

EXPERIMENT #3 · MENTAL ACCOUNTING

# **Does Where Money Comes From Affect How Much of It People Are Willing to Spend?**

*A pilot study on mental accounting and windfall spending*

---

Vidit Chhajer

October 2026

## **Abstract**

This study examines whether the source of a person's money affects how much of it they are willing to spend, extending prior mental accounting research to an everyday online survey. Drawing on Thaler's (1985) mental accounting theory, we predicted that participants shown a \$10,000 tax refund as part of their \$90,000 would report a higher trip budget than those shown the same \$90,000 entirely as salary savings. Twenty-nine participants were recruited via Reddit, Instagram, and my personal network and were randomly assigned to one of two conditions: a tax-refund condition (\$80,000 in savings plus a \$10,000 unexpected tax refund in checking) or a savings-only condition (\$90,000 in savings). Each participant was given the same income, the same paid time off, the same total amount of money, and the same 8-day trip to Costa Rica. A one-tailed Welch's independent samples t-test revealed a statistically significant difference in mean trip budgets between the two groups ( $t = 2.488$ ,  $df = 12.95$ ,  $p = .0136$ ), with the tax-refund group reporting a mean budget of \$9,666.67 compared to \$5,558.82 in the savings-only group. Effect size was large (Cohen's  $d = .999$ ), and a 95% confidence interval for the difference in means excluded zero (\$539.97, \$7,675.73), consistent with the result. These findings are consistent with mental accounting and prior work by Arkes et al. (1994), suggesting that money which arrives unexpectedly is spent more readily than money that was earned and saved, even when the total is identical. Limitations of this study, including small sample size, convenience sampling, a skewed tax-refund group, and a possible anchoring confound, are discussed alongside directions for future research.

## Introduction

In theory, money is fungible. A dollar from a paycheck and a dollar from a tax refund buy the exact same things, so where a dollar came from shouldn't change how it gets spent. In practice, people rarely treat money this way. Instead, the mind sorts money into separate mental "accounts" based on where it came from and where it sits, and each account gets its own spending rules. This is mental accounting, formally developed by Richard Thaler in 1985. Simply put, your brain guards the money it worked for, and treats money that showed up unexpectedly as free to spend.

Tversky and Kahneman (1982) proved the real world implications of this. A sample was split into two conditions, with one group of people being told they lost a \$40 ticket they already paid for and the other group being told they lost \$40 cash earlier. Both group of people were asked if they'd buy a \$40 ticket at the entry to a movie theatre. A statistically significantly grater percentage of people who lost \$40 cash said the would buy a \$40 ticket than the ones who already lost a \$40 ticket. Both groups lost \$40 net, yet the way in which they lost it impacted how they reacted to it.

Much of the classic mental accounting research has relied on controlled lab settings and student samples. This study tests the effect on everyday people via an online survey, asking whether showing part of a person's money as an unexpected tax refund changes how much a general sample is willing to budget for a vacation. If mental accounting holds, participants shown the tax refund should report a significantly higher trip budget than those shown only salary savings, even though both groups have the exact same amount of money.

## Research Question

*Does where money comes from affect how much of it people are willing to spend?*

**Hypothesis**

Respondents shown a \$10,000 tax refund as part of their \$90,000 will report a higher trip budget, on average, than those shown the same \$90,000 entirely as salary savings, consistent with Thaler's (1985) mental accounting theory.

## Method

Participants were given a hypothetical situation of planning an 8-day, 7-night trip to Costa Rica and were shown a snapshot of their finances. In both groups the total amount of money was \$90,000, the monthly take-home pay was \$9,000, and there were 12 PTO days remaining, but the money was split across accounts differently.

