

Azami Idrissi, Abdelkarim; Fanghella, Valeria; Gonzalez-Fernandez, Pedro; Hauser, Karl

Working Paper

A comment on "Predictable Effects of Visual Saliency"

I4R Discussion Paper Series, No. 321

Provided in Cooperation with:

The Institute for Replication (I4R)

Suggested Citation: Azami Idrissi, Abdelkarim; Fanghella, Valeria; Gonzalez-Fernandez, Pedro; Hauser, Karl (2026) : A comment on "Predictable Effects of Visual Saliency", I4R Discussion Paper Series, No. 321, Institute for Replication (I4R), s.l.

This Version is available at:

<https://hdl.handle.net/10419/343422>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

No. 321

I4R DISCUSSION PAPER SERIES

A comment on “Predictable Effects of Visual Salience”

Abdelkarim Azami Idrissi

Valeria Fanghella

Pedro Gonzalez-Fernandez

Karl Hauser

September 2026

I4R DISCUSSION PAPER SERIES

I4R DP No. 321

A comment on “Predictable Effects of Visual Salience”

Abdelkarim Azami Idrissi¹, Valeria Fanghella², Pedro Gonzalez-Fernandez³, Karl Hauser⁴

¹Lorraine University, Nancy/France

²EDHEC Business School, Roubaix/France

³Heidelberg University, Heidelberg/Germany

⁴Grenoble Alpes University, Grenoble/France

SEPTEMBER 2026

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the [ZBW – Leibniz Information Centre for Economics](https://www.zbw.eu/), and [RWI – Leibniz Institute for Economic Research](https://www.rwi-essen.de/), I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur
University of Ottawa

Jörg Ankel-Peters
RWI – Leibniz Institute for Economic Research

A comment on “Predictable Effects of Visual Salience”

Abdelkarim Azami Idrissi (Lorraine University)

Valeria Fanghella (EDHEC Business School, France)

Pedro Gonzalez-Fernandez* (Heidelberg University)

Karl Hauser* (Grenoble Alpes University)

Abstract

Li and Camerer (2022) examine whether attention automatically captured by visual features can influence economic decisions. In their first experiment, they find that visual salience affects people's choices between items, leading to systematic mistakes even when salience is unrelated to the items' actual value. First, we successfully computationally reproduce the main claims of the experiment, but uncover one minor coding error, which has no effect on the studies' main results. Second, using a larger, representative US sample (N = 418) we replicate both the significance and direction of the main estimates, finding that salience had an even stronger effect on economic choices. Furthermore, these replicated results withstand our robustness test.

*Corresponding authors: Pedro Gonzalez-Fernandez (pedro.gonzalez-fernandez@awi.uni-heidelberg.de) and Karl Hauser (karl.hauser@univ-grenoble-alpes.fr). All authors contributed equally to the paper. The authors acknowledge funding from The Institute for Replication I4R, and are thankful to Paolo Crosetto and Abel Brodeur for organising the Replication Games workshop in Grenoble (2024) and making this project possible.

1. Introduction

In the 2022 paper “Predictable Effects of Visual Saliency in Experimental Decisions and Games”, published in the *Quarterly Journal of Economics*, X. Li and C. F. Camerer investigate whether the visual saliency within choice situations may help explain economic decisions. Across three studies and various economic games, they test whether visual saliency may interfere with the economic choices made within these games. This replication focuses on the first study: “Saliency and Induced Value in Visual Fruit Displays”.

1.1 Experimental Protocol and Treatments

The Game. The first study uses a simple decision setting. Participants have to choose between two groups of fruits (a mix of apples and oranges) displayed on the right and left side of an image. Each type of fruit has a predetermined monetary value and the fruits are arranged in a way that one of the two food heaps has an objectively higher monetary value. Participants earn money based on the monetary value of the fruit heaps they chose. They make 20 choices in total and have a 10-second time limit per image to make their choice.

Assessing visual saliency and congruence. The study manipulates two types of saliency: goal-driven saliency and stimulus-driven saliency. “Goal-driven saliency” is the type of saliency induced by the monetary value of the fruit heaps, while “stimulus-driven saliency” is the attention captured only by eye-catching visual features such as color, contrast or size of the fruit heaps. To this end, each of the fruit heaps within the 20 images was analyzed using a Saliency Attentive Model (SAM) Algorithm. The SAM provides an estimate for the difference in stimulus-driven saliency between both sides of the image. This saliency score is computed between 0 and 1. Finally, half of the images are labeled as “congruent”, because they are constructed in a way where the visually more salient fruit heap is also the monetarily more valuable heap, while the other half of the images are labeled as “incongruent”.

