The Social Desirability Atlas

Leonardo Bursztyn   Ingar Haaland
Nicolas Roever   Christopher Roth

Experimental Methods, July 11, 2025

Why Social Desirability Bias Matters

- Economists increasingly rely on survey data to measure outcomes such as income, financial decisions, welfare participation, and political preferences.
- Yet because humans are inherently social animals, their responses are often shaped by concerns about how they are perceived by others.
- **Social Desirability Bias (SDB)** arises when respondents adjust their answers to align with social norms rather than reveal their true views or behaviors (Tourangeau and Yan, 2007).
- Left unaddressed, SDB can:
  - distort empirical conclusions,
  - misinform policy decisions,
  - and obscure real-world behavior.

When Does SDB Arise?

- The prevalence and direction of **socially desirable responding** vary widely across topics and survey designs.
- Three key factors shape the likelihood and form of SDB:
    - **Question Sensitivity:** Does the content evoke emotional responses or touch on controversial issues?
    - **Social Context:** Are others present? Are responses identifiable or anonymous?
    - **Motivations:** Does the respondent seek to protect self-image or avoid reputational or material costs?
- These dimensions often interact – jointly determining whether, and how, SDB will manifest.

How Prevalent Is Social Desirability Bias?

- We synthesize empirical evidence on the prevalence and magnitude of SDB across diverse domains.
- Focus is on studies using **individual-level data** and **objective ground-truth measures** to avoid distortions from false positives/negatives.
- SDB is most pronounced when:
  - **Norms are strong**, and
  - **Personal stakes of disclosure are high**
- Especially salient in:
  - Self-reports of criminal behavior
  - Voting behavior
  - Measures of honesty
- Severity of bias is closely tied to the **perceived sensitivity** of the topic.

Pitfalls of Common Mitigation Strategies

- We critically assess strategies used to reduce SDB, such as:
  - *List experiments*
  - *Randomized response techniques*
- While well-intentioned, these methods often **fail to improve data accuracy** in practice.
- Evidence using individual-level benchmarks reveals two key issues:
  - **Confusion:** Complexity can overwhelm or mislead respondents (John et al., 2018).
  - **Backfire effects:** These methods may increase the **perceived sensitivity** of the topic (Loewenstein, 1999).

Tailoring SDB Mitigation to Respondent Motives

- To navigate the challenges of SDB, researchers must **align mitigation strategies with the underlying motivations for misreporting**.

- **Three key motives** behind socially desirable responding:
  - **Material Costs:** Fear of punishment or institutional backlash
    $\Rightarrow$ Use strong anonymity protections
  - **Social-Image Concerns:** Fear of reputational harm
    $\Rightarrow$ Use privacy-enhancing tools that don't confuse participants.
  - **Self-Image Concerns:** Discomfort with personal truths
    $\Rightarrow$ Use indirect framing (e.g., third-person or forgiving outcome framing)

- $\Rightarrow$ Effective SDB mitigation requires identifying **what respondents are trying to protect**—and adapting tools accordingly.

Does SDB Distort Treatment Effects?

- Economists are focused on estimating **treatment effects**.
- Key concern: **SDB can bias treatment effects** if the intervention alters salience and perceptions of what is socially desirable.
- **Best practices to mitigate differential SDB** while preserving power:
    - Use *obfuscated follow-up surveys*, anonymized outcomes, incentivized behavioral outcomes.

Related Literature

- This review builds on interdisciplinary work on SDB, beginning with **personality-based explanations** (Crowne and Marlowe, 1960; Edwards, 1957).
- Classic survey methodology highlights how **question wording, survey mode, and social setting** shape responses on sensitive topics (Bradburn, 1978; Sudman and Bradburn, 1974).
- In economics, our review complements work on the limitations of self-reported data (Bertrand and Mullainathan, 2001; Harrison, 2006; Harrison and List, 2004).
    - We provide a **deeper focus on SDB** and how and why it **distorts prevalence estimates and treatment effects**.

Conceptualizing Social Desirability Bias

What is Social Desirability Bias (SDB)?