### Group 1 saw this text:

*You are 26 years old, unmarried, and work full-time. You net \$9000 every month after expenses, and your savings account consists entirely of your own salary saved up. You recently got a huge, unexpected tax refund worth \$10,000, which deposited straight into your checkings account. Your 12 PTO (Paid time off) days expire on December 31 and do not roll over, so you wish to use them soon. You're going on an 8-day, 7-night trip to Costa Rica. What is your budget for this trip, in USD? What are you willing to spend out of your accounts to make this trip memorable yet financially feasible?*

### Group 2 saw this text:

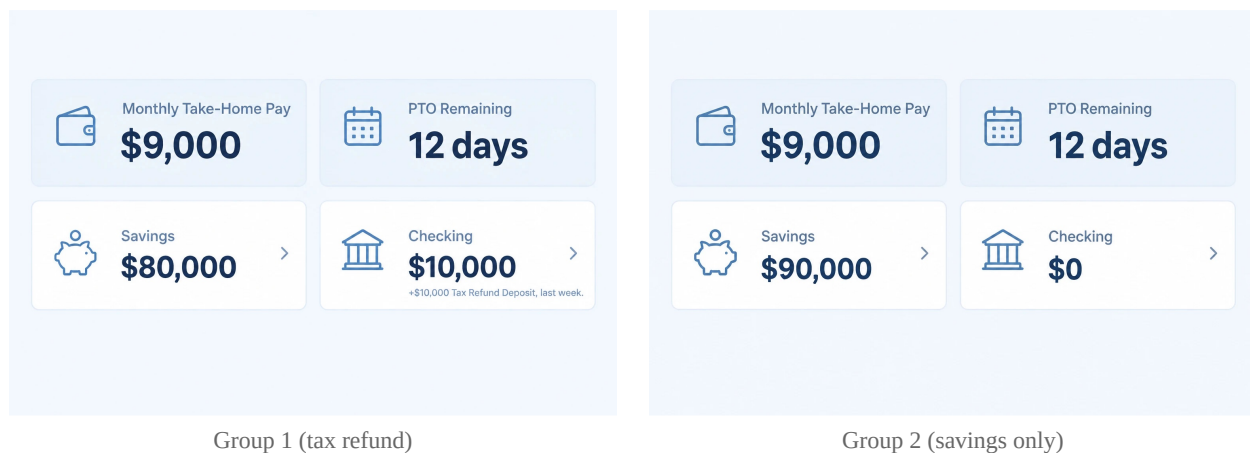
*You are 26 years old, unmarried, and work full-time. You net \$9000 every month after expenses, and your savings account consists entirely of your own salary saved up. Your 12 PTO (Paid time off) days expire on December 31 and do not roll over, so you wish to use them soon. You're going on an 8-day, 7-night trip to Costa Rica. What is your budget for this trip, in USD? What are you willing to spend out of your accounts to make this trip memorable yet financially feasible?*

Along with the text, each group saw an account dashboard. Group 1's dashboard showed \$80,000 in savings and \$10,000 in checking, with a note under the checking balance reading "+\$10,000 Tax Refund Deposit, last week." Group 2's dashboard showed \$90,000 in savings and \$0 in checking. Both dashboards showed the same \$9,000 monthly take-home pay and the same 12

PTO days remaining. All participants were also shown the same graphic of four trip budget levels: Budget (about \$3,000), Regular (about \$6,000), Luxury (about \$10,000), and Splurge (\$15,000 and up), each with the hotels, activities, and restaurants at that level.

