Codebook — Mind the gap: Partisan bias in justifying political violence in the United States

Conditionally accepted at Public Opinion Quarterly (2025)

Version: 1.0 (2025-10-29)
Data file: NO26467145_Power_of_Party_data_Final (1).sav
Format: SPSS (.sav) with value labels where applicable
Study location: United States
Field dates: 2024-01-25 to 2024-01-31
Sample size (raw): 2,030
Analysis sample (attentive partisans): 1,521

This codebook documents variables used in the manuscript and appendix, and the derived variables created by the replication script. Variable names reflect the .sav file and the R script.

Notes:

Analyses are unweighted (sample average treatment effects via random assignment).

The replication script filters to attentive partisans (QAttention == 4) and excludes non-partisans.

By design, only one of B1–B6 should be non-missing for each respondent (single vignette assignment).

1. **Study overview**

Design: 2 × 3 between-subjects vignette experiment embedded in a web survey (YouGov matched panel).

Factor 1: Party condition

No Party (control; partisan labels omitted)

Inparty (victim = respondent's own party)

Outparty (victim = opposing party)

Factor 2: Provocation severity

Mild (shouting/screaming)

Severe (physical obstruction of venue entrance)

Primary outcome: Perceived justifiability of a beating administered by rally attendees to disruptors (1–7 scale).

2. Variable-level documentation (original survey fields)

For each variable: name, label/description, response options/coding, and notes.

QAttention

Label: Attention check

Question: “Please select the number four below.”

Values: 1, 2, 3, 4, 5

Pass criterion: QAttention == 4

Notes: Used to retain attentive respondents.

A1

Label: Party identification (baseline)

Question: “Generally speaking, do you usually think of yourself as a Democrat, a Republican, an Independent, or what?”

Values:

1 = Democrat

2 = Republican

3 = Independent

4 = Some other party

5 = Do not want to answer

Notes: Follow-ups depend on A1 response.

A2

Label: Partisan strength (asked if A1 ∈ {Democrat, Republican})

Question: “Would you call yourself a strong [Democrat/Republican] or a not very strong [Democrat/Republican]?”

Values:

1 = Strong

2 = Not very strong

A3

Label: Partisan leaning (asked if A1 ∈ {Independent, Some other party, Do not want to answer})

Question: “If you had to choose, do you think of yourself as closer to the Democratic Party or the Republican Party?”

Values:

1 = Closer to the Democratic Party

2 = Closer to the Republican Party

3 = Neither

pid3

Label: Party family (tri-categorized)

Values:

1 = Democrat

2 = Republican

3 = Independent

Notes: Used with A2/A3 to classify partisanship and party family.

B1–B6

Label: Vignette response (1–7 justifiability)

Question stem (group-specific; one group per respondent):

B1 (Control, mild): Rally disrupted by shouting/screaming; attendees beat disruptors. Response scale 1–7.

B2 (Control, severe): Entrance physically blocked; attendees beat disruptors. Response scale 1–7.

B3 (Own party vs. opposing, mild): [inparty] rally; disruptors are [outparty] shouting/screaming; [inparty] attendees beat them. Response 1–7.

B4 (Own party vs. opposing, severe): [inparty] rally; [outparty] physically block entrance; [inparty] attendees beat them. Response 1–7.

B5 (Opposing vs. own, mild): [outparty] rally; disruptors are [inparty] shouting/screaming; [outparty] attendees beat them. Response 1–7.

B6 (Opposing vs. own, severe): [outparty] rally; [inparty] physically block entrance; [outparty] attendees beat them. Response 1–7.

Values: 1 = Completely unjustified … 7 = Completely justified

Notes: Exactly one of B1–B6 should be non-missing. The [inparty]/[outparty] tokens are personalized using A1/A3.

C1_Dark_personality

Label: Dark personality — exploitation (SIDDD)

Question: “I tend to exploit others towards my own end.”

Values: 1 = Strongly disagree … 7 = Strongly agree

Notes: Single-item measure (Dirty Dozen core).

Q_Gender

Label: Gender

Labeled values (from data distribution):

Male

Female

Other

Notes: The script converts labeled values to factors; for covariates, restricted to Male/Female.

age

Label: Age in years

Values: Integer (18–93 observed)

Notes: Used directly and as age10 = age/10.

profile_gross_household

Label: Household income (categorical)

Values (observed categories):

Less than $10,000

$10,000 - $19,999

$20,000 - $29,999

$30,000 - $39,999

$40,000 - $49,999

$50,000 - $59,999

$60,000 - $69,999

$70,000 - $79,999

$80,000 - $99,999

$100,000 - $119,999

$120,000 - $149,999

$150,000 - $199,999

$200,000 - $249,999

$250,000 - $349,999

$350,000 - $499,999

$500,000 or more

Don’t know / Prefer not to say

Notes: “Don’t know/Prefer not to say” treated as missing in analyses; also grouped to Low/Middle/High in robustness.

race

Label: Race/ethnicity (categorical)

Values (observed categories):

White, Black, Hispanic, Asian, Native American, Middle Eastern, Two or more races, Other

Notes: Often recoded to White vs. Non-White for covariates.

Additional panel/profile variables may be present in the .sav but are not used in the manuscript’s analyses.

3. Derived variables (constructed by the replication script)

The script constructs the following variables; they are not stored in the raw .sav unless you save them yourself.

response

Definition: Coalesced justifiability rating from B1–B6.