- **Definition:** SDB refers to the tendency of respondents to misreport their true attitudes, beliefs, or behaviors to appear more socially acceptable (Tourangeau and Yan, 2007).
- **Distinct from Classical Measurement Error:** Unlike random error, SDB introduces *systematic* and *predictable* distortions in survey data.
- **Cost-Benefit Trade-Off:** Respondents weigh the perceived benefits of misreporting against the costs of honesty-making SDB a utility-driven form of measurement error.

SDB versus Other Response Biases

Experimenter Demand Effects.
Respondents adjust answers based on perceived researcher expectations, even if not socially desirable (de Quidt et al., 2018).
⇒ May overlap with SDB when expectations align with socially valued behavior.

Acquiescence Bias (AB).
Tendency to agree regardless of content (Hurd, 1999). Often due to low effort, confusion, or politeness.
⇒ Unlike strategic SDB, AB reflects passive or unengaged responding.

Hawthorne Effects.
Behavior changes due to being observed (Mayo, 2003). Not tied to social norms.
⇒ Inflated responses may resemble SDB, but stem from reactivity, not impression management.

Why Do People Respond in Socially Desirable Ways?

- **Material costs:** Respondents may misreport to avoid tangible consequences like legal sanctions, job loss, or loss of benefits (Blair et al., 2020).
- **Social image concerns:** When anonymity is uncertain, individuals distort answers to conform with social norms and avoid reputational harm (e.g., Bursztyn et al., 2023b; Bursztyn and Jensen, 2017; Cortés et al., 2024).
- **Self-image concerns:** Even with full privacy, people may report in socially desirable ways to maintain a positive self-image (e.g., Bénabou and Tirole, 2002; Henkel and Zimpelmann, 2023).

Conscious vs. Non-Conscious Social Desirability Bias

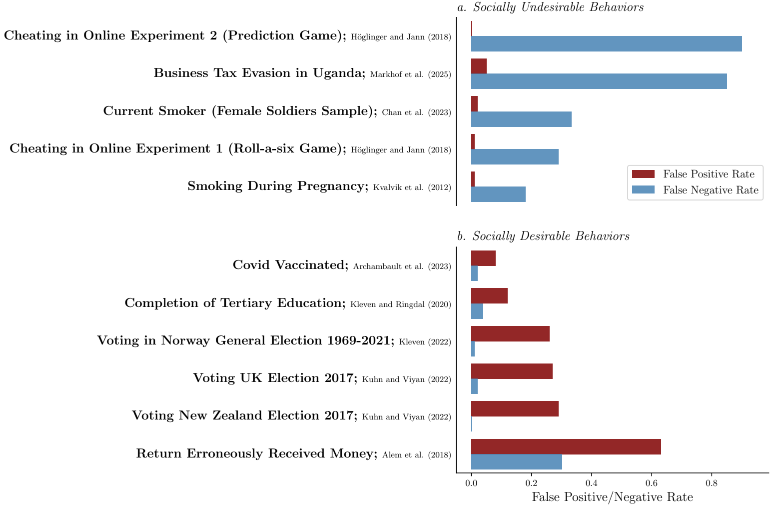
- **Conscious SDB:** Respondents deliberately misreport to avoid judgment or appear favorable (Paulhus, 1984).
  Example: Exaggerating charitable donations or claiming regular voting despite rarely participating.
- **Non-Conscious SDB:** Respondents internalize social norms and unconsciously distort their answers (Paulhus, 1986).
  Example: Genuinely believing one holds egalitarian views while forgetting contradictory past behavior.
- **Key distinction:** Conscious SDB is strategic; non-conscious SDB reflects self-deception.

