# Age and Framing Effects in the Balloon Analogue Risk Task (BART)

## Supplementary Materials

Adam T. Schulman, Amy W. Chong, and Corinna E. Löckenhoff

Department of Human Development

Cornell University

Correspondence: Corinna Löckenhoff, CEL72@cornell.edu

A. Participants

B. BART task

C. Supplemental results

## A. Participants

We pre-registered a sample size of 200 based on a-priori power analyses (conducted with G*Power; Faul et al., 2009) for a 2 (age: young vs. old) by 2 (frame: gain vs. loss) between-subjects ANOVA. For $\alpha = .05$ a sample size of 200 yields a power of .94 for detecting main and interaction effects when assuming medium effect sizes ($f = 0.25$).

We planned to recruit 80% of participants (n = 160) through the online survey company Qualtrics.com. Although this company follows industry guidelines (www.esomar.org) to verify participant demographics, monitor participation, and screen out low-quality responses, we opted to obtain the remaining 20% of participants (n = 40) through direct channels (i.e., existing participant databases and volunteer convenience sampling) where we could directly confirm participant identity and demographics.

Data collection deviated from the pre-registered plans in two respects. First, pilot testing revealed that including the working memory task would have exceeded the time limit for the Qualtrics survey panel. We therefore recruited an additional 47 participants through direct channels (including 44 participants who participated via group-based laboratory sessions at Cornell's Laboratory for Experimental Economics & Decision Research, LEEDR, http://leedr.dyson.cornell.edu/) to meaningfully examine working memory as covariate. Second, since invitations for participation were sent out in batches and response rates were slightly higher than expected, we obtained 7 additional participants in the Qualtrics sample.

The final sample consisted of three sources of participants: Participants recruited via Qualtrics who participated online (86 younger and 81 older participants), participants recruited through direct channels who participated online (20 younger and 23 older participants), and participants recruited through LEEDR (23 younger and 21 older participants). Preliminary

analyses examined the pre-registered effects of age and framing on BART responses for the full sample, the pre-registered sample, and the subsample of Qualtrics participants. The pattern of significant effects was the same across all three samples. Given the small size of the two other subsamples, they were not sufficiently powered to conduct separate analyses. However, when we analyzed the full sample and included sample source (Qualtrics vs. other) as a covariate, the pattern of significant effects also remained the same. Subsequent analyses therefore included the full sample.

We used two attention checks. The first was hidden among the future time perspective items and asked: “Please select number five to show you are paying attention”. The second was hidden in the numeracy assessment and asked: “At the FUN IN THE SUN music festival everybody gets a door prize. Out of 1,000 visitors how many are expected to get a door prize?”. Participants who failed the checks were automatically screened out by the survey system. In addition, we manually removed: 5 participants whose age (collected at the beginning of the study) did not match their birth year (collected at the end of the study), 1 participant who pumped only a single time for each balloon indicating they had misunderstood instructions, and 2 participants whose BART responses could not be linked to their questionnaire responses due to a technical error.

Participants compensation depended on the modality of recruitment: $20 for LEEDR laboratory participants, $5 for paid community participants, and no compensation for community volunteers. Qualtrics Recruitment Panel participants were compensated directly by Qualtrics.

## B. BART task

The automatic BART was implemented following Pleskac et al. (2008). Gain versus loss framing was implemented following Benjamin and Robbins (2007). Below, we describe each aspect of the task in more detail.

### Pilot testing

Pilot testing enrolled thirty-nine younger (aged 18 - 27 years) and 6 older participants (aged 62 - 71 years). Initial piloting stages recruited undergraduate students for iterative testing that focused on the general flow of the tasks in order to fine-tune the wording and location of the pop-up instructions. The final stage recruited a sample of older adults to confirm that the task flow and instructions were equally effective in that age group.

### Practice trials

Four practice trials allowed participants to familiarize themselves with the task. During the first two practice trials, pop-up directions explained the function of each button, the scale on the side, and the running total. The practice trials were programmed to expose the participant to a non-pop on the first balloon and a pop on the second balloon to ensure familiarity with either outcome. The last two practice trials did not have pop-up directions or pre-programmed outcomes.

