

# Anvendelse af listeeksperimenter

*Giver smileys seriøse svar?*

*Utraditionelle spørgsmålstyper og listeeksperimenter.*

Selskab for Surveyforskning

Zoom, 15. april, 2021

Kim Mannemar Sønderskov

Institut for Statskundskab  
Aarhus Universitet  
ks@ps.au.dk



# Baggrund/Motivation

# Baggrund/Motivation

(Nogle) Generelle udfordringer ved indsamling af surveydata:

# Baggrund/Motivation

(Nogle) Generelle udfordringer ved indsamling af surveydata:

- Ekstern validitet
  - Repræsentativitet
    - (Systematisk) Survey non-response
    - (Systematisk) Item non-response
  - Målingsvaliditet
    - Response bias, fx pga. social desirability bias, dårlig hukommelse (= målefejl)
    - Tvedydige spørgsmål, tvetydige svarkategorier
- Reliabilitet
  - Støj, fx pga. uopmærksomhed eller dårlig hukommelse

# Baggrund/Motivation

(Nogle) Generelle udfordringer ved indsamling af surveydata:

- Ekstern validitet
  - Repræsentativitet
    - (Systematisk) Survey non-response
    - (Systematisk) Item non-response
  - Målingsvaliditet
    - Response bias, fx pga. social desirability bias, dårlig hukommelse (= målefejl)
    - Tvetydige spørgsmål, tvetydige svarkategorier
- Reliabilitet
  - Støj, fx pga. uopmærksomhed eller dårlig hukommelse

Liste-eksperimentet som løsning?

# Baggrund/Motivation

(Nogle) Generelle udfordringer ved indsamling af surveydata:

- Ekstern validitet
  - Repræsentativitet
    - (Systematisk) Survey non-response
    - (Systematisk) Item non-response
  - Målingsvaliditet
    - Response bias, fx pga. **social desirability bias**, dårlig hukommelse (= målefejl)
    - Tvetydige spørgsmål, tvetydige svarkategorier
- Reliabilitet
  - Støj, fx pga. uopmærksomhed eller dårlig hukommelse

Liste-eksperimentet som løsning?

# Response bias er et alvorligt problem

## Spørgsmål om kriminel adfærd (Preisendörfer og Wolter, 2014)

Et eksempel fra Tyskland, 2009-2010:

Spørgeskemainterview med tidligere dømte lovovertrædere (typisk butikstyre, spritbilister, og skattesnydere).

Andel af dømte, der indrømmede at de er straffede:

# Response bias er et alvorligt problem

## Spørgsmål om kriminel adfærd (Preisendörfer og Wolter, 2014)

Et eksempel fra Tyskland, 2009-2010:

Spørgeskemainterview med tidligere dømte lovovertrædere (typisk butikstyre, spritbilister, og skattesnydere).

Andel af dømte, der indrømmede at de er straffede:

63%



# Response bias er et alvorligt problem

## Spørgsmål om kriminel adfærd (Preisendörfer og Wolter, 2014)

Et eksempel fra Tyskland, 2009-2010:

Spørgeskemainterview med tidligere dømte lovovertrædere (typisk butikstyre, spritbilister, og skattesnydere).

Andel af dømte, der indrømmede at de er straffede:

63% → 37% lyver

## (Nogle) Mulige survey-baserede løsninger på response bias

## (Nogle) Mulige survey-baserede løsninger på response bias

(A) Spørg eksplicit, men skab tryghed

## (Nogle) Mulige survey-baserede løsninger på response bias

(A) Spørg eksplicit, men skab tryghed

- Undgå personlig kontakt

## (Nogle) Mulige survey-baserede løsninger på response bias

- (A) Spørg eksplicit, men skab tryghed
- Undgå personlig kontakt

Et eksempel fra Tyskland, 2009-2010:

Spørgeskemainterview med tidligere dømte lovovertrædere (typisk butikstyre, spritbilister, og skattesnydere).

Andel af dømte, der indrømmede at de er straffede:

63% → 37% lyver

## (Nogle) Mulige survey-baserede løsninger på response bias

(A) Spørg eksplicit, men skab tryghed

- Undgå personlig kontakt

Et eksempel fra Tyskland, 2009-2010:

Spørgeskemainterview med tidligere dømte lovovertrædere (typisk butikstyre, spritbilister, og skattesnydere).