SDB Across Domains

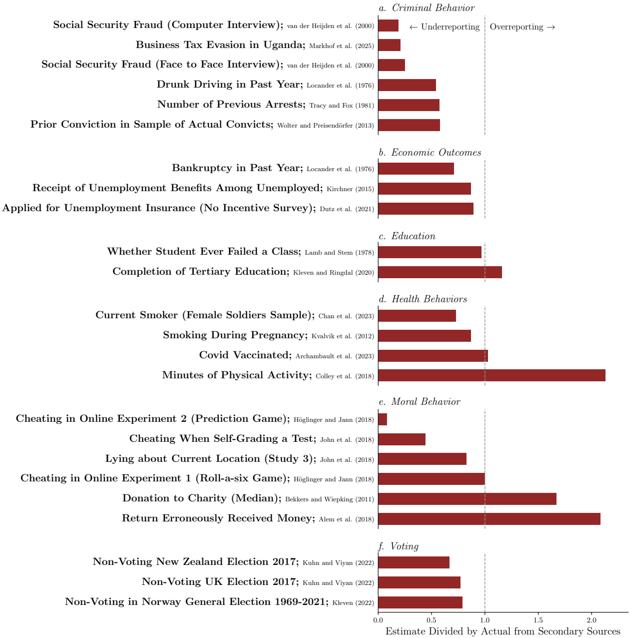
Overview of studies with individual-level benchmarks

- We synthesize high-quality validation studies to identify topics especially prone to SDB.
- Focus is restricted to studies using *individual-level ground truth data* (e.g., administrative records).
- This avoids issues with:
  - **False positives** in method comparisons (e.g., direct vs. list experiment).
  - **Strong assumptions** in aggregate-level benchmarks about survey representativeness.
- While this narrows the scope (i.e. no coverage of topics like racism and sexual behaviour due to missing objective benchmarks), it ensures **more accurate inferences about the presence and magnitude of SDB**.

SDB Across Domains I



SDB Across Domains



When Is Social Desirability Bias Most Severe?

SDB is quantitatively large in three domains:

- **Criminal Behavior:** Large misreporting on tax evasion, fraud, etc., driven by *material risks*, *social-image*, and *self-image concerns*.
- **Voting:** Widely seen as a civic duty, voting is often overreported – especially when social norms are salient (Karp and Brockington, 2005).
- **Honesty:** Studies reveal gaps between reported and actual cheating (Höglinger and Jann, 2018; John et al., 2018).
    - Even subtle wording changes can affect responses, showing strong *context dependence*.

Takeaway: Expect SDB when disclosure carries consequences or social norms are strong.

Mitigation Strategies and their Effectiveness

Mitigating Social Desirability Bias

Key mitigation strategies:

- List experiments
- Randomized Response Techniques (RRT)
- Anonymity guarantees
- Third-person/social-circle questions
- Forgiving outcome framing
- Incentives

List Experiments

What is a List Experiment?

- Also called the *item count technique*
- Indirectly infers prevalence of sensitive behaviors
- Respondents randomized into:
    - Control group: List of non-sensitive items
    - Treatment group: Same list + one sensitive item
- Reduces direct connection between individual and sensitive answer

List Experiment Example: Drug Use

- Control group sees:
  - Likes coffee
  - Owns a pet
  - Has traveled abroad
- Treatment group sees same list + "Used illicit drugs"
- Average number of endorsed items:
  - Control: 1.8
  - Treatment: 2.3
- Estimated drug use prevalence: $2.3 - 1.8 = 0.5$

List Experiment: Key Assumptions

1. Adding the sensitive item should not change responses to non-sensitive items (Coffman et al. 2017)

2. Respondents who endorse the sensitive item must include it in their count

List Experiment: Potential Pitfalls

1. Does not mitigate self-image bias
2. May increase salience of sensitivity (Loewenstein et al. 1999)
3. Misunderstanding or miscounting
4. Floor/ceiling effects
5. High variance; requires large sample sizes

List Experiment: Evidence on Effectiveness

- Often produces higher estimates than direct questions (Blair et al. 2020; Li et al. 2022; Ehler 2021)

Two cautionary examples where individual-level ground truth data is available:

- Markhof et al. (2024): 18% false positive rate in tax evasion data from Uganda
- Kuhn and Viyanathan (2022): Lower accuracy for voter turnout due to satisficing

Randomized Response Technique

RRT: Procedure

- Introduced by Warner (1965) to address social desirability bias.
- Respondents follow a random rule (e.g., coin flip) before answering.
- The researcher never learns which question was answered.
- Enables plausible deniability and reduces misreporting.