**Figure 1**

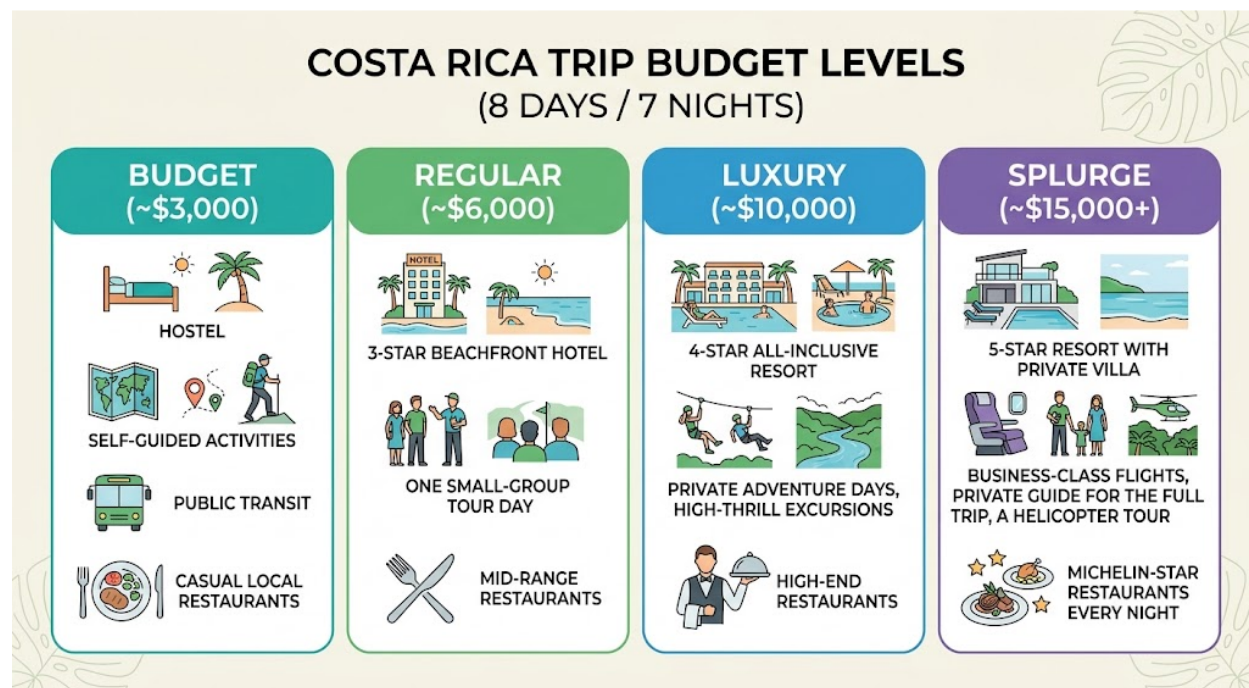
*Account dashboards shown to each condition.*



Note. Income, paid time off, and the \$90,000 total are identical. Only the split between accounts and the deposit note differ.