Andel af dømte, der indrømmede at de er straffede:

63% → 37% lyver

Instrument-eksperiment:

Mail-survey: 67%

Face-to-face: 58%

Forskellen mellem metoder er statistisk signifikant ( $p < 0,05$ )

## (Nogle) Mulige survey-baserede løsninger på response bias

- (A) Spørg eksplicit, men skab tryghed
  - Undgå personlig kontakt

## (Nogle) Mulige survey-baserede løsninger på response bias

(A) Spørg eksplicit, men skab tryghed

- Undgå personlig kontakt
- Brug proxy subjects: "Mange mennesker har svært ved at forholde sig politik"



## (Nogle) Mulige survey-baserede løsninger på response bias

(A) Spørg eksplicit, men skab tryghed

- Undgå personlig kontakt
- Brug proxy subjects: "Mange mennesker har svært ved at forholde sig politik"

(B) Spørg ultra-eksplicit

- Fx: Nævn navnene på de, af dine venner, der har indvandrerbaggrund (Smith, 2002)

## (Nogle) Mulige survey-baserede løsninger på response bias

### (A) Spørg eksplicit, men skab tryghed

- Undgå personlig kontakt
- Brug proxy subjects: "Mange mennesker har svært ved at forholde sig politik"

### (B) Spørg ultra-eksplicit

- Fx: Nævn navnene på de, af dine venner, der har indvandrerbaggrund (Smith, 2002)

### (C) Spørg implicit

- IAT: Implicit Association Test (Greenwald, McGhee og Schwartz, 1998)  
Prøv fx: <https://implicit.harvard.edu/implicit/denmark/>
- AMP: Affect Misattribution Procedure  
Fx: Tucker m.fl., 2018
- Se flere muligheder i Gawronski og De Houwer (2014)

## (Nogle) Mulige survey-baserede løsninger på response bias

### (A) Spørg eksplicit, men skab tryghed

- Undgå personlig kontakt
- Brug proxy subjects: "Mange mennesker har svært ved at forholde sig politik"

### (B) Spørg ultra-eksplicit

- Fx: Nævn navnene på de, af dine venner, der har indvandrerbaggrund (Smith, 2002)

### (C) Spørg implicit

- IAT: Implicit Association Test (Greenwald, McGhee og Schwartz, 1998)  
Prøv fx: <https://implicit.harvard.edu/implicit/denmark/>
- AMP: Affect Misattribution Procedure  
Fx: Tucker m.fl., 2018
- Se flere muligheder i Gawronski og De Houwer (2014)

### (D) Brug et **liste-eksperiment** (Spørg eksplicit, men skab tryghed)

- Effektiv metode til reduktion af social desirability bias, der nemt lader sig implementere i surveys (Blair, Imai og Lyall, 2014)

# Det grundlæggende setup i et liste-eksperiment

# Det grundlæggende setup i et liste-eksperiment

Inddel respondenterne tilfældigt i 2 (eller flere) grupper.

# Det grundlæggende setup i et liste-eksperiment

Inddel respondenterne tilfældigt i 2 (eller flere) grupper. Stil dem et spørgsmål á la følgende "Folk har forskellige holdninger til forskellige ting. Her følger en liste med forskellige ting, som nogle mennesker ikke ville bryde sig om. Angiv hvor mange af tingene, du ikke ville bryde dig om. Du skal altså ikke angive, hvilke af tingene du ikke ville bryde dig om, men blot et tal mellem 0 og [4/5]"

Liste:

Kontrolgruppe	Stimulusgruppe
At lytte til musik	At lytte til musik
At blive ringet op af en telefonsælger	At blive ringet op af en telefonsælger
At sende skattekroner til Afrika	At sende skattekroner til Afrika
At gøre grænsekontrollen permanent	At gøre grænsekontrollen permanent
-	At få [X] som nabo