Choice Accuracy. The dependent variable is accuracy, coded as a binary variable that indicates whether (1) or not (0) the participant chose the heap with the highest monetary value.

1.2 Participant population

97 study participants were recruited through Prolific. The prescreening criteria include screening out participants with less than 70% prior approval rate on Prolific, as well as those participants using tablets or smartphones. Furthermore, five comprehension questions we used to screen out inattentive participants.

1.3 Time Period

The experiment was conducted online, but the specific time period for data collection was not detailed in the summary.

1.4 Main Scientific Claims

The experiment shows that visual salience significantly influences decision-making, often leading participants to choose more visually prominent but lower-value options, especially under time pressure. When visual salience aligns with value, accuracy improves, but when it conflicts, mistakes increase. The study highlights the difficulty of overriding automatic, stimulus-driven attention, even when it leads to suboptimal decisions (Appendix I).

1.5 Methodology and Main Findings

To analyze whether salience influences accuracy, the authors estimate a logistic regression in which the dependent variable y_{ij} equals 1 if participant i chose the more valuable set in image j . The key independent variables are: congruency $(S_j^H - S_l^L)$, defined as the difference between the maximum salience level of the more valuable set and that of the less valuable set (positive when the higher-value set is also more salient); the absolute value

difference $abs(V_j^L - V_j^R)$, that is, the induced value gap between the left and right sets; and their interaction term. Additional covariates $(X)_i$ include participants' education, gender, income, and self-reported fruit preference. The estimated model is:

$$y_{ij} = \beta_1(S_j^H - S_j^L) + \beta_2 abs(V_j^L - V_j^R) + \beta_3((S_j^H - S_j^L) * abs(V_j^L - V_j^R)) + \beta_4 X_i + \varepsilon_{ij}$$

The authors find that both congruency and the absolute value difference are significantly associated with accuracy, with similar effect sizes. However, two boundary conditions emerge where the salience effect disappears: when the value difference is large (accuracy is around 94% regardless of congruency) and when choices are untimed (accuracy reaches ceiling levels in both congruent and incongruent trials).

2. Computational Reproducibility

We first test the computational reproducibility using the replication packaged provided here:

<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/9LCYK>

The cleaning codes were provided in the replication package as well as the analysis data. We successfully computationally reproduced all the main results (i.e., Table 3 in Appendix I) from the full data.

	Fully	Partial	No
Raw data provided	x		
Cleaning code provided	x		
Analysis data provided	x		
Analysis code provided	x		
Reproducible from raw data	x		
Reproducible from analysis data	x		

We record one minor coding error that we uncovered while reproducing the study. When attempting to generate a summary table using the “stargazer” function in R, the following error appeared “Error in if (is.na(s)) { : the condition has length > 1” which is due to a misstep within the “stargazer” function, where it attempts to evaluate a condition that unfurls into multiple truths rather than a singular certainty. An additional detail is related to the name of the dataset, which was “controls” on the R code and “fruitDatasets” in the replication package.

However, the error mentioned above did not prevent us from reproducing the results for the regression displayed in Table 1. The structure of the table is the same as in the original study. We find that the point estimates are almost identical in sign, magnitude, and statistical significance.

2.1 Discrepancies Between Pre-registration Plan and Article

The authors state that their study was pre-registered, but the pre-registration is not publicly available. We contacted them to request access but did not

receive a response. The authors mention that “*Ninety-seven participants did [...] a study on Prolific [...], following a preregistration process on the Open Science Foundation website (OSF)*” but no pre-registration was provided.

3. Replication with New Data

We now report the results of our replication, conducted with a new and larger sample of participants. We first developed a Pre-Analysis Plan (see Appendix III) that was sent to the *Institute for Replication* (I4R) for ethical approval, which was granted in 2025.

3.1 New Data collection and differences to the original study:

Our replication differs from the main study in three main ways.

First, we increased the sample size. Funding from I4R allowed us to recruit 418 participants via Prolific from a representative U.S. sample, while still mirroring the two screening conditions described in the original study.