RRT Example: Measuring Drug Use

- Coin flip:
  - Heads: answer "Have you ever used illegal drugs?"
  - Tails: answer "Is your birthday in January or February?"
- Only "yes" or "no" is reported.
- If non-sensitive prevalence is known, the sensitive behavior can be estimated.

RRT: Assumptions

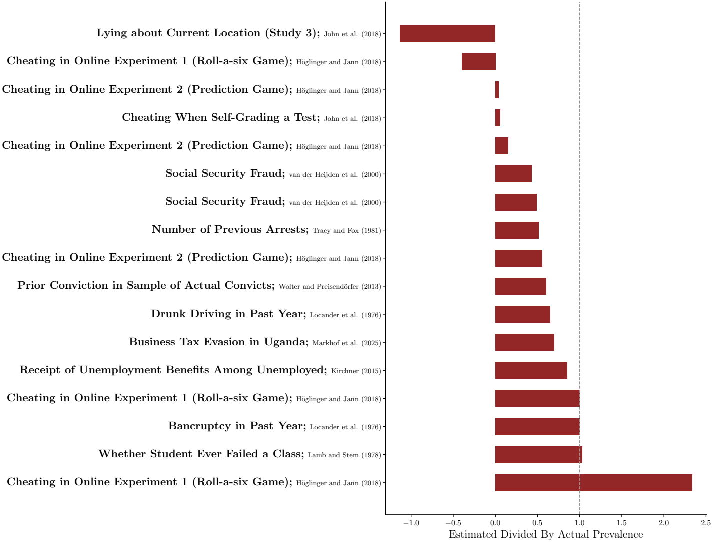
- Respondents must follow the randomization protocol accurately.
- Truthful responses are required when instructed.
- Randomization must be independent of true answers.
- Trust in anonymity is crucial.

Pitfalls and Limitations

- Misunderstanding or distrust biases estimates.
- Non-compliance with instructions among surveyed individuals appears prevalent and not easy to characterize (Chuang et al., 2021).
- Coutts and Jann (2011) show limited trust even with correct understanding.
- Increased noise due to probabilistic element.

RRT: Evidence on Effectiveness

- Studies with individual-level benchmarks show poor performance.



Why RRT May Fail

- **Salience:** RRT may highlight sensitivity, increasing underreporting (John et al., 2018).
- **Complexity:** RRT may induce non-classical measurement error by creating confusion.

Anonymity guarantees

Anonymity Guarantees

Objective: Reduce Social Desirability Bias (SDB) by assuring respondents that answers remain untraceable.

Core Idea:

- Emphasize anonymity to reduce external judgment concerns.
- Common methods include self-paced, anonymous online surveys.

Procedure

- Administer surveys in formats without collecting identifying information.
- Explicitly communicate confidentiality.
- Ensure surveys are conducted anonymously and online.

Assumptions and Potential Pitfalls

Key Assumptions:

- Respondents must perceive responses as genuinely anonymous.
- Anonymity should not lead to careless or strategic misreporting.

Potential Pitfalls:

- Reduced engagement or unintended behavior changes in experiments.
- Excessive anonymity emphasis may signal sensitivity, causing defensive answering or increased non-response (Loewenstein, 1999).

Evidence on Effectiveness

- Mixed empirical support for anonymity effectiveness:
  - Increased willingness to admit sensitive behaviors (e.g., drug use, tax evasion) (Tourangeau and Yan, 2007).
  - Murdoch et al. (2014) find no anonymity effect in mailed surveys on sensitive veteran topics.
  - Experimental economics evidence is mixed (Hoffman et al., 1996; List et al., 2004).
- Alternative strategies include obscuring survey intent or deflecting attention.

Third-person and social circle questions

Procedure

- Galesic et al. (2018) asked respondents about voting intentions in their social circles, finding this yields more accurate election outcome predictions.
- Another variant systematically varies the reference group, contrasting sensitive behaviors among in-group vs. out-group members to detect SDB (Chakravarty et al., 2022).
- Bursztyn et al. (2023a) differentiate between what respondents believe others "would say they agree with" vs. what they "truly agree with," using discrepancies to measure SDB.

