(D)IY – Data Integration Project

Data cleaning & showcases
Who *claps* the most?

Name duplicates

- If deputies have similar names, this might be hard for the algorithms to interprete correctly.
- Cleaning step: Search for possible name similarities, then treat them manually
- Result:
  - 88 dagerous name similarities found
  - (only) 2 similarities in deputies of the 19th Bundestag:
    - Michael Grosse-Brömer and Michael Groß
    - Martina Renner and Martin Renner
- Manual correction: Remove unnecessary deputy-event links:
  - If Michael Grosse-Blömer and Michael Groß both appear on an event, it is more likely that Michael Grosse-Blömer was the mentioned deputy
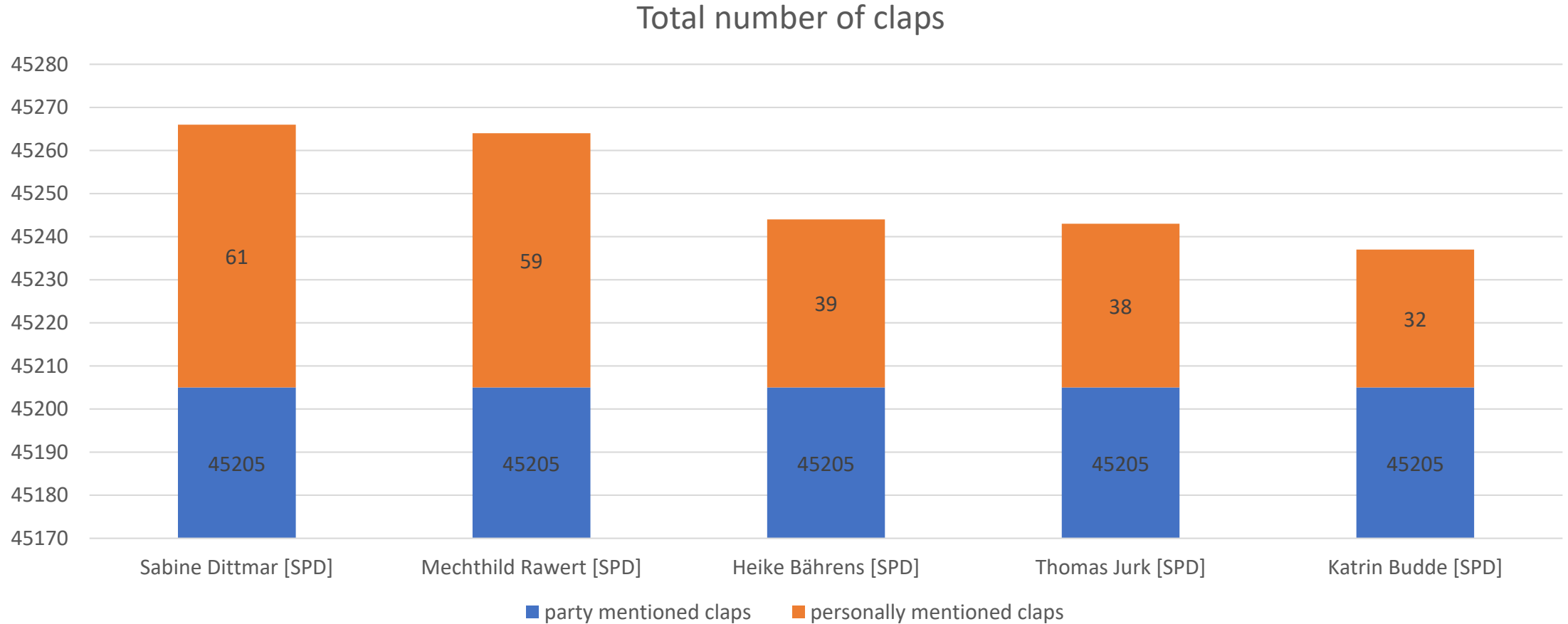
Deputies with changing attributes