### Main task

The main task consisted of 20 trials. Each balloon was set to explode between 1 and 50 pumps. On the first pump, the balloon had a 1/50 chance of exploding. On the second pump, it had a 1/49 chance of exploding, and so on until the 50th pump, which had a 1/1 chance of exploding. The balloon exploded after 25 pumps on average and explosion points were randomly distributed across the trials. Participants were not informed of these probabilities.

## Incentives

As an incentive to optimize performance, participants were told at the beginning of the task: "We will select a small percentage of participants at random and offer them extra payment, according to their total points, up to $25." After study completion, one participant from each sample (i.e., Qualtrics, LEEDR/paid community, and volunteer) was selected at random and received an online gift card based on the % of possible points achieved (e.g., 20% of points = $5).

## Framing

As seen in Figure 1 of the main document, the G-BART began with zero points and participants gained points with each pump (up to 1000 points over 20 trials). In contrast, the L-BART began with 1000 points and participants avoided losing points with each pump. Below, we show the instructions for each stage of the task with the gain versus loss versions shown in the following format: [*gain/loss*].

**General instructions.** "In this experiment you will blow up 20 balloons, one pump at a time. You will start the game with [*0 points/1000 points*]. Remember, balloons can pop. Some may pop after just one pump, while others may not pop until they are full. The explosion point varies across balloons. You will start each trial with [*0 points/-50 points*]. If you inflate the balloon without popping it, you will [*gain 1 point/ avoid losing 1 point*] for every pump you make. If you pop the balloon, you will [*not gain any points from that trial/fail to avoid the 50 point loss from that trial*]. The most you can [*gain/lose*] from each balloon is 50 points and the least is 0 points.

**Pop-up instructions during practice trials**. The interface for the task is shown in Figure 1 of the main manuscript. Instructions for the pump button read: "Click 'Pump' to tell the

computer the size of the balloon that you desire." Instructions for the slider read: "This shows how much [*you are gaining/loss you are avoiding*]". Instructions for the collect button read: "When you are done pumping, click 'Collect' to fill up the balloon to the desired size." Instructions shown in the center of the screen read: "The computer will fill the balloon to the desired size. You will see here if the balloon pops or not."

**Post-trial feedback**. If the participant did not pop the balloon the screen read: "Congratulation! You pumped the balloon [*x*] times and it did not pop. You [*earned/avoided*] [*x*] of the potential 50 point [*gain/loss*] in this balloon. Your total points are [*x*]." After a pop, the screen read: "Oh no! You pumped the balloon [*x*] times, but it popped after [*x*] pumps. You [*earned/avoided*] 0 of the potential 50 points [*gain/loss*] in this balloon. Your total points are [*x*]."

## C. Supplemental results

### Bayesian Analyses

Bayesian analyses examining the pre-registered hypotheses were conducted following van den Bergh et al. (2020) and used the default priors provided in JASP 0.16 and a matched-models approach (i.e., comparing models containing the effect to equivalent models without the effect). Results are summarized in Table S2.

For average pumps there was moderate evidence for the exclusion of the main effects for age ($BF_{incl}$ = 0.177) and framing ($BF_{incl}$ = 0.246) and strong evidence for the inclusion of the age by frame interaction ($BF_{incl}$ = 18.04). For total score, there was moderate evidence for the exclusion of main effects for age ($BF_{incl}$ = 0.222) and framing ($BF_{incl}$ = 0.177) and extremely strong evidence for the exclusion of the age by frame interaction ($BF_{incl}$ = 0.009).

Table S1

*Bayesian Analyses Examining Effects of Age and Frame on BART Pumps and Total Scores*

| | BART Average Pumps | | | BART Total Score | | |
|---|---|---|---|---|---|---|
| | *P(incl)* | *P(incl\|data)* | $BF_{incl}$ | *P(incl)* | *P(incl\|data)* | $BF_{incl}$ |
| Age group | 0.400 | 0.099 | 0.177 | 0.400 | 0.222 | 0.222 |
| Frame | 0.400 | 0.129 | 0.246 | 0.400 | 0.156 | 0.156 |
| Age Group * Frame | 0.200 | 0.346 | 18.040 | 0.200 | 0.009 | 0.009 |

*Note. P(incl)* = prior inclusion probability, *P(incl|data)* = posterior inclusion probability, $BF_{incl}$ = inclusion Bayes factor. Following a matched-models approach, models that include the effect are compared to equivalent models without the effect.

