

Moral Licensing: An Empirical Audit and Review

Kristin Donnelly, Smriti Mehta, Tory Taylor, Paige Amormino, Julia Bottesini, Sandy Campbell, Dane Cannon, Beth Clarke, Randy Y. Gao, Stephanie Hong, Beidi Hu, Yunhao Huang, Yikun Jiang, Giulia Maimone, Jacob Miranda, Andres Montealegre, Aaron Nichols, Alex Park, William Ryan, Ameer Sarwar, Sarah Schiavone, Charlie Townsend, Ozge Ugurlu, Yixiang Xu, Don A. Moore, and Leif D. Nelson

Is this a familiar situation?



Is this a familiar situation?



Is this a familiar situation?



Is this a familiar situation?



Moral Licensing



Moral Licensing

- A psychological phenomenon where **prior good behavior** makes people **more likely to engage in questionable behavior**
- Core idea: People maintain a kind of "moral balance sheet"
 - Virtuous actions build up "credit" that can later be "spent" on questionable behaviors

Original Demonstration

- Monin & Miller (2001): After establishing **non-racist credentials**, participants were more likely to make **discriminatory hiring decisions**
- *Key finding*: Participants who selected a Black applicant for one job were subsequently more likely to recommend a White applicant for another position
- Similar pattern found with gender
 - Establishing non-sexist credentials increased likelihood of gender-biased decisions

Original Demonstration

Moral credentials and the expression of prejudice.

[B Monin](#), [DT Miller](#)

Journal of personality and social psychology, 2001 • psycnet.apa.org

Abstract

Three experiments supported the hypothesis that people are more willing to express attitudes that could be viewed as prejudiced when their past behavior has established their credentials as nonprejudiced persons. In Study 1, participants given the opportunity to disagree with blatantly sexist statements were later more willing to favor a man for a stereotypically male job. In Study 2, participants who first had the opportunity to select a member of a stereotyped group (a woman or an African American) for a category-neutral

SHOW MORE ▾

☆ Save [Cite](#) [Cited by 1589](#) [Related articles](#) [All 20 versions](#)

Moral Licensing Across Different Domains

1. Prejudice & Discrimination

- After recalling past non-racist behavior, increased willingness to express potentially racist views (Effron et al., 2009; Effron et al., 2012)
- After believing they'd made progress toward egalitarian goals, participants sat farther from Black confederates (Mann & Kawakami, 2012)
- Proxy credentials: Even having a close friend select a minority candidate provided licensing (Bradley-Geist et al., 2010)

Moral Licensing Across Different Domains

2. Environmental Behavior

- Purchasing (or “purchasing”) "green" products led to:
 - Lower pro-environmental tendencies (Geng et al., 2016)
 - Reduced interest in learning about carbon footprint (Gholamzadehmir et al., 2019)
 - More cheating and less generosity (Mazar & Zhong, 2010)

Moral Licensing Across Different Domains

2. Environmental Behavior

- Just imagining pro-environmental behavior reduced willingness to engage in similar actions later (Chatelain et al., 2018)
- Pattern applied to both real and hypothetical green behaviors

Moral Licensing Across Different Domains

3. Dishonesty

- Imagined charitable donations increased cheating (Brown et al., 2011)
- Recalling past helping increased dishonesty (Greene & Low, 2014)
- Anticipated charitable giving licensed current cheating (Cojoc & Stoian, 2014)

Moral Licensing Across Different Domains

4. Charitable giving

- Lower donations to charity after affirming a moral identity by...
- Sending “thoughts and prayers”(Thunström, 2020)
- Writing about themselves using positive (vs. negative) words (Sachdeva et al., 2009)