- Deputies change their party during a parliament period
- Some deputies quit or start their parliament job during a parliament period
- Some events may occur during a speech, but not related to that speech, but another event
→Everything represented in the original data, but not in the integrated data (because of the complexity)
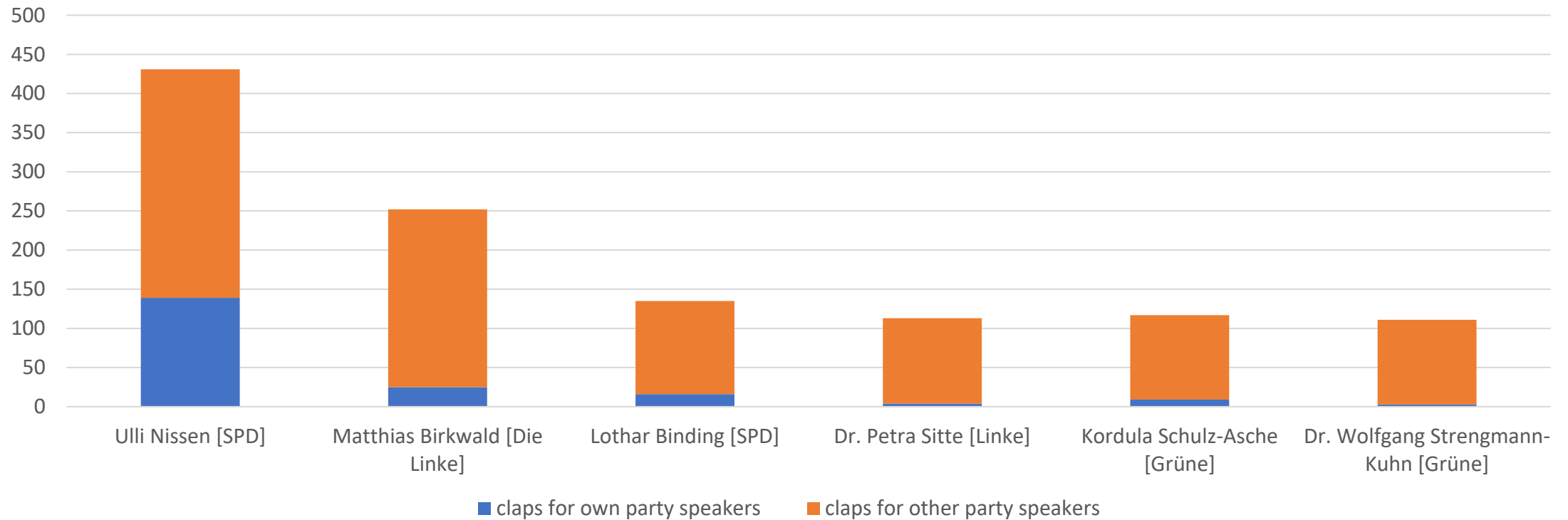→No guarantee for correctness!!

Showcases

- Obviously data integration enabled showcases at all

- Profits from data cleaning: not that relevant. Data cleaning, in this case, means doing the integration step more and more carefully.

