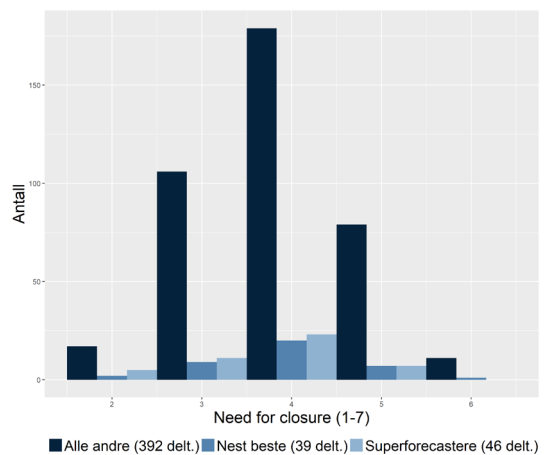
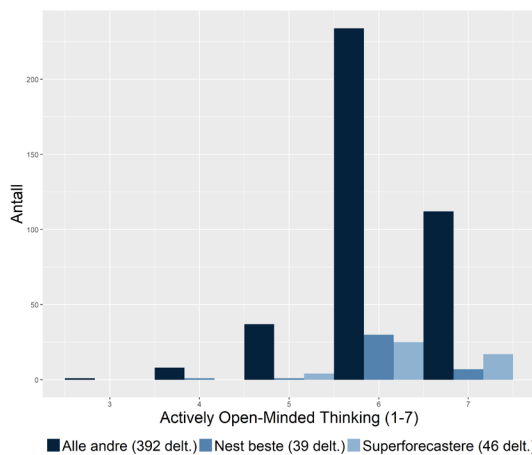


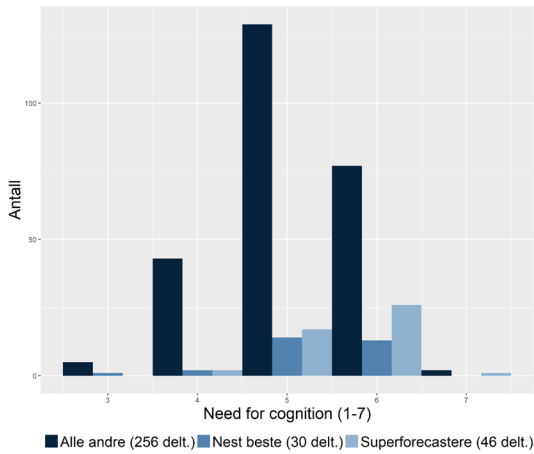
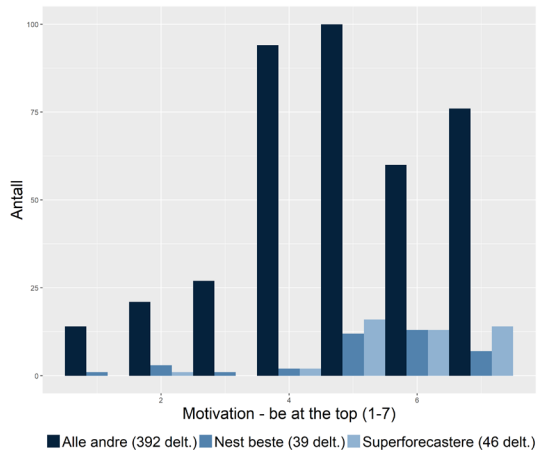
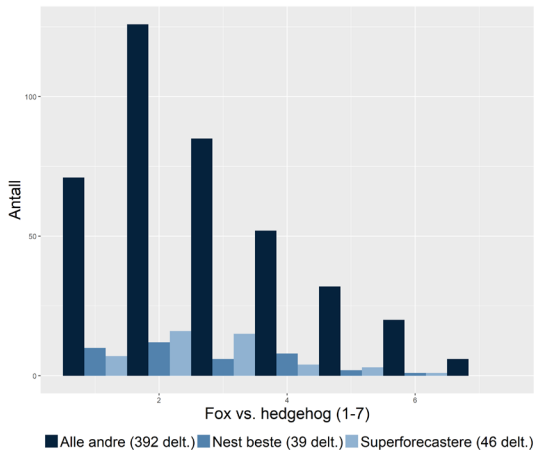