Moral Licensing Across Different Domains

5. Indulgent Consumer Choices

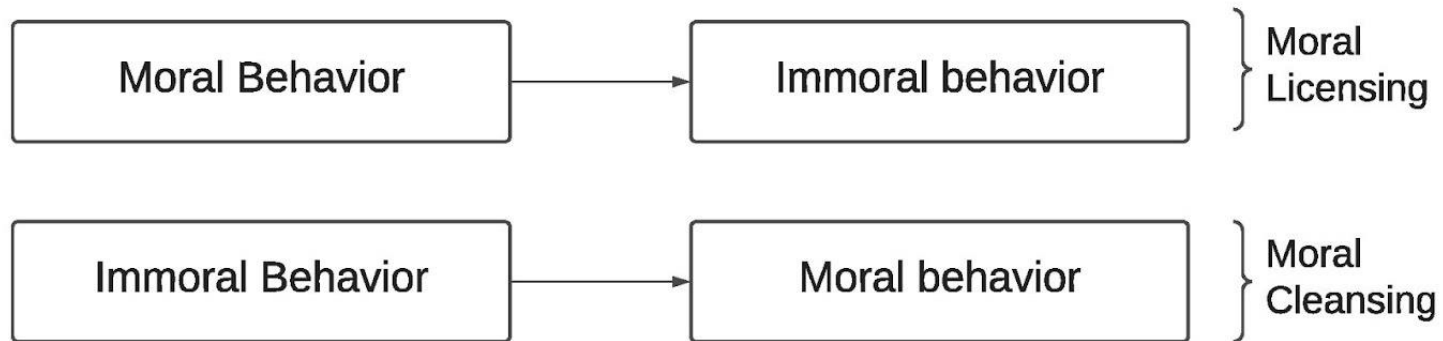
- Imagined hypothetical community service, more willing to hypothetically buy luxury items (Khan & Dhar, 2006)
- Imagined breaking a hypothetical diet for a good reason made people more willing to hypothetically indulge in chocolate pie (Prinsen et al., 2016)

The Complement to Moral Licensing: Moral Cleansing

- **Moral Cleansing**: Immoral behavior leads to compensatory virtuous behavior
- Examples:
 - After playing violent video games, increased interest in hygiene products (Gollwitzer & Melzer, 2012)
 - Physical cleansing after unethical actions (Zhong & Liljenquist, 2006)
- Together with moral licensing, suggests people maintain a "moral equilibrium"

A Hydraulic Relationship

Moral Balancing



The Opposite of Licensing: Consistency

The Opposite of Licensing: Consistency

Moral Licensing:

The Opposite of Licensing: Consistency

Moral Licensing:

Good → Bad (or Bad → Good, for cleansing)

The Opposite of Licensing: Consistency

Moral Licensing:

Good → Bad (or Bad → Good, for cleansing)

Moral Consistency:

The Opposite of Licensing: Consistency

Moral Licensing:

Good → Bad (or Bad → Good, for cleansing)

Moral Consistency:

Good → Good (or Bad → Bad)

The Opposite of Licensing: Consistency

Moral Licensing:

Good → Bad (or Bad → Good, for cleansing)

Moral Consistency:

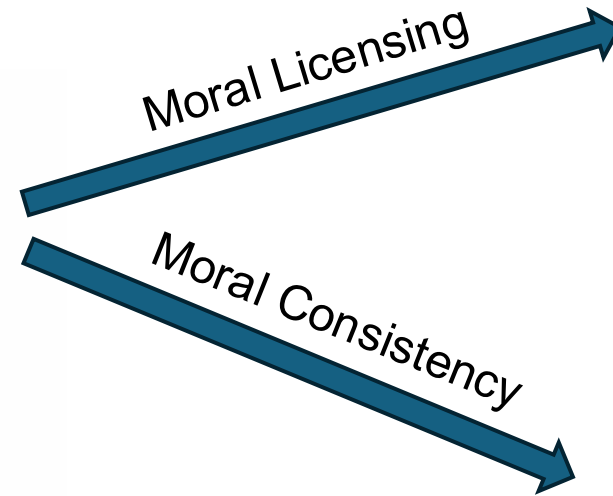
Good → Good (or Bad → Bad)