→ No comparison between cleaned and uncleaned data possible.
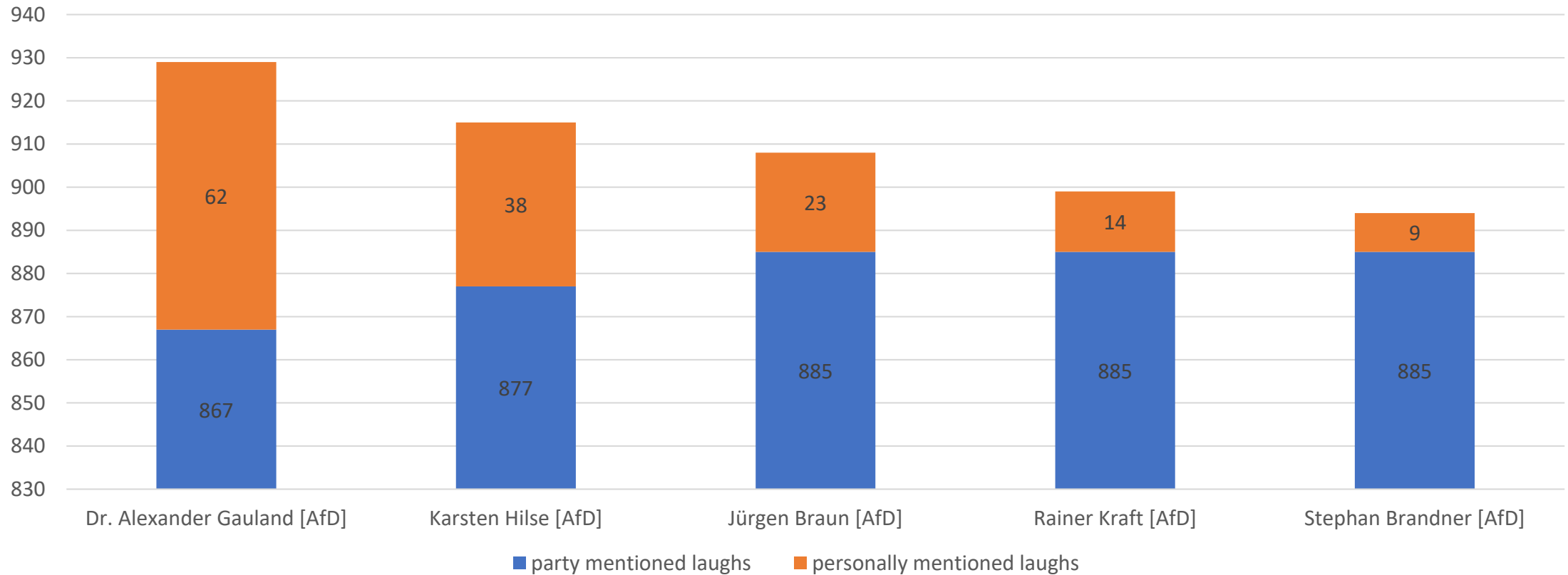
Who claps the most?

Total number of claps

Who claps for other parties?

Deputies with highest total number of personally mentioned claps for other other party's speaker

- claps for own party speakers
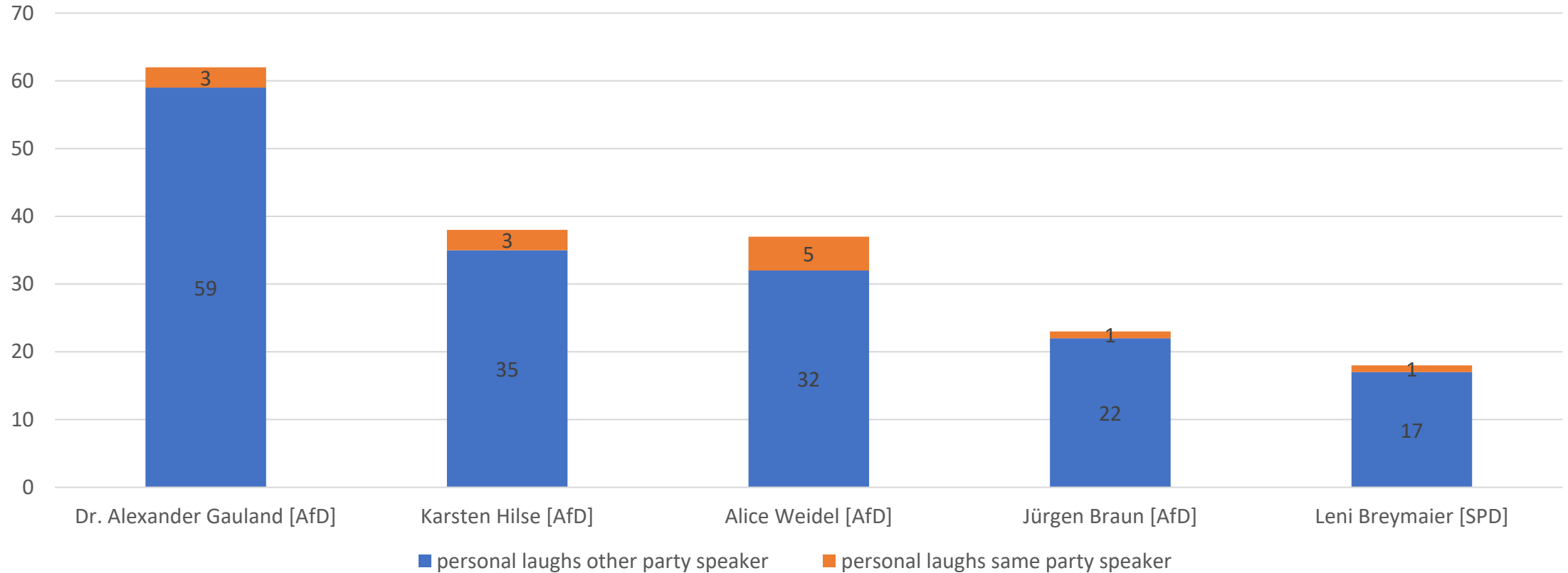- claps for other party speakers