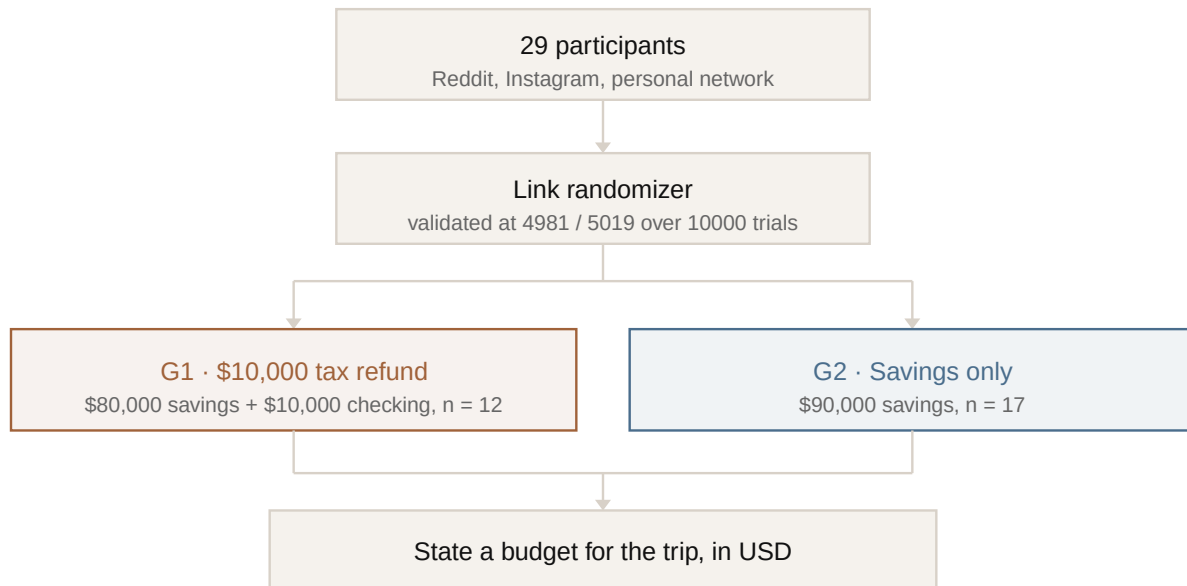
**Figure 2**

*Trip budget levels shown to both conditions.*



The study used two between-subjects treatment conditions. Participants in the tax-refund condition were told \$10,000 of their \$90,000 came from an unexpected tax refund sitting in their checking account. Participants in the savings-only condition were told all \$90,000 was their own salary saved up. Then all participants were asked for their trip budget in USD. Participants were randomized to treatments via a link randomizer which I designed. It essentially takes 2 links as inputs and randomly redirects to one of the two links. While testing the randomizer with  $n = 10000$ , I got 4981 redirects leading to link 1 and 5019 leading to link 2, essentially confirming the tool works properly, and any variance is random.

The survey ended with 29 participants that were recruited via Reddit posts in r/sample-size, r/surveyexchange, and r/behavioraleconomics (adhering to subreddit guidelines), an Instagram reel, and my personal network. All participants ended up with the same total amount of money; the manipulated variable was how \$10,000 of it was labeled and which account it sat in.

**Figure 3***Study design and treatment assignment.*

Note. The two vignettes differ by a single sentence. Every other word is identical.

## Results

G1 = group shown the \$10,000 tax refund (\$80,000 savings + \$10,000 checking). G2 = group shown savings only (\$90,000 savings).

**Table 1**

*Descriptive statistics for trip budgets by condition.*

|                    | G1 (\$10,000 tax refund) | G2 (Savings only)         |
|--------------------|--------------------------|---------------------------|
| Mean               | \$9,666.67               | \$5,558.82                |
| Median             | \$7,500                  | \$6,000                   |
| Standard deviation | \$5,482.76               | \$1,935.54                |
| Range              | \$17,000                 | \$7,000                   |
| IQR                | \$6,500                  | \$2,500                   |
| n                  | 12                       | 17                        |
| Outliers           | None                     | 1 outlier, value \$10,000 |

Note. Outliers identified using the 1.5x IQR rule. Budgets are in US dollars.

## Hypotheses and Test

$H_0$ : There is no difference in mean trip budgets between the tax-refund and savings-only groups.  $\mu_1 = \mu_2$

$H_1$ : Mean trip budget will be higher for the tax-refund group than the savings-only group.  $\mu_1 > \mu_2$