Assumptions and Potential Pitfalls

- Relies on higher-order beliefs, imposing higher cognitive load.
- Effective when social networks are homogeneous and clearly defined.
- Heterogeneous or poorly defined groups increase noise and measurement error.
- Issues like pluralistic ignoranceâsystematic misperceptions of peer attitudesâcan undermine reliability (Bursztyn and Jensen, 2017; Roth et al., 2024).

Evidence on Effectiveness

- Empirical evidence remains scarce and mixed.
- Positive results: Galesic et al. (2018) demonstrated accuracy in predicting election outcomes.
- Bursztyn et al. (2023a) found minimal evidence of SDB using their third-person approach, suggesting method viability under certain conditions.

Forgiving Outcome Framing

What is Forgiving Outcome Framing?

Forgiving outcome framing is a pragmatic, low-cost method to reduce Social-Desirability Bias (SDB) in self-reports.

Core Idea: Embed questions within normalizing narratives, indicating that various behaviors are common and acceptable.

Example:

- Instead of: "Do you always follow public-health guidelines?"
- Use: "People vary in how closely they follow public-health recommendations. How often did you...?"

Procedure

Implementing forgiving framing involves:

1. **Normalizing Statement:** Start with a brief statement acknowledging variation.
2. **Attribution to External Factors:** Attribute non-desirable behaviors to external, benign factors (e.g., workplace culture, time constraints).
3. **Even-Handed Response Scales:** Use balanced scales (e.g., never, rarely, sometimes, often, always) to normalize less desirable responses.

Assumptions and Pitfalls

Key Assumptions:

- Respondents are more candid if their self-esteem is protected and external judgment concerns are reduced.

Potential Pitfalls:

- Excessively explicit preambles may increase perceived sensitivity and thus bias.
- Forgiving outcome framing in itself could induce a demand effect.

Evidence on Effectiveness

Empirical evidence is mixed and context-dependent:

- Higher disclosure only under strict social norms; otherwise null effects (Näher and Krumpal, 2012).
- Increased truthful reporting among high-SDB adolescents (Peter and Valkenburg, 2011).
- Potentially counterproductive in interviewer-led contexts (Kaplan and Yu, 2015).
- Modest improvements in accurate turnout reporting with normalizing disclaimers (Belli et al., 2006).

Incentives

Why Incentives?

- Incentives are widely used in experimental economics to elicit truthful responses
- Key idea: make misreporting financially costly
- Easy to implement when outcomes are verifiable
- Many surveys lack objective benchmarks
- Some methods attempt to adapt incentives to settings without external benchmarks (e.g., Bayesian Truth Serum)

Procedure

- When an objective benchmark exists, respondents can be rewarded for accuracy
- Discrete: reward for correct answer
- Continuous: reward based on proximity (e.g., within a certain range)
- Effective for beliefs about:
  - Economic indicators
  - Second-order normative and descriptive beliefs

Assumptions and Pitfalls

- Respondents must:
  - Understand the incentive structure
  - Trust that payouts are based on accuracy
- Complex incentives may reduce truthfulness (Danz et al. 2022) as a result of confusion

- Respondents may prioritize expressive motives over money

- Risk of gaming or searching for correct answer instead of revealing true belief

Evidence on Effectiveness

- Reduce partisan bias in beliefs (Bullock et al., 2015; Prior et al., 2015)
- Lower political polarization in some settings (Peterson and Iyengar, 2021)
- Weaker or absent effects for rumors and misinformation (Allcott et al., 2020; Berinsky, 2018)
- Reduce motivated forgetting in image-related information (Zimmermann, 2020)

When SDB distorts treatment effects

SDB and Treatment Effects

- In some experiments, the treatment itself may alter what is perceived as the **socially desirable response**.
    - For instance, providing information about immigrants may shift respondents' beliefs about socially acceptable answers.
- This raises the concern that **SDB may distort treatment effects**, as the treatment changes the salience or perception of what is socially desirable.
- In such cases, **SDB may operate differently across treatment and control**, biasing estimates of causal effects.