Kilde: Hjorth, Sønderskov & Dinesen, 2019

# Det grundlæggende setup i et liste-eksperiment

## Fortolkning

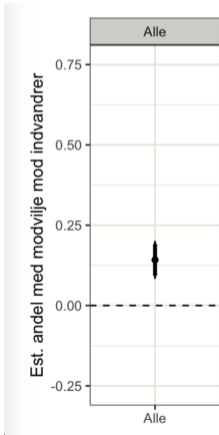
Kontrolgruppe	Stimulusgruppe
At lytte til musik	At lytte til musik
At blive ringet op af en telefonsælger	At blive ringet op af en telefonsælger
At sende skattekroner til Afrika	At sende skattekroner til Afrika
At gøre grænsekontrollen permanent	At gøre grænsekontrollen permanent
-	At få [X] som nabo

Forskellen mellem de to gruppers gennemsnit kan - givet den tilfældige inddeling og under visse antagelser - fortolkes som andelen i stimulusgruppen, der har valgt item 5 (det kritiske item). Givet tilfældig tildeling gælder dette resultat for hele stikprøven. Standardfejlen for forskel i gennemsnit kan bruges til statistisk inferens til en population. (Kuklinski, Cobb og Gilens, 1997)

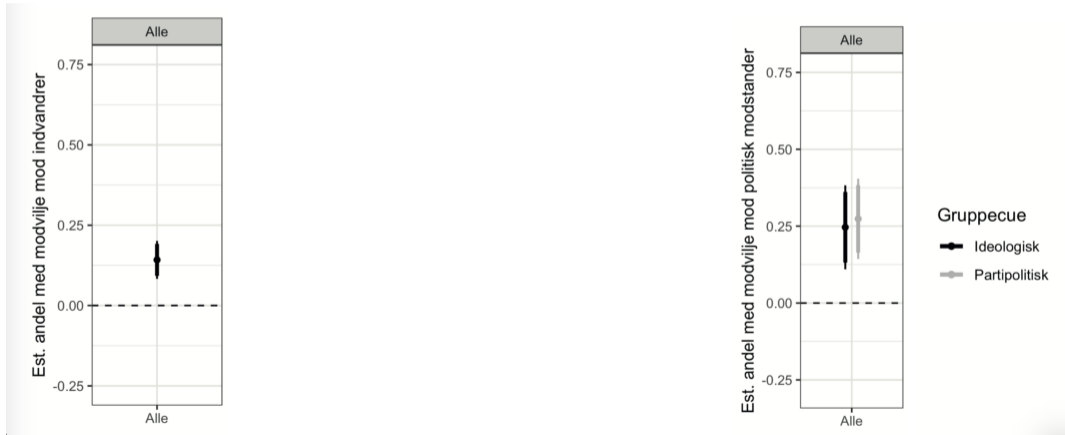
# Eksempler på resultater fra Danmark (Hjorth, Sønderkov og Dinesen, 2019)



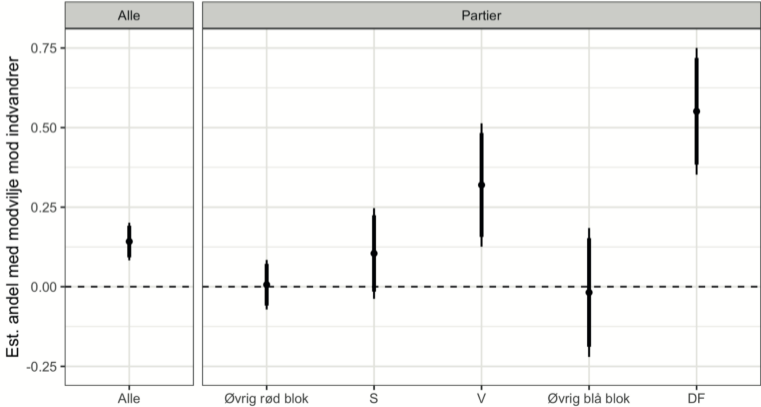
# Eksempler på resultater fra Danmark (Hjorth, Sønderskov og Dinesen, 2019)