Second, we relied on the images provided in the replication package, even though it was unclear which specific subset had been used in the original analysis. As different files gave different indications, we decided to use the subset outlined in files “*fruits.csv*” and “*saliency_values.csv*”. As a robustness check, we also conducted our analysis excluding those 5 conflicting images from the data set. The main results replicate in both specifications.

Finally, we wrote our own instructions and comprehension questions for participants (see Appendix II) as the exact instructions used in the original study were not provided. Some discrepancies from the original instructions may therefore exist.

3.2 Results of the Replication

We successfully replicate the main results of the experiment.

Table 1 reports our replication estimates, as reported in Li and Camerer (2022) (Table 1 p. 1862). In both studies, accuracy in the value-based choice task is modeled as a function of congruency between salience and value, the absolute value difference, and their interaction. We find that congruency strongly increases accuracy: our coefficients range from 2.54 to 3.90 ($SE \approx 0.11$ to 0.16), compared to 0.83 to 1.26 ($SE \approx 0.29$ to 0.41) in the original study. Thus, the replication confirms the direction of the effect but yields substantially larger magnitudes, suggesting that in our sample, salience has an even stronger influence on correct choices. Consistent with this, mean accuracy was 85% versus 79% for congruent and incongruent images in the original study, compared to 59% versus 32% in ours. Hence, even though participants in our sample performed worse overall, the relative influence of salience on accuracy was considerably stronger.

The effect of absolute value difference is likewise positive and highly significant, consistent with larger value gaps making the correct option easier to identify. However, the estimated coefficients are notably smaller in the replication (≈ 0.25 – 0.30) than in the original study (≈ 0.77 – 0.80). For the interaction term, the original study reported a negative but statistically insignificant coefficient (-0.55 , SE 0.63), while our replication yields a large and highly significant negative estimate (-1.27 , SE 0.10). This provides stronger evidence that salience exerts the greatest impact when value differences are small, and that its influence diminishes as the task becomes easier. Finally, the replication is based on a substantially larger sample ($\approx 7,300$ vs. $\approx 1,300$ observations), which improves precision and likely accounts for the narrower standard errors and the statistical significance of the interaction effect.

Overall, the replication confirms the patterns found in Li and Camerer (2022). That is, congruency and value differences independently improve accuracy, and that salience is most relevant when value signals are weak. Additionally, we highlight differences in effect size magnitudes across samples, which may

partly reflect differences in sample size and composition, as our replication relies on a larger, nationally representative U.S. sample.

3.3. Robustness check

As a robustness test, we re-estimate the main specification after excluding the five images for which the replication package provided conflicting information about their inclusion in the original analysis (see Table 2). This restriction leaves a reduced sample, but ensures that our results are not driven by the specific image set.

The results, reported in Table 2, are highly consistent with our baseline findings. The coefficient on *Congruency* remains positive, large, and highly significant across all specifications (≈ 1.86 – 3.28 , $SE \approx 0.10$ – 0.16), confirming that accuracy is substantially higher when salience and value are aligned. Similarly, the coefficient on the absolute value of the induced value difference between left and right sets *abs(valueDiff)* is positive and significant (≈ 0.28 – 0.31), indicating that larger value differences facilitate correct choices. The interaction between *Congruency* and *abs(valueDiff)* continues to be negative and significant (-1.43 , $SE 0.11$), suggesting that the effect of congruency diminishes when value differences are large. Thus, our main conclusions are robust to the exclusion of images used in training, reinforcing the stability of the replication.

4. Conclusion

Our replication of Li and Camerer (2022) fully reproduces their main results. We confirm that visual salience systematically shapes economic choices, particularly when value differences are small, and that both congruency and value differences independently improve accuracy. Moreover, the effects remain robust across alternative specifications and image subsets. Taken together, these findings provide strong and consistent evidence that perceptual salience influences decision accuracy in value-based choice tasks

References

- Li, X. and Camerer C. F. (2022), Predictable Effects of Visual Saliency in Experimental Decisions and Games, *The Quarterly Journal of Economics*, 137(3), pp. 1849 - 1900. 10.1093/qje/qjac025

Tables

Table 1: Replication results of Study 1 in Li and Camerer (2022). Clustered std. errors.