Who laughs the most?

Deputies with highest total number of laughs

Who laughs the most alone?

Deputies with highest total number of personal mentioned laughs

Legend: ■ personal laughs other party speaker ■ personal laughs same party speaker

Dr. Alexander Gauland [AfD]: 59, 3
Karsten Hilse [AfD]: 35, 3
Alice Weidel [AfD]: 32, 5
Jürgen Braun [AfD]: 22, 1
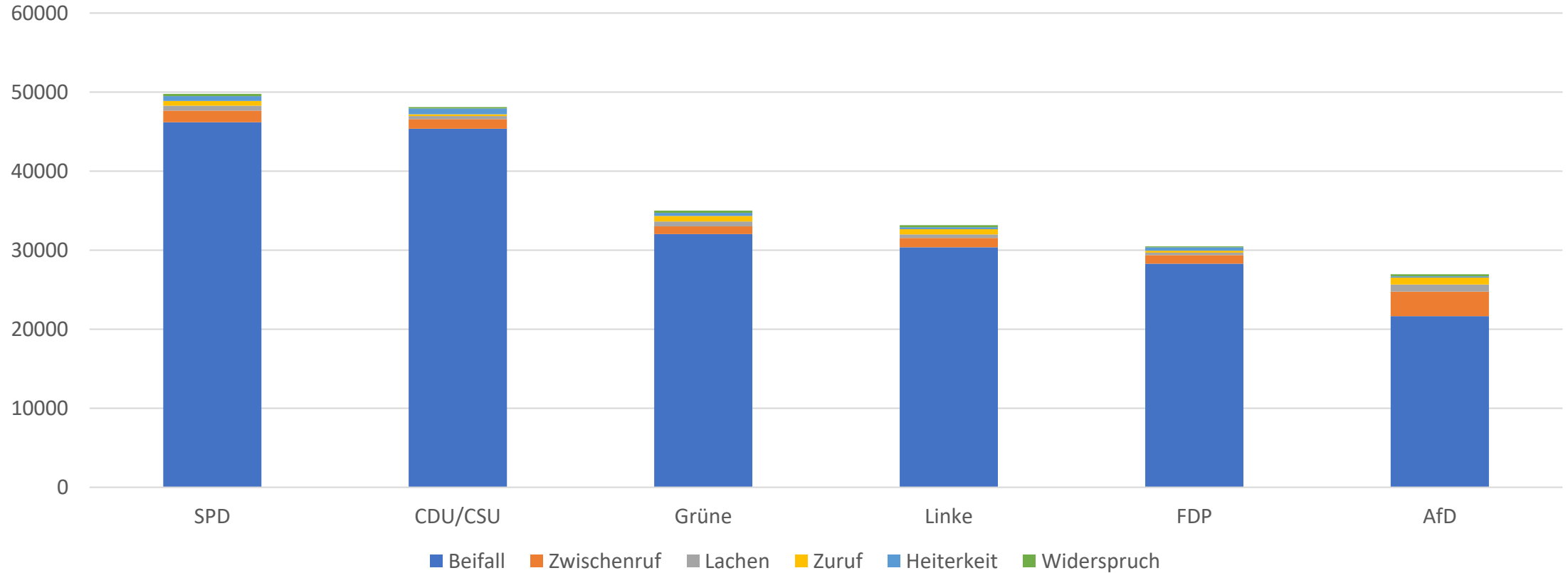Leni Breymaier [SPD]: 17, 1

Who disturbs the most?