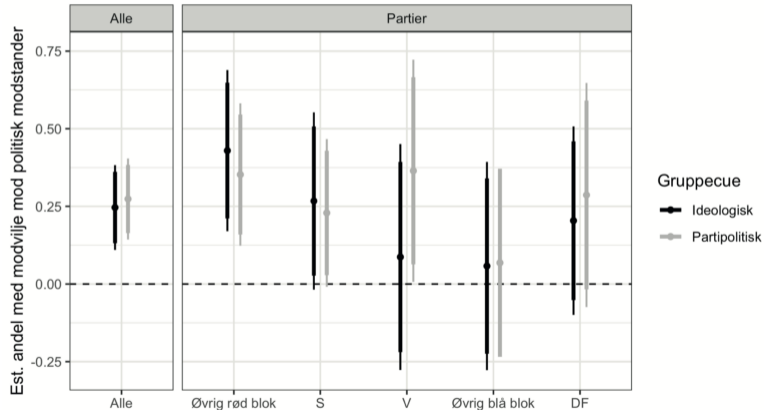
# Eksempler på resultater fra Danmark (Hjorth, Sønderkov og Dinesen, 2019)



# Eksempler på resultater fra Danmark (Hjorth, Sønderskov og Dinesen, 2019): Sammenligning af grupper - Holdninger til indvandrere



# Eksempler på resultater fra Danmark (Hjorth, Sønderskov og Dinesen, 2019): Sammenligning af grupper - Holdninger til politiske modstandere



## Overvejelser om de (4) øvrige items "kontrol-items"

Kontrolgruppe	Stimulusgruppe
At lytte til musik	At lytte til musik
At blive ringet op af en telefonsælger	At blive ringet op af en telefonsælger
At sende skattekroner til Afrika	At sende skattekroner til Afrika
At gøre grænsekontrollen permanent	At gøre grænsekontrollen permanent
-	At få [X] som nabo

## Overvejelser om de (4) øvrige items "kontrol-items"

Ideal: Antallet af valgte kontrol-items skal have minimal varians på tværs af respondenter og være upåvirket af gulv- og loftseffekter.

Kontrolgruppe	Stimulusgruppe
At lytte til musik	At lytte til musik
At blive ringet op af en telefonsælger	At blive ringet op af en telefonsælger
At sende skattekroner til Afrika	At sende skattekroner til Afrika
At gøre grænsekontrollen permanent	At gøre grænsekontrollen permanent
-	At få [X] som nabo

## Overvejelser om de (4) øvrige items "kontrol-items"

Ideal: Antallet af valgte kontrol-items skal have minimal varians på tværs af respondenter og være upåvirket af gulv- og loftseffekter.

→ Ønske: Alle vælger 2 kontrol-items

Kontrolgruppe	Stimulusgruppe
At lytte til musik	At lytte til musik
At blive ringet op af en telefonsælger	At blive ringet op af en telefonsælger
At sende skattekrone til Afrika	At sende skattekrone til Afrika
At gøre grænsekontrollen permanent	At gøre grænsekontrollen permanent
-	At få [X] som nabo

# Overvejelser om de (4) øvrige items "kontrol-items"

Ideal: Antallet af valgte kontrol-items skal have minimal varians på tværs af respondenter og være upåvirket af gulv- og loftseffekter.

→ Ønske: Alle vælger 2 kontrol-items

→ Anbefaling: (Blair og Imai, 2012)

Item-type	Kontrolgruppe	Stimulusgruppe
Lavfrekvent	At lytte til musik	At lytte til musik
Højfrekvent	At blive ringet op af en telefonsælger	At blive ringet op af en telefonsælger
Negativt korreleret 1	At sende skattekrone til Afrika	At sende skattekrone til Afrika
Negativt korreleret 2	At gøre grænsekontrollen permanent	At gøre grænsekontrollen permanent
Følsomt	-	At få [X] som nabo



# Udbygninger

- Estimation af Response bias/Social desirability bias:  
Man kan som sidegevindst analysere omfanget af Response bias/Social desirability bias. Inkluder et eksplicit spørgsmål efter listen i kontrolgruppeversionen af spørgeskemaet (Gilens, Sniderman og Kuklinski, 1998)

# Udbygninger

- Estimation af Response bias/Social desirability bias:  
Man kan som sidegevindst analysere omfanget af Response bias/Social desirability bias. Inkluder et eksplicit spørgsmål efter listen i kontrolgruppeversionen af spørgeskemaet (Gilens, Sniderman og Kuklinski, 1998)
- Øget efficiens og mulighed for at estimere sensitive holdninger på respondent-niveau (Blair og Imai, 2012)

# Antagelser

# Antagelser

Antagelser:

- Fravær af design effekter: respondenternes valg blandt kontrol-items er urelateret til treatment-status.