Minimizing SDB in Treatment Effect Estimation

- This section outlines best practices for minimizing the impact of Social Desirability Bias (SDB) on treatment effect estimates in experimental settings.
- We focus on methods that maintain sufficient **statistical power**.
  - Some techniques discussed earlier (e.g., list experiments) introduce **substantial noise** and are therefore impractical for treatment effect estimation due to large required sample sizes.

Overview: Methods to Mitigate SDB

- Natural Field Experiments
- Use of Behavioral or Administrative Outcomes
- Implicit Measures
- Anonymity (Group-Level Outcomes)
- Obfuscated Follow-Ups
- Purpose Obfuscation
- Measuring Perceived Social Norms
- Heterogeneity by Sensitivity to Social Norms

Natural Field Experiments

Natural field experiments embed interventions in real-world settings without participants being aware of the study (Harrison and List, 2004). This reduces SDB by:

- Avoiding experimenter demand effects.
- Increasing the cost of misrepresentation due to real-world stakes.

Behavioral or Administrative Outcomes

Objective outcomes help isolate treatment effects from self-report bias:

- Behavioral data
- Administrative records

These measures avoid contamination from differential reporting across groups.

Implicit Measures

Psychometric tools such as IATs aim to measure subconscious attitudes:

- Bypass conscious self-presentation (Greenwald et al., 1998).
- May indicate where SDB is particularly strong.
- Validity concerns remain (Blanton et al., 2015; Oswald et al., 2013).

Anonymity via Group-Level Outcomes

Outcomes observable only at aggregate level help reduce SDB:

- e.g., Online petition signatures measured only by group (Grigorieff et al., 2020).
- Preserves privacy but limits individual-level analysis.

Obfuscated Follow-Ups

Follow-up surveys that obscure connection to initial treatment reduce SDB (Haaland and Roth, 2020):

- Break link between treatment and outcome measurement.
- Reduce tailored responses.

Purpose Obfuscation

Avoid revealing the normative alignment of the researcher:

- e.g., Random assignment to ideologically opposed donation options (Bursztyn et al., 2020).
- Use cover stories to mask research goals.

Maintains uncertainty about socially desirable answers.

Measuring What is Perceived as Socially Desirable

Directly elicit beliefs about socially desirable behavior:

- Use open-ended questions at the end (Haaland et al., 2025).
- Measure multiple belief types (1st/2nd order, normative/descriptive) (Harrison and Swarthout, 2025).
- Assess whether these beliefs differ across treatment conditions.

Heterogeneity by Norm Sensitivity

Test whether treatment effects vary by individual sensitivity to norms:

- Psychometric scales: Marlowe-Crowne (Crowne and Marlowe, 1960), Self-Monitoring (Snyder, 1974).
- Identify who is more susceptible to SDB (Allcott and Taubinsky, 2015; Dhar et al., 2022).
- Hard to interpret these heterogeneity tests given correlation of scales with levels of outcomes.

Conclusion

Key Takeaways

- **SDB is domain-specific:** Its prevalence and magnitude vary depending on topic sensitivity.
- **Mitigation is tricky:** Some mitigation strategies, such as RRT, fail by confusing respondents or heightening sensitivity.
  - Complexity of method comes at the cost of confusion.
- **Biased treatment effects:** SDB can distort not just levels but also treatment effects, especially when treatments shift perceived norms.

Key Takeaways

Guidance for Researchers:

- Treat SDB as a *context-dependent distortion*: anticipate, measure, and address it.
- Match mitigation strategies to the *underlying mechanisms*: self-image, social-image, or material cost concerns.
- Tailor approaches to the specific goals and constraints of the research design.

Thank you!

Questions and Comments?

References I

Allcott, H., Boxell, L., Conway, J., Gentzkow, M., Thaler, M., and Yang, D. Y. (2020). Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of Public Economics*, 191:104254.

Allcott, H. and Taubinsky, D. (2015). Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. *American Economic Review*, 105(8):2501–2538.