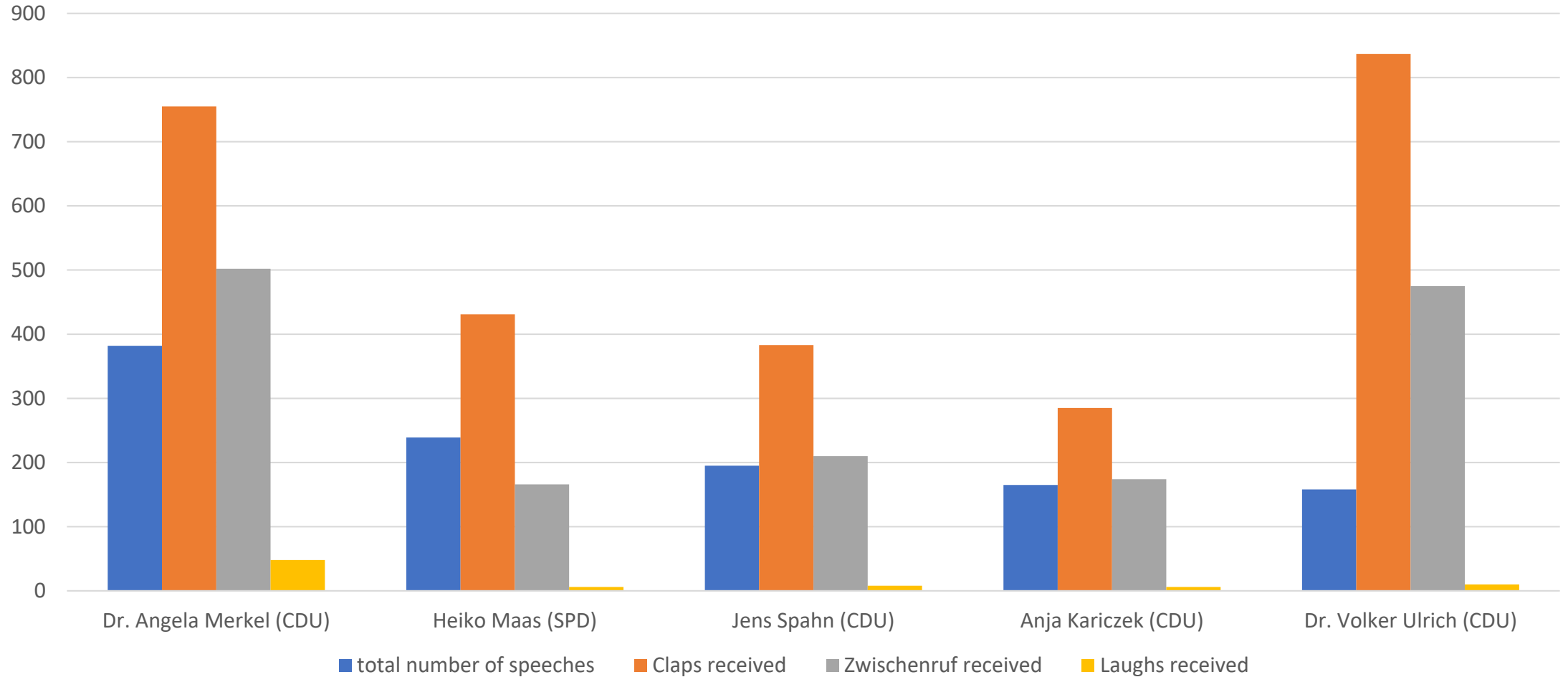Deputies with highest total number of personal mentioned disturbing events