Table S2

Correlations Among Covariates and Dependent Variables

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Age group | - | | | | | | | | | | |
| 2. Female | .00 | - | | | | | | | | | |
| 3. White | .20** | .01 | - | | | | | | | | |
| 4. Income | .20** | -.14* | .13* | - | | | | | | | |
| 5. Education | .09 | -.04 | .01 | .28** | - | | | | | | |
| 6. Self-rated physical Health | .14* | -.03 | .07 | .25** | .18** | - | | | | | |
| 7. Self-rated emotional health | .40** | -.06 | .09 | .32** | .19** | .53** | - | | | | |
| 8. Current affect: valence | .35** | -.03 | .07 | .30** | .15* | .42** | .63** | - | | | |
| 9. Current affect: Activation | .19** | -.07 | .02 | .09 | .10 | .21** | .31** | .43** | - | | |
| 10. Future time perspective | -.36** | .05 | -.05 | .13* | .04 | .33** | .25** | .32** | .26** | - | |
| 11. Growth goals | -.25** | .03 | -.10 | -.14* | .00 | .20** | .06 | .10 | .23** | .43** | - |
| 12. Information goals | .10 | -.29** | -.10 | .15* | .10 | .02 | .17** | .14* | .05 | -.08 | -.14* |

Table S1 (continued)

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 13. Neuroticism | -.33** | .20** | -.03 | -.27** | -.09 | -.29** | -.52** | -.42** | -.21** | -.15* | .01 |
| 14. Extraversion | .22** | .06 | .15* | .25** | .15* | .20** | .34** | .27** | .32** | .22** | .13* |
| 15. Openness | .01 | .06 | -.03 | -.10 | .11 | .01 | -.04 | .03 | .03 | .07 | .17** |
| 16. Agreeableness | .19** | .09 | -.03 | .02 | .12 | .18** | .23** | .30** | .11 | .19** | .05 |
| 17. Conscientiousness | .34** | .06 | .18** | .24** | .15* | .42** | .41** | .47** | .30** | .24** | .07 |
| 18. Self-rated learning | -.14* | -.17** | .03 | .15* | .19** | .29** | .32** | .29** | .20** | .34** | .24** |
| 19. Self-rated memory | -.14* | -.10 | .11 | .09 | .18** | .36** | .30** | .21** | .19** | .28** | .15* |
| 20. Numeracy | -.19** | -.18** | .05 | .09 | .33** | .02 | -.04 | -.10 | -.09 | -.04 | .02 |
| 21. N-back | -.22* | .05 | -.06 | -.07 | -.11 | -.07 | -.01 | -.07 | -.16 | .15 | .15 |
| 22. Bart: Pumps | .05 | -.04 | .00 | .17** | .19** | .21** | .10 | .13* | .17** | .05 | .10 |
| 23. Bart: Total score | -.08 | -.07 | .06 | .10 | .12 | -.05 | .01 | .00 | .04 | .05 | .07 |
| 24. Bart: Pops | .08 | -.04 | -.05 | .13* | .17** | .24** | .09 | .10 | .12 | .02 | .06 |

