*Results*

The Luth survey company recruited participants to take part in the study. We set a target sample size of 1,500 participants, to be collected in two waves of 750. Luth collects data from participants in 20 minute sessions, and considers a complete survey to be one in which the participant finishes the entire 20-minute session. Because each session contains a number of studies, there is inevitably dropout between the first and last study in the session; this means that although ~1500 participants completed the total session, 2,942 participants completed the target study (“Fairness”). Similarly, although we refer to participants being split into two waves of 750, more than 750 participants completed the target study in each wave. The waves of the study are identified by when data collection began (3/21/18 for wave 1 and on 3/26/2018 for wave 2.)

**Wave 2 Results**

1,493 Wave 2 participants completed the target study and correctly answered the pre-registered attention check item. This item was placed at the end of the study and required participants to correctly recall the names of the two characters they read about in the study scenario.

This design was a 2 X 2 mixed design with one between subjects factor and a repeated-measure within subject factor. We calculated a difference score from time 1 to time 2 by subtracting the time 2 rating on the critical DV from the time 1 rating on the critical DV. We then conducted a two-sample t-test to determine if these difference scores differed between conditions. This t-test revealed a statistically significant difference *t*(1,491) = -9.5407.

**Wave 1 Results**

1,449 Wave 1 participants completed the target study and correctly answered the pre-registered attention check item. This item was placed at the end of the study and required participants to correctly recall the names of the two characters they read about in the study scenario.

This design was a 2 X 2 mixed design with one between subjects factor and a repeated-measure within subject factor. We calculated a difference score from time 1 to time 2 by subtracting the time 2 rating on the critical DV from the time 1 rating on the critical DV. We then conducted a two-sample t-test to determine if these difference scores differed between conditions. This t-test revealed a statistically significant difference *t*(1447) = -7.6607

**Total Sample Results**

2,942 participants completed the target study and correctly answered the pre-registered attention check item. This item was placed at the end of the study and required participants to correctly recall the names of the two characters they read about in the study scenario.

This design was a 2 X 2 mixed design with one between subjects factor and a repeated-measure within subject factor. We calculated a difference score from time 1 to time 2 by subtracting the time 2 rating on the critical DV from the time 1 rating on the critical DV. We then conducted a two-sample t-test to determine if these difference scores differed between conditions. This t-test revealed a statistically significant difference *t*(2,940) = -12.1939.