	Dependent variable: Accuracy (0,1)				
	(1)	(2)	(3)	(4)	(5)
Congruency	2.67*** (0.11) t = 25.13	2.54*** (0.11) t = 24.08	2.70*** (0.11) t = 25.18	2.58*** (0.11) t = 24.13	3.90*** (0.16) t = 24.20
abs(valueDiff)		0.30*** (0.02) t = 16.98		0.30*** (0.02) t = 17.22	0.25*** (0.02) t = 14.01
Congruency:abs(val ueDiff)					-1.27*** (0.10) t = -12.505
Constant	-0.17*** (0.02) t = -7.73	-0.56*** (0.03) t = -20.02	-0.36*** (0.08) t = -3.61	-0.77*** (0.08) t = -12.14	-0.70*** (0.08) t = -10.08
Covariates	No	No	Yes	Yes	Yes
Observations	7,441	7,441	7,341	7,341	7,341
Log likelihood	-4737.004	-4595.976	-4666.765	-4524.307	-4431.805
Akaike inf. crit.	9478.007	9197.952	9349.531	9066.615	8883.609

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 2: Robustness Check — Accuracy Models without Five Images

	Dependent variable: Accuracy (0,1)				
	(1)	(2)	(3)	(4)	(5)
Congruency	1.88*** (0.10) t = 18.04	1.86*** (0.10) t = 17.30	1.90*** (0.10) t = 18.02	1.88*** (0.10) t = 17.28	3.28*** (0.16) t = 18.04
abs(valueDiff)		0.30*** (0.01) t = 16.72		0.31*** (0.01) t = 16.97	0.27*** (0.01) t = 14.91
Congruency:abs(valueDiff)					-1.43*** (0.10) t = -1.43
Constant	-0.06*** (0.02) t = -2.63	-0.51*** (0.03) t = -16.06	-0.04*** (0.08) t = -0.84	-0.50*** (0.08) t = -9.22	-0.45*** (0.08) t = -2.63
Covariates	No	No	Yes	Yes	Yes
Observations	5524	5524	5451	5451	5451
Log likelihood	-3677.16	-3537.32	-3625.12	-3484.14	-3373.65
Akaike inf. crit.	7358.32	7080.63	7266.24	6986.27	6767.29

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Appendix

Appendix I – The original results (Li & Camerer 2022, p. 1862)

Table 3: Original results of Study 1 in Li and Camerer (2022). Clustered std. errors.

	Dependent variable: Accuracy (0,1)				
	(1)	(2)	(3)	(4)	(5)
Congruency	0.83*** (0.32) t = 2.59	0.90*** (0.29) t = 3.11	0.89*** (0.33) t = 2.65	0.97*** (0.31) t = 3.15	1.26*** (0.41) t = 3.05
abs(valueDiff)		0.80*** (0.23) t = 3.54		0.80*** (0.23) t = 3.45	0.77*** (0.23) t = 3.39
Congruency:abs(valueDiff)					-0.55*** (0.63) t = -0.86
Constant	1.54*** (0.10) t = 15.92	0.78*** (0.17) t = 4.60	1.99*** (0.61) t = 2.73	1.25** (0.57) t = 1.79	1.26** (0.57) t = 1.81
Covariates	No	No	Yes	Yes	Yes
Observations	1,382	1,382	1,307	1,307	1,307
Log likelihood	-644.7	-591.8	-607.5	-556.7	-556.3
Akaike inf. crit.	1,293	1,189	1,239	1,139	1,141

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Appendix II – Instructions and Comprehension Questions

II.A. Instructions

You are going to play a game, called the **Fruit game**.

The game consists of choosing between two piles of fruits.

A pile of fruit consists of apples and oranges. **Apples and oranges have different monetary values.**

Your objective is to **choose the pile with the highest monetary value, as it may affect your bonus payment.**

You will play 25 rounds of the game, with different piles of apples and oranges.

Pay close attention to the instructions, as you will have to answer correctly some comprehension questions to be allowed to participate.