Belli, R. F., Moore, S. E., and VanHoewyk, J. (2006). An experimental comparison of question forms used to reduce vote overreporting. *Electoral Studies*, 25(4):751–759.

Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics*, 117(3):871–915.

Berinsky, A. J. (2018). Telling the truth about believing the lies? evidence for the limited prevalence of expressive survey responding. *The Journal of Politics*, 80(1):211–224.

Bertrand, M. and Mullainathan, S. (2001). Do people mean what they say? implications for subjective survey data. *American Economic Review*, 91(2):67–72.

Blair, G., Coppock, A., and Moor, M. (2020). When to worry about sensitivity bias: A social reference theory and evidence from 30 years of list experiments. *American Political Science Review*, 114(4):1297–1315.

Blanton, H., Jaccard, J., Strauts, E., Mitchell, G., and Tetlock, P. E. (2015). Toward a meaningful metric of implicit prejudice. *Journal of Applied Psychology*, 100(5):1468–1481.

Bradburn, N. (1978). Respondent burden. In *Proceedings of the Survey Research Methods Section of the American Statistical Association*, volume 35, pages 35–40. American Statistical Association Alexandria, VA, USA.

Bullock, J. G., Gerber, A. S., Hill, S. J., and Huber, G. A. (2015). Partisan bias in factual beliefs about politics. *Quarterly Journal of Political Science*, 10(4):519–578.

Bursztyn, L., Cappelen, A. W., Tungodden, B., Voena, A., and Yanagizawa-Drott, D. H. (2023a). How are gender norms perceived? NBER Working Paper No. 31049, National Bureau of Economic Research.

Bursztyn, L., Egorov, G., and Fiorin, S. (2020). From extreme to mainstream: The erosion of social norms. *American Economic Review*, 110(11):3522–3548.

Bursztyn, L., Egorov, G., Haaland, I., Rao, A., and Roth, C. (2023b). Justifying dissent. *Quarterly Journal of Economics*, 138(3):1403–1451.

References II

Bursztyn, L. and Jensen, R. (2017). Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics*, 9:131–153.

Chakravarty, A., Ghosh, A., Lowe, M., and Nellies, G. (2022). Learning about outgroups: The impact of broad versus deep interactions. CESifo Working Paper No. 11363.

Chuang, E., Dupas, P., Huillery, E., and Seban, J. (2021). Sex, lies, and measurement: Consistency tests for indirect response survey methods. *Journal of Development Economics*, 148:102582.

Cortés, P., Koşar, G., Pan, J., and Zafar, B. (2024). Should mothers work? how perceptions of the social norm affect individual attitudes toward work in the us. *Review of Economics and Statistics*, pages 1–28.

Coutts, E. and Jann, B. (2011). Sensitive questions in online surveys: Experimental results for the randomized response technique (rrt) and the unmatched count technique (uct). *Sociological Methods & Research*, 40(1):169–193.

Crowne, D. P. and Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4):349–354.

de Quidt, J., Haushofer, J., and Roth, C. (2018). Measuring and bounding experimenter demand. *American Economic Review*, 108(11):3266–3302.

Dhar, D., Jain, T., and Jayachandran, S. (2022). Reshaping adolescentsâ gender attitudes: Evidence from a school-based experiment in india. *American Economic Review*, 112(3):899–927.

Edwards, A. L. (1957). *The Social Desirability Variable in Personality Assessment and Research*. Dryden Press, New York.

Galesic, M., Bruine de Bruin, W., Dumas, M., Kapteyn, A., Darling, J. E., and Meijer, E. (2018). Asking about social circles improves election predictions. *Nature Human Behaviour*, 2:187–193.

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6):1464–1480.

Grigorieff, A., Roth, C., and Ubfal, D. (2020). Does information change attitudes toward immigrants? *Demography*, 57:1117–1143.

Haaland, I. and Roth, C. (2020). Labor market concerns and support for immigration. *Journal of Public Economics*, 191:104256.

References III

Haaland, I. K., Roth, C., Stantcheva, S., and Wohlfart, J. (2025). Understanding economic behavior using open-ended survey data. *Journal of Economic Literature*.