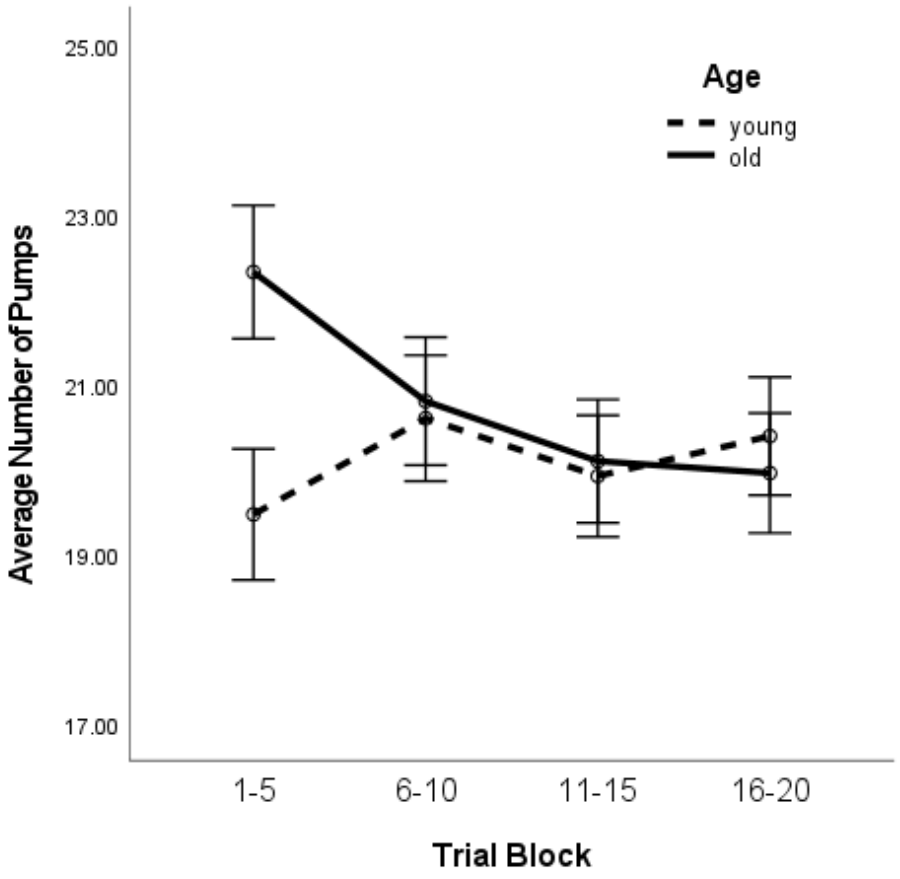
Table S1 (continued)

| Variable | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 13. Neuroticism | -.24** | - | | | | | | | | | | |
| 14. Extraversion | .09 | -.28** | - | | | | | | | | | |
| 15. Openness | -.15* | .07 | -.04 | - | | | | | | | | |
| 16. Agreeableness | -.12 | -.23** | .13* | .13* | - | | | | | | | |
| 17. Conscientiousness | .16* | -.33** | .27** | -.03 | .25** | - | | | | | | |
| 18. Self-rated learning | .16* | -.19** | .16* | .10 | .09 | .28** | - | | | | | |
| 19. Self-rated memory | .12 | -.19** | .20** | .02 | .05 | .19** | .63** | - | | | | |
| 20. Numeracy | .18** | .08 | -.03 | .00 | -.09 | -.05 | .08 | .04 | - | | | |
| 21. N-back | -.05 | -.02 | .00 | -.22* | -.27* | -.15 | .17 | .20 | .17 | - | | |
| 22. Bart: Pumps | .07 | -.07 | .09 | .08 | .02 | .12 | .10 | .06 | .11 | -.10 | - | |
| 23. Bart: Total score | -.07 | -.06 | .06 | .06 | -.04 | .01 | -.06 | -.02 | .17** | -.11 | .41** | - |
| 24. Bart: Pops | .15* | -.09 | .05 | .05 | .06 | .10 | .12 | .06 | .13* | .05 | .79** | -.10 |

*Note.* * $p < .05$, ** $p < .01$

Figure S1

*Age Differences in BART Pumping Rates Across Trial Blocks*




Note. Error bars show standard error of the mean.

## References

Benjamin, A. M., & Robbins, S. J. (2007). The role of framing effects in performance on the Balloon Analogue Risk Task (BART). *Personality and Individual Differences*, *43*(2), 221-230. https://doi.org/10.1016/j.paid.2006.11.026

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Pleskac, T. J., Wallsten, T. S., Wang, P., & Lejuez, C. W. (2008). Development of an automatic response mode to improve the clinical utility of sequential risk-taking tasks. *Experimental and Clinical Psychopharmacology*, *16*(6), 555–564. https://doi.org/10.1037/a0014245

van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E., Derks, K .. & Wagenmakers, E. (2020). A Tutorial on Conducting and Interpreting a Bayesian ANOVA in JASP. *L'Année psychologique*, 120, 73-96. https://doi.org/10.3917/anpsy1.201.0073