# Antagelser

## Antagelser:

- Fravær af design effekter: respondenternes valg blandt kontrol-items er urelateret til treatment-status.  
De må gerne lyve om deres holdning til kontrol-items
- Fravær af løgn ifht. det sensitive item

# Opsamling

## Liste-eksperimentet

...er en effektiv metode til reduktion af social desirability bias, der nemt lader sig implementere i surveys (Blair, Imai og Lyall, 2014)

# Opsamling

## Liste-eksperimentet

...er en effektiv metode til reduktion af social desirability bias, der nemt lader sig implementere i surveys (Blair, Imai og Lyall, 2014)

...kan også reducere non-response: "In addition, direct questioning of sensitive items generally leads to greater nonresponse rates. For example, in the Multi-Investigator Survey discussed in Section 2.2 where the sensitive question about affirmative action is asked both directly and indirectly, the nonresponse rate is 6.5% for the direct questioning format and 0% for the list experiment." (Blair og Imai, 2012)

# Opsamling

## Liste-eksperimentet

...er en effektiv metode til reduktion af social desirability bias, der nemt lader sig implementere i surveys (Blair, Imai og Lyall, 2014)

...kan også reducere non-response: "In addition, direct questioning of sensitive items generally leads to greater nonresponse rates. For example, in the Multi-Investigator Survey discussed in Section 2.2 where the sensitive question about affirmative action is asked both directly and indirectly, the nonresponse rate is 6.5% for the direct questioning format and 0% for the list experiment." (Blair og Imai, 2012)

Tak for jeres opmærksomhed!



# References I



Blair, Graeme og Kosuke Imai (2012). "Statistical analysis of list experiments". I: *Political Analysis* 20.1, s. 47–77.



Blair, Graeme, Kosuke Imai og Jason Lyall (2014). "Comparing and combining list and endorsement experiments: Evidence from Afghanistan". I: *American Journal of Political Science* 58.4, s. 1043–1063.



Gawronski, Bertram og Janx De Houwer (2014). "Implicit measures in social and personality psychology". I: *Handbook of research methods in social and personality psychology*. Udg. af Harry T. Reis og Charles M. Jud. Cambridge: Cambridge University Press, s. 283–310.



Gilens, Martin, Paul M Sniderman og James H Kuklinski (1998). "Affirmative action and the politics of realignment". I: *British Journal of Political Science* 28.1, s. 159–183.



Greenwald, Anthony G, Debbie E McGhee og Jordan LK Schwartz (1998). "Measuring individual differences in implicit cognition: the implicit association test." I: *Journal of personality and social psychology* 74.6, s. 1464.



Hjorth, Frederik, Kim Mannemar Sønderskov og Peter Thisted Dinesen (2019). "Affektiv Polarisering i Danmark: Et Listeesperiment om Social Distance til Politiske Modstandere". I: *Økonomi og Politik* 92.3, s. 31–40.



Kuklinski, James H, Michael D Cobb og Martin Gilens (1997). "Racial attitudes and the "New South"". I: *The Journal of Politics* 59.2, s. 323–349.



Preisendörfer, Peter og Felix Wolter (2014). "Who is telling the truth? A validation study on determinants of response behavior in surveys". I: *Public Opinion Quarterly* 78.1, s. 126–146.



Smith, Tom W (2002). "Measuring inter-racial friendships". I: *Social Science Research* 31.4, s. 576–593.

# References II



Tucker, Raymond P m.fl. (2018). "Implicit Association with Suicide as Measured by the Suicide Affect Misattribution Procedure (S-AMP) Predicts Suicide Ideation". I: *Suicide and Life-Threatening Behavior* 48.6, s. 720–731.