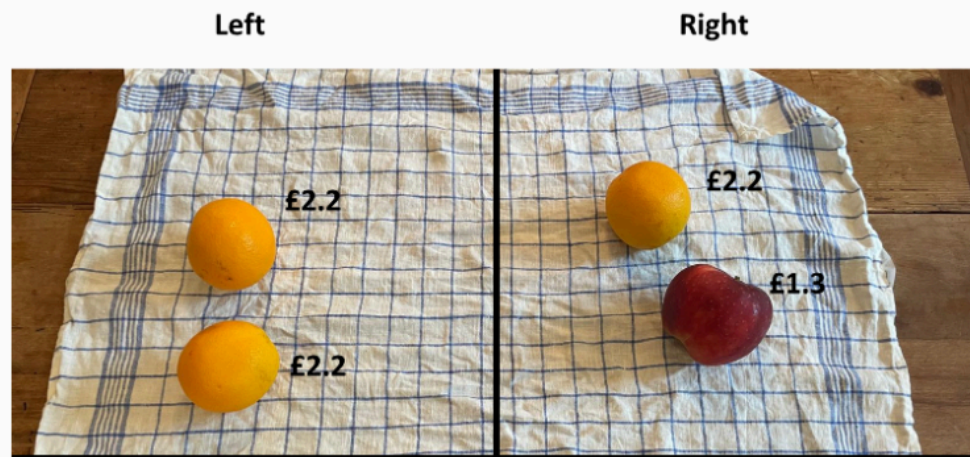
In the next page you will see an example.

Page break

In the Fruit game, the value of apples and oranges is **always** as follows:

- **An apple is worth £1.3**
- **An orange is worth £2.2**

Example:



Left Pile Worth: £2.2 + £2.2 = £4.4 Right Pile Worth: £2.2 + £1.3 = £3.5

Thus, in this example, the **left pile has a higher value than the right pile.**

-----Page break-----

In addition to your participation payment, you may receive an additional bonus payment that will depend on **your** choices. One of your choices will be randomly selected. You have a 10% chance of receiving the monetary value of that choice. Therefore, it is in your best interest to think carefully at each choice you make during the Fruit game.

-----Page break-----

II. B Comprehension questions

To ensure you have understood the fruit game, please answer the following 5 questions. If you give more than 1 wrong answer, you will be excluded from the survey and will receive no participation payment. Pay therefore attention while answering them.

You can find the comprehension questions below. **Remember that Apples are worth £ 1.3 and Oranges are worth £ 2.2**

What is your objective in the fruit game?

- ☐ Select the pile of fruits with the highest monetary value
- ☐ Select the pile of fruits with the lowest monetary value
- ☐ Select the pile of fruits with more apples

Which of the following statements is true?

- ☐ Your choices in the fruit game can never have consequences on your bonus payment
- ☐ Your choices in the fruit game can have consequences on your bonus payment

Suppose that:

- The left side pile is composed of 2 Apples
- The right side pile is composed of 2 Oranges

What is the monetary value of the left-side pile?

Suppose that:

- The left side pile is composed of 2 Apples
- The right side pile is composed of 2 Oranges

Which pile is worth more?

☐ Left Pile

☐ Right Pile

Suppose that:

- The left-side pile is composed of 4 Apples and 1 Orange
- The right-side pile is composed of 1 Apple and 3 Oranges

Which pile is worth more?

☐ Left Pile

☐ Right Pile

Appendix III – Pre-Analysis Plan

Introduction

This replication article focuses on the paper “Predictable Effects of Visual Salience in Experimental Decisions and Games” by Li & Camerer (2022) published in The Quarterly Journal of Economics. We focus on Study 1, as it “measures[s] the empirical importance of visual salience in a simple [and playful] setting that is lifelike” (p. 1859). It appears to us to be the most relevant and most easily replicable experiment compared to the other two studies, which require complex interactions among participants. We will replicate only the main condition, which includes a time limit. In this within-subject experiment, participants are shown images featuring fruits displayed on the two sides (L and R) of the screen. Based on the fruits displayed, each side of the screen has a specific monetary value. Participants are tasked with selecting the side of the image that represents the highest value. The observed variable is the accuracy of correctly choosing the highest value side. The within-subject variation includes the congruence between monetary value and visual salience of the two sides for every image. Our analysis will primarily examine the main effects of congruency and the absolute differences in monetary value on accuracy, as well as their interaction.

Hypothesis to replicate

H1: There is a positive effect of congruence (*the same side of the image has the highest saliency score AND the highest monetary value*) on accuracy.

H2: This positive effect remains even when taking into account the absolute monetary value difference between the two sides of the screen.

H3: The interaction between congruence and absolute monetary value is not statistically significant.

Criteria for replication

The criteria for replication are an effect in the same direction as the original study and a p -value < 0.05 in a two-sided test for significance of coefficients.

Power analysis

The original study has a small sample size ($N=97$) with repeated measures at the participant level (20 choices per participant), and reports results significant at the 0.01 level. In our replication, we will maintain the same number of repeated observations ($N=20$) but increase the sample size to 400 participants. This larger sample will enhance statistical power while adhering to budget constraints.