- Literature shows evidence for both patterns

Moral Licensing vs. Moral Consistency

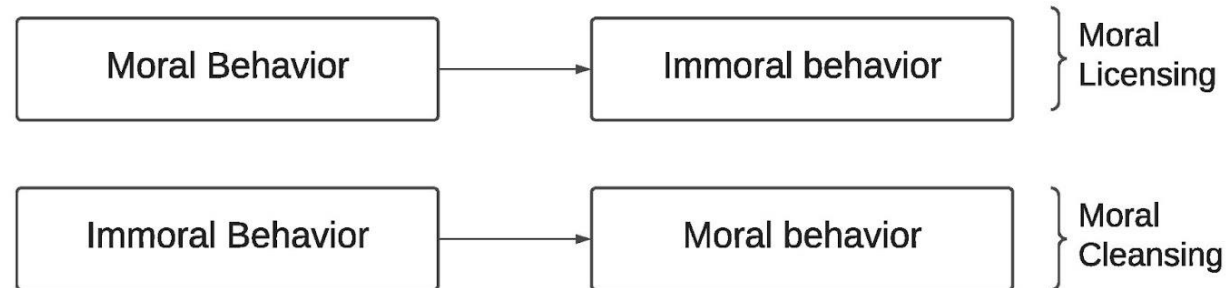


Moral Licensing vs. Moral Consistency

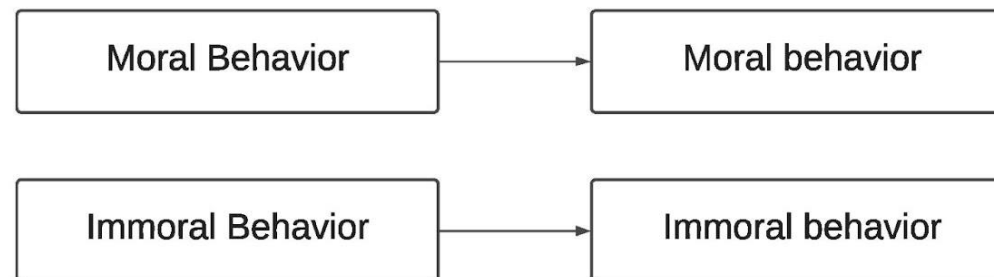


Moral Licensing vs. Moral Consistency

Moral Balancing



Moral Consistency



When Does Licensing vs. Consistency Occur? It's not well-understood...

- **Construal level:** Concrete actions → licensing; abstract traits → consistency (Conway & Peetz, 2012)
- **Temporal distance:** Recent past → licensing; distant past → consistency (Conway & Peetz, 2012)
- **Mindset:** Progress towards moral value → licensing; commitment to moral value → consistency (Susewind & Hoelzl, 2014)
- **Cost:** Costless prosocial behavior → licensing; costly behavior → consistency (Gneezy et al., 2012)
- **Individual differences:** Environmental consciousness moderates environmental licensing (Meijers et al., 2019)

Previous Assessments of Moral Licensing

- Meta-analyses found **small but reliable effects**
 - Blanken et al. (2015): Average effect size $d = 0.31$
 - Simbrunner & Schlegelmilch (2017): Culture moderated licensing effects (stronger effects in Western samples)
- But, meta-analyses have limitations...
 - Vulnerable to publication bias and selective reporting
 - Can only analyze what has been published
- Many Labs 3 successfully replicated Monin & Miller (2001), but with smaller effect size
- No comprehensive, systematic assessment of the entire literature

Our Approach: Empirical Audit and Review

- Novel method introduced by O'Donnell et al. (2021), which involved:
 - Random sampling of studies from well-defined literature
 - Direct replication of selected studies
- Approach aimed to assess overall empirical support for the theory
- Has advantages over meta-analysis and targeted replications

“Seed” Papers and Research Team

- Started with papers that:
 - a) Were included in a meta-analysis by Blanken et al. (2015) and/or a review paper by Mullen & Monin (2016)
 - b) Had cited the meta-analysis (Blanken et al., 2015)
- This left a total sample of 224 articles
- Co-authors/replicators were 22 students in an PhD Open Science seminar (taught by Don Moore and Leif Nelson at Berkeley Haas)

Study Selection

- Two replicators examined each paper, filtering for those that:
 1. Reported experiments
 2. Used manipulations to induce moral licensing
 - Defined by Mullen & Monin (2016) as “when a positive initial behavior yields less positive target behavior than a neutral baseline condition”
 3. Could be conducted online
- Reviewers had to agree on all three dimensions
 - Left ~ 110 studies from 53 unique papers

More on Study Selection

- Randomized papers and then assigned them to replicators in alphabetical order
 - If multiple studies were eligible for replication, we randomly selected one
- Second check by the lead author and the replicator
 - Excluded and reassigned studies based on implementation challenges like design features, time horizons, or insufficient detail
 - Replicators contacted the original authors for materials that were not openly available

A Quick Note:

- Of the 22 studies included for replication, one was re-replicated later to better align with the original study's timing (holiday gift-giving scenarios)
- Another study was moved to the supplement due to changes in the meaning of the stimulus materials over time
 - This study was excluded from the cumulative analysis

Replicators were also “Verifiers”

- Each replicator replicated one study and was randomly paired with another replicator to serve as a “verifier”
- Verifiers contributed at multiple stages:
 - Co-authoring pre-registration
 - Pilot-testing the survey
 - Double-checking analyses
 - Providing feedback throughout the process
- Aimed to ensure replications closely adhered to the original studies