Harrison, G. W. (2006). Hypothetical bias over uncertain outcomes. In List, J. A., editor, *Using Experimental Methods in Environmental and Resource Economics*. Edward Elgar Publishing.

Harrison, G. W. and List, J. A. (2004). Field experiments. *Journal of Economic Literature*, 42(4):1009–1055.

Harrison, G. W. and Swarthout, J. T. (2025). Causal inferences over unobservables. CEAR Working Paper 2025-02, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University.

Henkel, L. and Zimpelmann, C. (2023). Proud to not own stocks: How identity shapes financial decisions. IZA Discussion Paper No. 16246, Institute of Labor Economics.

Hoffman, E., McCabe, K., and Smith, V. L. (1996). Social distance and other-regarding behavior in dictator games. *American Economic Review*, 86(3):653–660.

Höglinger, M. and Jann, B. (2018). More is not always better: An experimental individual-level validation of the randomized response technique and the crosswise model. *PloS one*, 13(8):e0201770.

Hurd, M. D. (1999). Anchoring and acquiescence bias in measuring assets in household surveys. *Journal of Risk and Uncertainty*, 19(1-3):111–136.

John, L. K., Loewenstein, G., Acquisti, A., and Vosgerau, J. (2018). When and why randomized response techniques (fail to) elicit the truth. *Organizational Behavior and Human Decision Processes*, 148:101–123.

Kaplan, R. L. and Yu, E. C. (2015). Measuring question sensitivity. *American Association for Public Opinion Research*, pages 4107–4121.

Karp, J. A. and Brockington, D. (2005). Social desirability and response validity: A comparative analysis of overreporting voter turnout in five countries. *The Journal of Politics*, 67(3):825–840.

List, J. A., Berrens, R. P., Bohara, A. K., and Kerkvliet, J. (2004). Examining the role of social isolation on stated preferences. *American Economic Review*, 94(3):741–752.

Loewenstein, G. (1999). Experimental economics from the vantage-point of behavioural economics. *The Economic Journal*, 109(453):25–34.

Mayo, E. (2003). *The Human Problems of an Industrial Civilization*. Routledge.

References IV

Murdoch, M., Simon, A. B., Polusny, M. A., Bangerter, A. K., Grill, J. P., Noorbaloochi, S., and Partin, M. R. (2014). Impact of different privacy conditions and incentives on survey response rate, participant representativeness, and disclosure of sensitive information: a randomized controlled trial. *BMC Medical Research Methodology*, 14:90.

Näher, A.-F. and Krumpal, I. (2012). Asking sensitive questions: The impact of forgiving wording and question context on social desirability bias. *Quality & Quantity*, 46:1601–1616.

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., and Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of iat criterion validity. *Journal of Personality and Social Psychology*, 105(2):171–192.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3):598–609.

Paulhus, D. L. (1986). Self-deception and impression management in test responses. In *Personality Assessment via Questionnaires: Current Issues in Theory and Measurement*, pages 143–165. Springer.

Peter, J. and Valkenburg, P. M. (2011). The impact of âforgivingâ introductions on the reporting of sensitive behavior in surveys: The role of social desirability response style and developmental status. *Public Opinion Quarterly*, 75(4):779–787.

Peterson, E. and Iyengar, S. (2021). Partisan gaps in political information and information-seeking behavior: Motivated reasoning or cheerleading? *American Journal of Political Science*, 65(1):133–147.

Prior, M., Sood, G., and Khanna, K. (2015). You cannot be serious: The impact of accuracy incentives on partisan bias in reports of economic perceptions. *Quarterly Journal of Political Science*, 10(4):489–518.

Roth, C., Schwardmann, P., and Tripodi, E. (2024). Depression stigma. CESifo Working Paper No. 11012, CESifo.

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, 30(4):526–537.

Sudman, S. and Bradburn, N. M. (1974). *Response Effects in Surveys: A Review and Synthesis*. Aldine Publishing, Chicago.

Tourangeau, R. and Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133(5):859–883.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69.

Zimmermann, F. (2020). The dynamics of motivated beliefs. *American Economic Review*, 110(2):337–361.