Alpha Level: .05

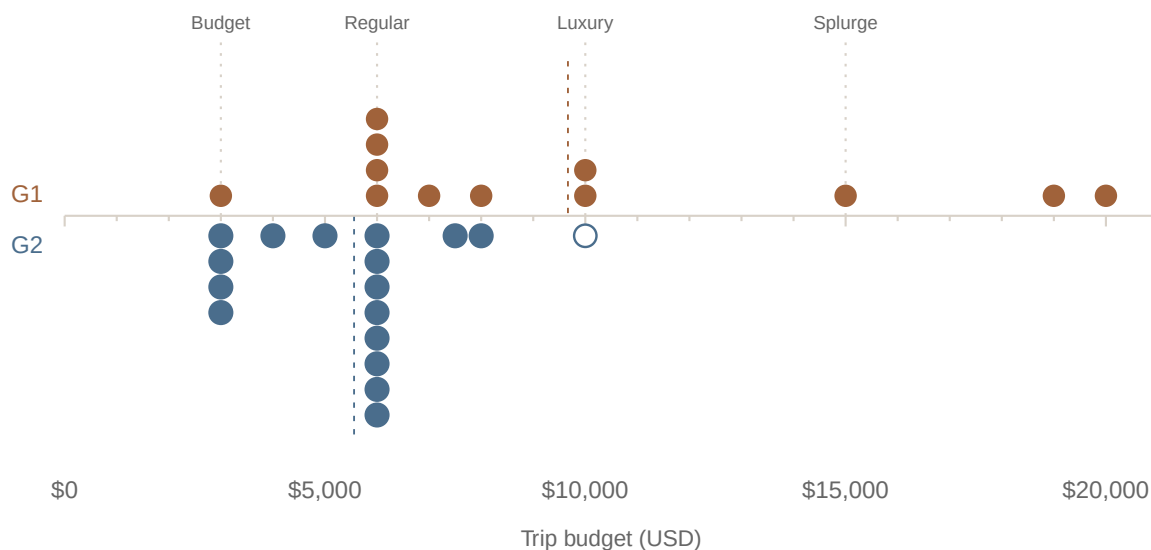
Name: one-tailed Welch's independent samples t-test

## Conditions

1. Random: Subjects weren't randomly chosen, rather it was a voluntary sample with some elements of a convenience sample. However, participants were randomly assigned to a treatment via a random link redirector — leading to a 12-17 split
2. 10%: Easily met,  $n = 12$  and  $n = 17$  is much lower than 10% of all potential vacation planners
3. Normal: Only weakly met. G2's responses are clustered tightly around \$6,000 with one outlier. G1's plot is right-skewed, with a few very high budgets (up to \$20,000) pulling the mean (\$9,666.67) well above the median (\$7,500), and  $n = 12$  is too small to make up for that. I proceeded with the test, but the p value should be read with caution

**Figure 4**

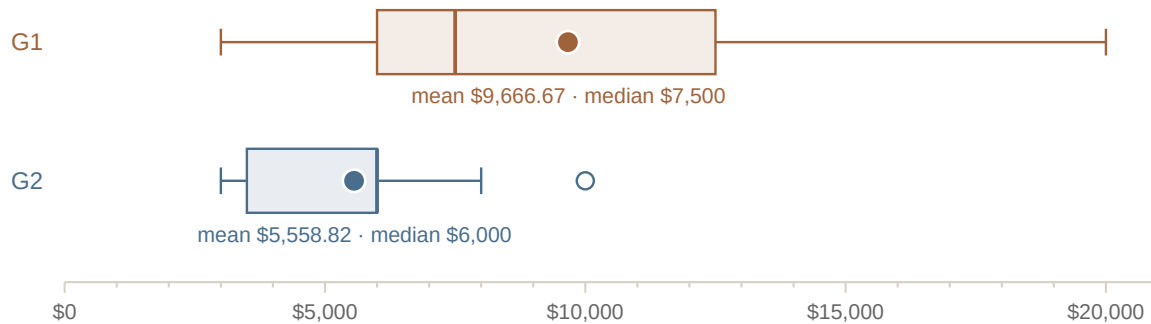
*Distribution of trip budgets by condition.*



Note. G1 stacks above the axis, G2 below. The hollow marker in G2 is the single outlier at \$10,000. Dashed verticals mark each group mean; dotted guides mark the four budget levels participants were shown.

**Figure 5**

Box plots of trip budgets by condition, with each group's mean (dot) and median (vertical rule).



Note. Boxes span the interquartile range. The means are \$4,107.85 apart; the medians are \$1,500 apart.

### Test Statistic

$t = 2.488$  with  $df = 12.95$

$p = .0136$

Since the calculated  $p$  value of .0136 is less than our stated alpha level of .05, we reject the null hypothesis. We found statistically significant evidence to suggest that the group shown the \$10,000 tax refund reported a higher average trip budget than the group shown savings only, consistent with mental accounting.