Replication Procedure

- Pre-registered all methods, analyses, and sample sizes
- Sample size: 2.5x original (following Simonsohn, 2015)
- Created surveys with minimal deviations from original
- Data collection via Amazon Cloud Research with demographic targeting when needed
- Multiple verification steps to ensure fidelity to original

Another Note:

- Some studies had an interaction term as the primary test of interest
 - When the interaction term reveals moral licensing in one condition, and moral consistency or a null effect in the other
 - i.e., reversal and attenuated interactions, respectively
- In such cases, we examine the outcome as two “separate” findings
- This left 25 findings in total

Outcomes of the Original Studies

Outcomes of the Original Studies

Original studies

- **73%** showed **moral licensing**
- **4%** showed **moral consistency**
- **24%** null

Moral Licensing vs. Moral Consistency



Moral Licensing

Moral Consistency



Results?

First of all...what do you think happened?

Do you think moral licensing generally replicated?

What about its opposite, moral consistency?

Do you think we mostly found null results?

Replication Outcomes

Original studies

- **73%** showed **moral licensing**
- **4%** showed **moral consistency**
- **24%** null

Replications

Replication Outcomes

Original studies

- **73%** showed **moral licensing**
- **4%** showed **moral consistency**
- **24%** null

Replications

- **8%** showed **moral licensing**

Replication Outcomes

Original studies

- **73%** showed **moral licensing**
- **4%** showed **moral consistency**
- **24%** null

Replications

- **8%** showed **moral licensing**
- **20%** showed **moral consistency**

Replication Outcomes

Original studies

- **73%** showed **moral licensing**
- **4%** showed **moral consistency**
- **24%** null

Replications

- **8%** showed **moral licensing**
- **20%** showed **moral consistency**
- **72%** null

Successful Replications

- 5 findings were classified as “successful replications”
 - This is purely based on their alignment with the original statistical conclusions

Successful Replications

- 5 findings were classified as “successful replications”
 - This is purely based on their alignment with the original statistical conclusions
- We replicated:
 - 1 finding consistent with moral licensing (with another reaching marginal significance)
 - 1 finding that originally reflected moral consistency
 - 3 findings that initially reported null results

Domain/Topic Differences?

Domain/Topic Differences?

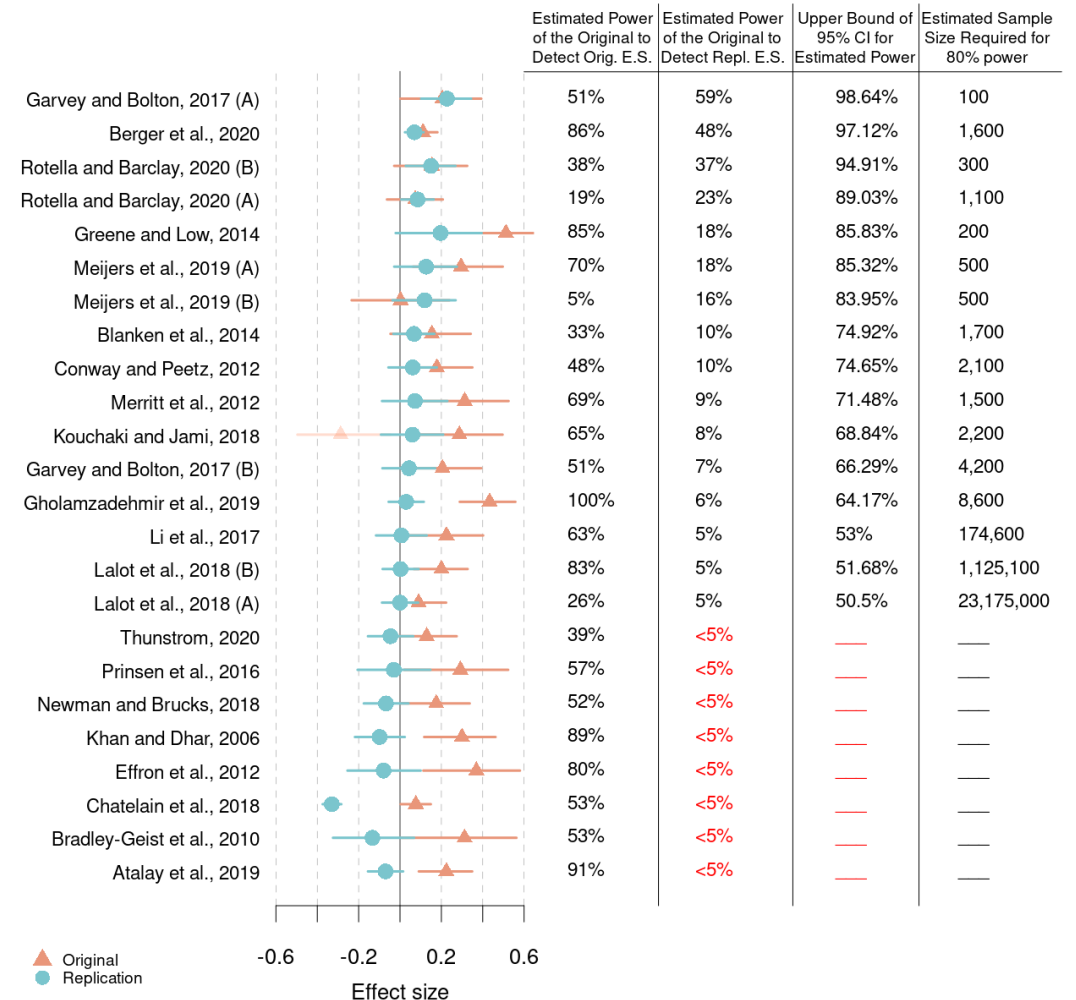
- **Moral licensing** (1-2 studies*):
 - Berger et al. (2020): Dishonest income reporting
 - Greene & Low (2014): Immoral actions (*marginal significance)