Construction: response = first non-missing value across B1, B2, B3, B4, B5, B6.

Scale: 1–7 (higher = more justified).

Notes: Exactly one B* should be non-missing per respondent.

disruption_severity

Type: Factor with levels: “Mild”, “Severe”

Construction:

Mild if B1 or B3 or B5 non-missing (shouting/screaming)

Severe if B2 or B4 or B6 non-missing (physical obstruction)

party_condition

Type: Factor with levels: “No Party”, “Inparty”, “Outparty”

Construction:

“No Party” if B1 or B2 non-missing (control)

“Outparty” if B3 or B4 non-missing (victim = outparty)

“Inparty” if B5 or B6 non-missing (victim = inparty)

Interpretation: Encodes the identity of the victim/target group in the vignette relative to the respondent’s party family.

partisanship

Type: Character (“Strong”, “Moderate”)

Construction:

Strong if (pid3 $\in$ {1,2} and A2 == 1)

Moderate if (pid3 $\in$ {1,2} and A2 == 2) OR (pid3 == 3 and A3 $\in$ {1,2})

Otherwise NA

Notes: Non-partisans/“Neither” leaners are NA and excluded from main analyses.

party_family

Type: Character (“Democrat”, “Republican”)

Construction:

Democrat if pid3 == 1 OR (pid3 == 3 and A3 == 1)

Republican if pid3 == 2 OR (pid3 == 3 and A3 == 2)

Otherwise NA

partisanship_dummy

Type: Numeric dummy (Strong = 1; Moderate = 0; NA otherwise)

Use: Convenience for models.

condone

Type: Logical

Definition: TRUE if response ≥ 5 (Somewhat justified / Justified Completely justified); else FALSE.

Use: Exploratory χ^2 contrasts for "explicit justification".

age10

Type: Numeric

Definition: age / 10

exploit_hi_67

Type: Numeric dummy (1 if C1_Dark_personality ≥ 6; 0 otherwise)

Use: Alternative specification for dark-personality models.

Covariate recodes for robustness models:

gender (factor): Male, Female (ref = Male)

race (factor): White, Non-White (ref = White), from race

income_group (factor):

Low: Less than $30,000

Middle: $30,000–$99,999

High: ≥ $100,000

"Don't know/Prefer not to say" are dropped in covariate-adjusted models.

Script-only convenience dummies (not saved):

female (1 = Female)

nonwhite (1 = Non-White)

party_out_vs_control (1 = Outparty)

party_in_vs_control (1 = Inparty)

partisan_strong (1 = Strong)

disruption_severity_mild (1 = Mild)

income_middle (1 = Middle)

income_high (1 = High)

republican (1 = Republican)

4. Missing data conventions

SPSS system-missing values are read as NA in R.

profile_gross_household "Don't know / Prefer not to say" treated as missing in analyses.

For the main outcome, only respondents with non-missing response are kept.

For the analysis sample, respondents must have QAttention == 4 and non-missing partisanship/party_family.

No imputation is performed.

5. Inclusion/exclusion rules (analysis sample)

Include:

Respondents with exactly one non-missing B1–B6 value (response available)

QAttention == 4

Partisans: partisanship ∈ {Strong, Moderate} and party_family ∈ {Democrat, Republican}

Exclude:

Attention failures (QAttention ≠ 4)

Non-partisans (true independents, "Neither" leaners, missing partisanship)

Respondents without a vignette response

6. Value-label crosswalk (summary)

Justifiability (B1–B6; response)

1 = Completely unjustified

2 = Unjustified

3 = Somewhat unjustified

4 = Neither unjustified nor justified

5 = Somewhat justified

6 = Justified

7 = Completely justified

Attention (QAttention)

4 = Correct answer (pass)

Party family (pid3)

1 = Democrat

2 = Republican

3 = Independent

- Partisan strength (A2)

1 = Strong

2 = Not very strong

Lean (A3)

1 = Closer to the Democratic Party

2 = Closer to the Republican Party

3 = Neither

Dark personality (C1_Dark_personality)

1 = Strongly disagree … 7 = Strongly agree

Gender (Q_Gender)

Labeled categories (Male, Female, Other)

Income (profile_gross_household)

Labeled categories as listed in Section 2

Race (race)

Labeled categories as listed in Section 2

Derived/factor levels (in-script):

disruption_severity: Mild, Severe

party_condition: No Party, Inparty, Outparty

party_family: Democrat, Republican

partisanship: Strong, Moderate

7. Quality checks to reproduce

One and only one of B1–B6 non-missing per respondent.

Retained (attentive partisans) N = 1,521.

Cell counts (attentive partisans):

party_condition: No Party = 498; Inparty = 501; Outparty = 522

disruption_severity: Mild = 744; Severe = 777

8. Ethical and confidentiality notes

IRB approval obtained (see appendix for redacted screenshot).

Data are de-identified survey responses from a non-probability online panel.

No direct identifiers are provided in the shared dataset.

9. Provenance and processing summary

Source: YouGov CAWI survey, matched sample; English only.

Processing:

Read via haven::read_sav

Constructed variables as described above

Filtering for attention and partisan inclusion

No weights applied; no imputation

Outputs: Tables (Word .docx) and figures (PNG) generated by the R script.