Which is the most active party?
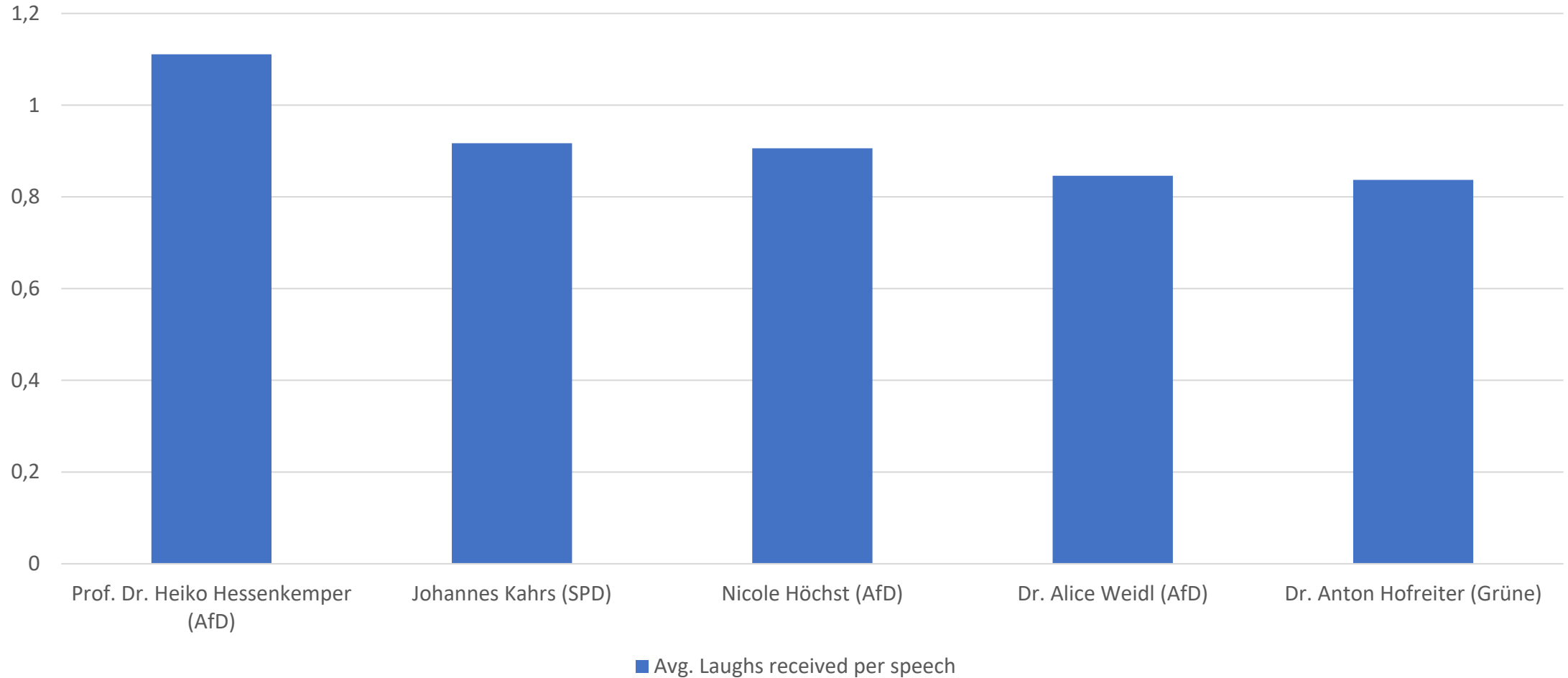

Total number of events per party

Who speaks the most?

Deputies with highest total number of speeches hold

Who is the „funniest"?

Deputies with highest average number of received laughs per speech

Who makes everyone clap?

Deputies with highest average number of received claps per speech

Andrea Nahles (SPD)	Sarah Wagenknecht (Linke)	Christian Lindner (FDP)	Dietmar Bartsch (Linke)	Ralph Brinkhaus (CD)

■ Avg. Laughs received per speech