Domain/Topic Differences?

- **Moral licensing** (1-2 studies*):
 - Berger et al. (2020): Dishonest income reporting
 - Greene & Low (2014): Immoral actions (*marginal significance)
- **Moral consistency** (5 studies):
 - Environmental behavior (Chatelain et al., 2018; Garvey & Bolton, 2017)
 - Charitable giving (Jones & Koenig, 2018; two studies from Rotella & Barclay, 2020)

Original vs. Replication Effect Sizes

- Effect sizes smaller in 80% of replications (blue circle)
- 64% of replications went in the same direction as original (red triangle)
- Only **5 findings** could be classified as “successful replications” based on their alignment with the original statistical conclusions.
 - E.g., replicating an original null, an original licensing effect, etc.



Original Studies were Underpowered

- Median power of original studies to detect:
 - Their own reported effect size: 58.6%
 - The replication effect size: 11.1%
- Only 25% of original studies reached 80% power threshold
- Median required sample size increase of 177%

Limitations

- **Sampling limitations:** Online-only replications (excluded in-person studies)
- **Replication fidelity:** Possible cultural/temporal shifts in stimuli meaning
 - Example: Racial attitude statements from 2011 might be perceived differently a decade later
- **WIDE variability** in operationalizations and domains...

Lots of ways to manipulate licensing

Lots of ways to manipulate licensing

Moral & Charitable Behavior

- Recalling and describing cooperative/moral behavior vs. uncooperative/immoral behavior.
- Considering moral actions and consequences of donations vs. just their financial impact.
- Writing about a positive experience with a Hispanic vs. non-Hispanic individual.
- Writing about a time they helped others vs. a time they hurt others.
- Asking Christians to pray before making a charitable donation vs. just priming them to consider the victims.
- Items sold attached to a social cause vs. items sold with no social cause.
- Indicating whether they agree that a company has made progress on ethical/humanitarian initiatives.

Lots of ways to manipulate licensing

Moral & Charitable Behavior

- Recalling and describing cooperative/moral behavior vs. uncooperative/immoral behavior.
- Considering moral actions and consequences of donations vs. just their financial impact.
- Writing about a positive experience with a Hispanic vs. non-Hispanic individual.
- Writing about a time they helped others vs. a time they hurt others.
- Asking Christians to pray before making a charitable donation vs. just priming them to consider the victims.
- Items sold attached to a social cause vs. items sold with no social cause.
- Indicating whether they agree that a company has made progress on ethical/humanitarian initiatives.

Hiring Decisions & Bias

- Endorsing Obama for presidency before making a hiring decision.
- Endorsement of negative statements described as “most Blacks” vs. “some Blacks” or “some people.”
- Hypothetically hiring a Black vs. White candidate.

...across many different domains

Pro-Environmental Behavior

- Asking participants if they have completed frequent vs. infrequent pro-environmental actions.
- Told their pro-environmental behavior was similar to vs. dissimilar from others.
- Showing participants a green (vs. non-green) advertisement.

