Policy makers, employers, and insurers often provide us with financial incentives to take actions that are good for us or for society (e.g., energy conservation, healthy living, safe driving). Even though incentives may induce good behavior, these payments can cause individuals to view their actions less positively due to the perceived incompatibility between intrinsic motives and financial incentives (Newman & Cain, 2014; Benabou & Tirole, 2003). As a result, we hypothesize that individuals may prefer to eschew financial rewards for good behavior, even after rewards have been earned, when reminded of that behavior’s intrinsic value. Two preregistered studies — an incentivized online experiment (N=454) and a large field experiment (N=17,968) — provide evidence in support of this hypothesis: emphasizing the intrinsic rewards of a past action can lead individuals to forgo or donate a financial incentive they were promised for taking the action. We highlight a novel potential benefit of offering incentives for good behavior: if actors are given the opportunity to forgo some or all of their incentives, they can strongly self-report that they acted for the right reason and consequently improve their positive self-image. Indeed, we find that participants likely to receive a larger self-image boost from signaling intrinsic motives are more likely to give up their incentives when intrinsic rewards are emphasized.

**Hypotheses**

**Hypothesis 1**: Emphasizing the non-monetary, intrinsic benefits of incentivized good deeds makes actors more willing to forgo incentives earned for those same acts.

**Hypothesis 2a**: The effect will be stronger when the intrinsic motives in question are more consistent with the values actors hold and the self-image they hope to project (Prelec & Boden, 2003)

**Hypothesis 2b**: The effect will be stronger when actors have expended more effort on the incentivized behavior (Gneezy et al., 2012)

**Study 1: Online Experiment**

**Method**:
- 454 MTurkers paid to write a letter to a sick child
- Measured self-reported Authentic Pro-sociality (adapted from Barasch et al., 2014)
- Participants randomized into treatment, active control, or control conditions

**Example Letter Written by a Participant**:

Dear sweet child,

I hope that this letter finds you in good spirits during the holiday season. I know that it’s not fun to be in the hospital, but you can find joy and cheer right where you are. The people taking care of you in the hospital care very much for you, and so do I. You are very special, and you deserve to be happy. So, let’s see your beautiful smile, and hear how well you can laugh out loud. Take care, and keep smiling, and know that you are truly loved!

**Instructions**: “You can keep this bonus or you can [treat the letter writing practice you’ve received as your reward and] [treat the joy and hope you’ve spread as your reward] and choose to forgo some or all of your monetary compensation.”

**Results**:

<table>
<thead>
<tr>
<th>Incentive Earnings, in Dollars</th>
<th>Treatment</th>
<th>Active Control</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.29</td>
<td></td>
<td>$0.13</td>
<td>$0.13</td>
</tr>
</tbody>
</table>

22.4% gave up some or all of their incentives in the treatment condition, 12.0% in the active control condition, and 12.5% in the control condition.

**Study 2: Field Experiment**

**Method**:
- 17,968 24 Hour Fitness members recruited through StepUp, a 28-day digital rewards program
- Participants randomized into either the treatment or control condition

**Results**:

<table>
<thead>
<tr>
<th>Percent Donating</th>
<th>Treatment Effect (%) by Visit Tercile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td></td>
</tr>
<tr>
<td>1st Tercile</td>
<td>91.1%</td>
</tr>
<tr>
<td>2nd Tercile</td>
<td>90.3%</td>
</tr>
<tr>
<td>3rd Tercile</td>
<td>88.8%</td>
</tr>
</tbody>
</table>

2.0% - 3.0% more in the treatment condition.

**Conclusion**

The tension between monetary and intrinsic rewards leads people to forgo earned financial rewards for good behavior when reminded of that behavior’s intrinsic value. Consistent with our theorizing, this effect is stronger when the self-image gains of giving up incentives are larger.

We extend past research on mental money laundering by documenting a motivation laundering effect (Imas, Loewenstein, & Morewedge, 2019). Our findings suggest that organizations can encourage positive behavior change via incentive programs and leave actors feeling that their motivations were internal.

**Questions and comments welcome.** Email Erika Kirgiros at ekirgiros@wharton.upenn.edu

**Acknowledgements**: We would like to thank the Behavioral Change for Good team for their research support and assistance. We are grateful for support from the Wharton Behavioral Lab and the Wharton Risk Management and Decision Processes Center.