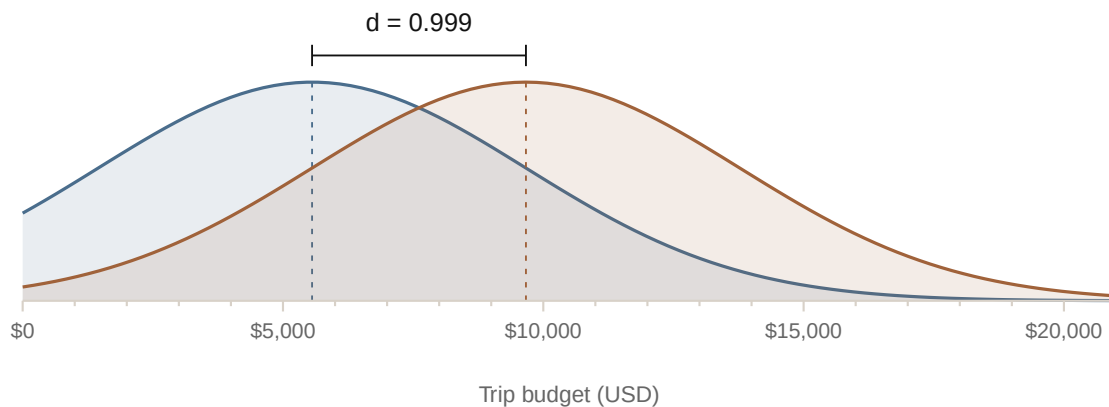
### Practicality

To assess the practical significance of the result, Cohen's  $d$  was calculated using the means and standard deviations of both groups. Cohen's  $d = 0.999$  using average variance formula (unpooled), which indicates a large effect by Cohen's (1988) conventions (small = 0.2, medium = 0.5, large = 0.8). This suggests that the difference in trip budgets between the two groups was not only statistically significant but also practically meaningful. One caution is that the medians (\$7,500 vs

\$6,000) are only \$1,500 apart, so a large part of the \$4,107.85 gap in means comes from a few very high budgets in G1.

**Figure 6**

*Effect size. The two conditions separated by 0.999 standard deviations.*



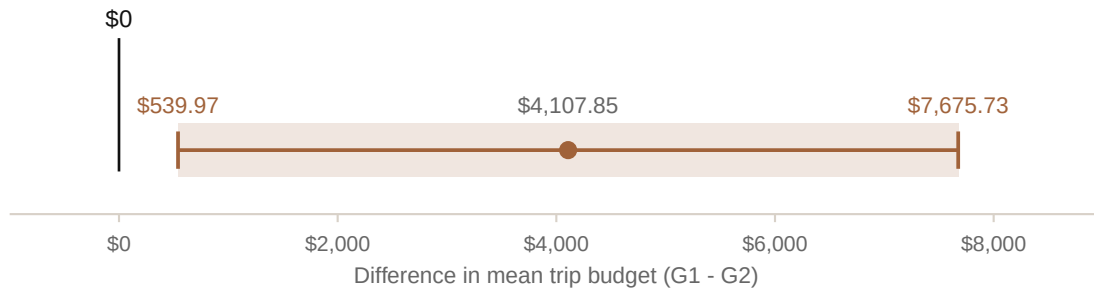
Note. Curves are drawn using the unpooled average-variance standard deviation on which Cohen's  $d$  was calculated.

### Confidence Interval

A 95% confidence interval for the difference in mean trip budgets (G1 - G2) was calculated as (\$539.97, \$7,675.73). This means we are 95% confident that the true difference in mean trip budgets between the tax-refund and savings-only groups falls between \$539.97 and \$7,675.73. 0 not being included within our interval is consistent with our decision to reject the null hypothesis. The interval is wide, which reflects how small the sample is.