...across many different domains

Pro-Environmental Behavior

- Asking participants if they have completed frequent vs. infrequent pro-environmental actions.
- Told their pro-environmental behavior was similar to vs. dissimilar from others.
- Showing participants a green (vs. non-green) advertisement.

Consumer & Financial Decision-Making

- Choosing an eco-product vs. a non-eco-product.
- Giving an excuse for choosing to buy a treat vs. having no excuse.
- Earning income from a prosocial-oriented P2P platform.
- Imagining they had volunteered for community service vs. control condition.

...across many different domains

Pro-Environmental Behavior

- Asking participants if they have completed frequent vs. infrequent pro-environmental actions.
- Told their pro-environmental behavior was similar to vs. dissimilar from others.
- Showing participants a green (vs. non-green) advertisement.

Family Support & Social Responsibility

- Describing a time they provided unsolicited help to others vs. describing one in which they didn't.

Consumer & Financial Decision-Making

- Choosing an eco-product vs. a non-eco-product.
- Giving an excuse for choosing to buy a treat vs. having no excuse.
- Earning income from a prosocial-oriented P2P platform.
- Imagining they had volunteered for community service vs. control condition.

...across many different domains

Pro-Environmental Behavior

- Asking participants if they have completed frequent vs. infrequent pro-environmental actions.
- Told their pro-environmental behavior was similar to vs. dissimilar from others.
- Showing participants a green (vs. non-green) advertisement.

Family Support & Social Responsibility

- Describing a time they provided unsolicited help to others vs. describing one in which they didn't.

Consumer & Financial Decision-Making

- Choosing an eco-product vs. a non-eco-product.
- Giving an excuse for choosing to buy a treat vs. having no excuse.
- Earning income from a prosocial-oriented P2P platform.
- Imagining they had volunteered for community service vs. control condition.

Framing of Financial Sources

- Identifying the source of lottery money as coming from a state lottery agency vs. a casino.

Even more concept creep in the outcome measures...

- Justifying racial bias or discrimination after demonstrating support for diversity.
- Hiring a less qualified candidate based on racial moral credentialing.
- Identifying a white suspect less often after establishing moral credentials.
- Donating less money to charity after recalling past moral behavior.
- Feeling justified in not engaging in further charitable giving after prayer.
- Choosing a luxury product over a necessity after imagining volunteering.
- Justifying indulgent purchases after recalling past good deeds.
- Reporting less taxable income after earning from a prosocial platform.
- Feeling less obligated to act environmentally friendly after a prior green action.
- Choosing a less sustainable product after being reminded of past environmental behavior.
- Using environmental filters less often after thinking about prior sustainability efforts.
- Allowing a company to receive moral credit for minimal ethical progress.
- Feeling less compelled to support social causes after previous engagement.
- Justifying less family support or helping behavior after recalling past good deeds.

Philosophical Considerations

- *Should* we rely on a random sample of studies to assess the credibility of a field?
- Not all studies are created equal...
 - Vary in terms of their quality, rigor, journal prestige, citation count, etc.
- But quality is subjective
 - Random sampling helps with that
- Readers can also focus on the original studies they find most compelling

Implications for Moral Licensing Theory

- Low replication rate challenges empirical foundation
- Two possible interpretations:
 - Effects exist but are smaller than originally estimated
 - Many original findings were false positives
- More evidence for moral consistency than licensing in our replication

Recommendations for Future Research

- Increase statistical power
 - Larger samples
 - More efficient designs (e.g., within-subject)
- Focus on main effects over interactions
- **Return to basics**
 - Establish a single reliable demonstration before generalizing...
- Reduce methodological variability
- Systematically identify boundary conditions

Recommendations for Future Research

- Moral licensing theory might have intuitive appeal but limited empirical support
- One could argue that we found *some* evidence for both licensing and consistency effects
 - More for consistency
- Need for stronger empirical foundation before further generalization
- Opportunity to rebuild with more rigorous methods

Thanks!

Special thanks to the research team: Smriti Mehta, Tory Taylor, Paige Amormino, Julia Bottesini, Sandy Campbell, Dane Cannon, Beth Clarke, Randy Y. Gao, Stephanie Hong, Beidi Hu, Yunhao Huang, Yikun Jiang, Giulia Maimone, Jacob Miranda, Andres Montealegre, Aaron Nichols, Alex Park, William Ryan, Ameer Sarwar, Sarah Schiavone, Charlie Townsend, Ozge Ugurlu, Yixiang Xu, Don A. Moore, and Leif D. Nelson