## 5.2.2 Kunnskapsnivå

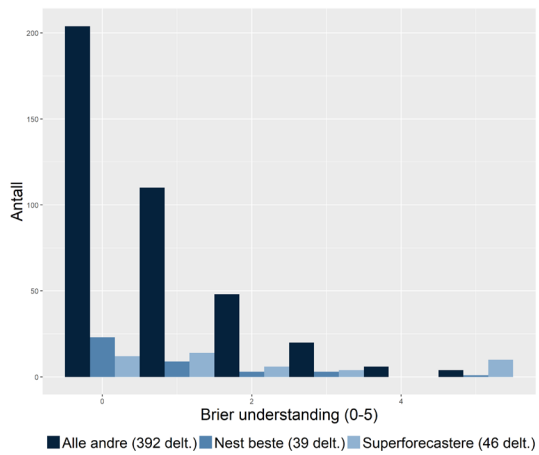
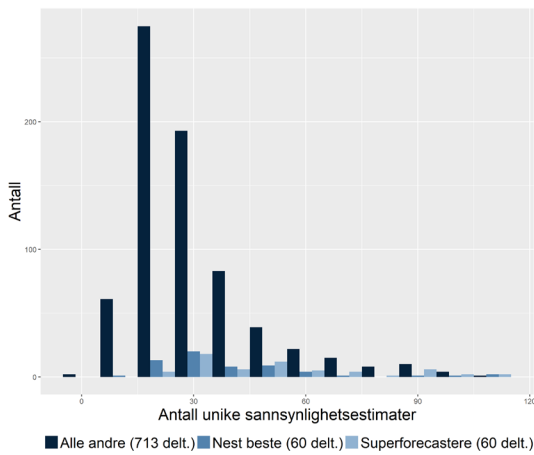


## 5.2.3 Tenkemåter

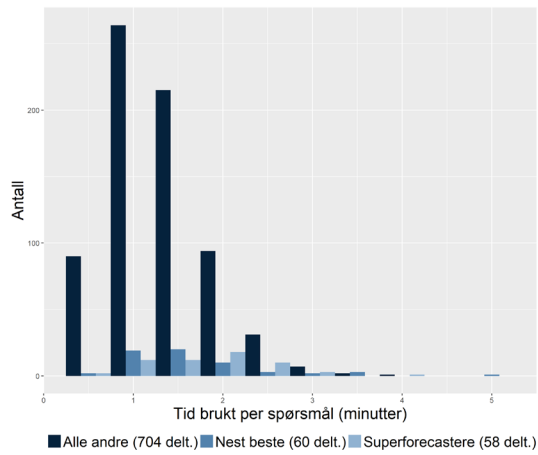
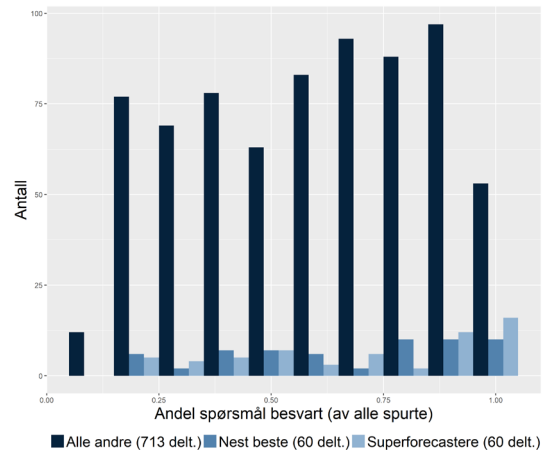
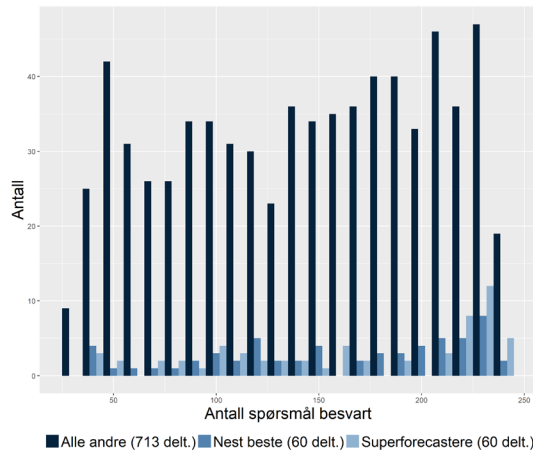




### 5.2.4 Oppgavespesifikke ferdigheter



## 5.2.5 Innsats



## Om FFI

Forsvarets forskningsinstitutt ble etablert 11. april 1946. Instituttet er organisert som et forvaltningsorgan. Med særskilte fullmakter underlagt Forsvarsdepartementet.

## FFIs formål

Forsvarets forskningsinstitutt er Forsvarets sentrale forskningsinstitusjon og har som formål å drive forskning og utvikling for Forsvarets behov. Videre er FFI rådgiver overfor Forsvarets strategiske ledelse. Spesielt skal instituttet følge opp trekk ved vitenskapelig og militærteknisk utvikling som kan påvirke forutsetningene for sikkerhetspolitikken eller forsvarsplanleggingen.

## FFIs visjon

FFI gjør kunnskap og ideer til et effektivt forsvar.

## FFIs verdier

Skapende, drivende, vidsynt og ansvarlig.