Consequently, we are confident that our sample will enable us to have 90% to detect at least 67% of the original effect size, as required by replication guidelines.

Sample

As in the original study, we will recruit participants via the Prolific platform.

Inclusion criteria

Following the original paper, we select only respondents with a prior Prolific approval rate of 70% or above.

Although it is not specified in the paper, we decide to restrict our analysis to UK participants who are native speakers in English.

Participants are allowed to do the experiment only once.

Exclusion procedure

Although not specified in the original paper, we include these exclusions criteria:

- Participants using tablets or phones.
- Participants failing the CAPTCHA at the beginning of the survey
- Participants failing an attention check after the task is complete

As in the original paper, we will remove participants that fail more than 1 of the 5 comprehension questions of the experimental instructions.

The final sample size for the replication will take into account any exclusions of participants before and/or during the experiment.

Materials

We will use the same material as in the original study, adapting the qsf files posted by the original authors. This includes the images and incentive structure used in the experiment as well as the SAM- generated salience estimates used for the analysis. However, experimental instructions (or comprehension checks) are not available; thus, we wrote them.

Procedure

We will closely follow the procedure of the original experiment. The following summary of the experimental procedure is therefore largely based on the description of the experiment in the article (p. 1860-61).

Participants will first provide informed consent and then complete a Captcha.

They will then see the instructions of the experiment; answer 5 comprehension checks of the instructions.

They will then do 25 rounds of the task. The first 5 are for the purpose of training and have no time limit. They are not included in the hypothesis testing. The remaining 20 represent the main task and are used for hypothesis testing have a time limit of 10 seconds. All the 25 rounds count for incentivization.

In the task, each image contains two fruit sets presented on the left and right parts of an image. Each fruit type (i.e. apples or oranges) has a unique, predetermined induced monetary value (1.3£ for apples and 2.2£ for oranges, as in the original study) that subjects are informed since the beginning of the study. Participants have to select the side with higher monetary value (L or R). Once they click on the R/L button, they are automatically redirected to the next image. In the main task, if they do not answer within 10 seconds, they are redirected automatically to the next image; their choice for that image is coded as missing value. We will also measure how much time they spend on each image (reaction time). The images are the same for all participants, but the order of the images is randomized.

At the end of the task, we will include an attention check and then collect socio-demographic characteristics of respondents: education, gender, income, self-reported fruit preference (for apples, oranges or equal preference for both fruits).

We estimate that the completion time will be between 10 and 15 minutes.

The analysis will be performed as in the original article. The main specification is logistic regression with repeated observations, so we will use robust standard errors clustered at individual level. The main variables are: (i) congruence, which is equal to 1 if the side with highest monetary value is also the one with highest visual salience, 0 otherwise; (ii) absolute difference in monetary value among the two sides (for each side, this is calculated as the sum of the values of the fruits on each side of pictures). We will follow the specification of the original paper: (i) only with congruence; (ii) with congruence and absolute difference; (iii) interaction between the two. And then repeat specifications (i), (ii), (iii) with additional controls (analysis (iii) with covariates is not part of the original analysis). In addition to the analyses in the paper, we will also check for order effects in our analysis (i.e., if accuracy improves over rounds).

Subject payment:

The budget allocated to this replication study by the *Institute for Replication* amounts to 2500\$. We estimate the total cost of our replication to amount to 1600£ (around 2,000\$ date 26.06.2024). This total cost is composed of the participation fee, the bonus payment as well as the 25% fees that the Prolific platform takes. It breaks down as follows:

We plan on having a total of 400 participants. We estimate the experiment to take on average 15 mins. Thus, we anticipate a show-up fee of 2.5 £ per participant (hourly wage of 10 £/h), amounting to 400 *

$$2.5 \text{ £} * 1.25 = 1250 \text{ £}.$$

The incentivization mechanism for the experiment involves selecting one round per participant, with a 10% probability of payment for the choice made in that round. Given that there are 400 participants, each making 20 choices, one choice per participant is randomly selected, resulting in a total of 400 selected choices. Out of these, 10% are implemented, leading to an expectation of 40 paid choices. We estimate the average monetary value for each side to be £7. Consequently, the total budget allocated for incentivization is calculated as $40 * £7 * 1.25$, amounting to £350.

Thus the expected budget amounts to $1250 \text{ £} + 350 \text{ £} = 1600 \text{ £}$.