**Figure 7**

*95% confidence interval for the difference in mean trip budget.*



Note. The interval lies entirely above zero.

## Discussion

My findings support the hypothesis. The group shown the tax refund reported a mean trip budget of \$9,666.67, while the group shown savings only reported a mean trip budget of \$5,558.82. Both groups had the exact same \$90,000, the same income, and the same trip, yet we found significant evidence that people budgeted more when part of that money was shown as an unexpected refund. The tax-refund group's mean budget also landed almost exactly on the size of the refund itself.

My findings are consistent with what Thaler described in 1985. He argued that people don't treat their money as one interchangeable pool, but instead sort it into separate mental accounts and make decisions inside each one. A \$10,000 refund sitting in checking seems to have been filed under "extra money," while the same \$10,000 inside a savings balance was filed under "money I worked for."

The findings also connect to Arkes et al.'s 1994 windfall experiments. Students who received money they weren't expecting bet more in a gambling game and spent more at a basketball game than students who knew the money was coming. My survey mirrors this. A trip budget was higher when part of the money was shown as an unexpected windfall, suggesting the effect extends from small amounts of real cash to large hypothetical budgets.

The real-world implications span a lot of domains. In a casino, gamblers treat their winnings as "house money" and bet it more freely than the cash they walked in with. During tax season, refunds get treated like a bonus, even though a refund is just the person's own paycheck money that was over-withheld all year. In retail, gift cards get spent on things people wouldn't pay cash for. In personal finance, people keep money in a low-interest savings account while carrying high-interest credit card debt, because the two sit in different mental accounts.

These are all examples where money gets treated differently because of its label, not its value. These are mental accounting in action.

### **Limitations**

1. I changed two things at once. In G1, \$10,000 was labeled as an unexpected tax refund and it was also moved from savings to checking. Both are mental accounting manipulations, but I can't separate how much of the effect came from the source of the money versus the account it sat in.
2. Anchoring is a competing explanation. G1 saw the number \$10,000 three times outside of the budget graphic (in the text, as the checking balance, and in the deposit note), while G2 never did. G1's mean of \$9,666.67 landing that close to \$10,000 could partly be an anchor on that number rather than mental accounting.
3. G1 was small ( $n = 12$ ) and right-skewed. The medians were only \$1,500 apart, so the gap in means was driven by a few very high budgets, and the normality condition was only weakly met. The p value is less trustworthy than it looks.
4. Because my Instagram account is framed as a behavioral economics account, and I posted in [r/behavioraleconomics](#) along with more general subreddits like [r/samplesize](#) and [r/surveyexchange](#), some participants likely went into the survey already expecting some kind of psychological trick.
5. If anyone took the survey twice and saw both treatments, it could contaminate responses.

### **Future Direction**

Ways this experiment could be improved are as follows. My sample size needs to be substantially larger, ideally upwards of 100 participants per group to improve statistical power and allow for more reliable replication. I would separate the two things I changed at once by running four groups (refund in checking, refund in savings, salary in checking, and salary in savings) to see whether the label or the location drives the effect. I would also use a refund amount that doesn't match any number on the budget graphic, like \$8,500, or remove the graphic entirely, to rule out anchoring. Other directions include testing whether the size of the windfall matters (\$1,000 vs \$10,000 vs \$50,000), comparing the tax refund against other sources of unexpected money like a bonus, a gift, or lottery winnings, asking participants to state their total money before giving a budget to see if making the full \$90,000 obvious weakens the effect, and running a Mann-Whitney U test alongside the t-test, since budget data tends to be skewed.

## References

- Arkes, H. R., Joyner, C. A., Pezzo, M. V., Nash, J. G., Siegel-Jacobs, K., & Stone, E. (1994). The psychology of windfall gains. *Organizational Behavior and Human Decision Processes*, 59(3), 331–347.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American*, 246(1), 160–173.
- Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science*, 4(3), 199–214.