Received: 7 April 2020 | Revised: 2 February 2021 | Accepted: 14 April 2021

DOI: 10.1111/1477-9552.12443

ORIGINAL ARTICLE

JAE Journal of Agricultural Economics

Do virtual reality experiments replicate projection bias phenomena? Examining the external validity of a virtual supermarket

Tiziana de-Magistris[1] | **Belinda López-Galán**[2] | **Petjon Ballco**[1]

[1]Agrifood Research and Technology Centre of Aragón (CITA, Zaragoza, Spain

[2]Departamento de Análisis Económico, Facultad de Economía y Empresa, Universidad de Zaragoza, Zaragoza, Spain

Correspondence
Tiziana de-Magistris, Agrifood Research and Technology Centre of Aragón (CITA), Zaragoza, Spain.
Email: tmagistris@cita-aragon.es

Funding information
Spanish National Institute of Agricultural and Food Research and Technology, Grant/Award Number: 2013-0092-00-00

Abstract

We examine whether projection bias is also present in virtual environments. Two hypotheses were tested using a between-subject experiment with three conditions: (i) experimental context (*virtual* vs. *real*); (ii) consumption periods (*today* vs. *tomorrow*); and (iii) appetite level (*hungry* vs. *satiated*). An exact replication of the virtual supermarket was performed in a real supermarket setting to test for robustness. The findings indicate the presence of a projection bias in the virtual reality setting. The robustness test used to compare a virtual and a real purchase at the supermarket showed that subjects in the virtual environment behave similarly to those in a real supermarket. These results validate our findings and highlight virtual reality use as a new tool for investigating consumer behaviour in food research.

KEYWORDS

consumers, external validity, projection bias, real supermarket, virtual reality

JEL CLASSIFICATION

D12; D91; E21; Q18

1 | INTRODUCTION

Individuals continuously anticipate intertemporal future behaviours for a variety of choices (Loewenstein et al., 2003). Such decisions depend on prediction accuracy. Standard economic approaches assume that people accurately predict future associated potential changes in their preferences (Loewenstein et al., 2003). However, psychological evidence suggests that economic decision models are inaccurate because they do not consider that individuals systematically fail to predict their future utility (Loewenstein & Schkade, 1999), that is, they suffer from projection bias.

Loewenstein et al., (2003) define projection bias as a general bias that arises whenever preferences change over time, that is, people incorrectly project their current state into the future. These authors demonstrate that people understand that their preferences can change in the future, but they systematically over-estimate the extent to which their future preferences will resemble their current ones. This is because they ignore the fact that their present state affects their assumptions about future preferences. The projection of these biased decisions is highly influenced by 'visceral factors' such as fear and anger (Loewenstein et al., 2003). Our focus here is on hunger. Several studies have tested projection bias in consumer behaviour research (Acland & Levy, 2015; Bienenstock & Ropaul, 2018; Buchheim & Kolasak, 2015). Projection bias has also been tested in the food sector and the under-appreciation of the effects of hunger is, perhaps, the most apparent evidence of this phenomenon (see de-Magistris & Gracia, 2016; Gilbert et al., 2002; Briz et al., 2015, for detailed analysis and discussion of testing for projection bias).

To illustrate, individuals who are subject to projection bias may choose to order too much food at the beginning of meals due to their current hunger state, to acknowledge afterwards of having fallen prey of unplanned impulsive over-purchase of food. For example, common advice is to eat a good meal before grocery shopping, to provide a defence against impulsive buying[1] assuming that one's current situation will remain constant in the future (Grable et al., 2006; Loewenstein & Schkade, 1999). But, how important is this potential projection bias for actual purchase decisions?

Most research on projection bias has been conducted under laboratory conditions, which are not necessarily reflective of live (field) experience (DellaVigna, 2009). Loewenstein (1999) suggests that all our choices (e.g., purchase decision-making) are context dependent, in which case testing for projection bias needs to consider how context affects behaviour (Harrison & List, 2004; Levitt & List, 2007). Harrison and List (2004) propose a taxonomy that differentiates natural from framed field experiments. In the latter, the field context is embodied in either the commodity, the task or information set that the subjects use (Harrison & List, 2004). A recent approach to embodying the field in the lab are virtual reality (VR) experiments, which can be considered as proper framed field experiments as they provide user or participant contexts under the experimenter's control. The critical element that differentiates VR from conventional laboratory experiments is the greater reflection of 'reality', which can provide original empirical evidence for various factors, such as the projection bias phenomenon.

Mol et al., (2019) claim that VR enables increased control in field experiments and improves laboratory studies' naturalistic context. Similarly, Gürerk and Kasulke (2018) illustrated that the injection of real-world frames and cues into virtual environments generates more natural responses than traditional laboratory experiments because the former provides a more emotionally engaging consumer experience (Burke, 2018). Likewise, Sinesio et al., (2019) pointed

[1]Impulsive buying is defined as the tendency of an individual to purchase goods and services without an advanced planning. When consumers take such buying decisions at the impulse of the moment, it is usually triggered by emotions and feelings (Weinberg and Gottwald 1982).

TABLE 1 Sociodemographic characteristics between experiments and treatments (%).

	Virtual supermarket (*n* = 161)					Real supermarket (*n* = 162)				
Demographics	**Hunger-today *n* = 40**	**Hunger-tomorrow *n* = 40**	**Satiated-today *n* = 40**	**Satiated-tomorrow *n* = 41**	**Pooled**	**Hunger-today *n* = 40**	**Hunger-tomorrow *n* = 41**	**Satiated-today *n* = 40**	**Satiated-tomorrow *n* = 41**	**Pooled**
Sex										
Male	47.5	47.5	47.5	48.8	47.8	47.5	46.3	47.5	46.3	46.9
Female	52.5	52.5	52.5	51.2	52.2	52.5	53.7	52.5	53.7	53.1
Age										
Mean age (SD)	44.7 (17.7)	43.7 (16.3)	44.7 (13.7)	44.4 (16.2)	44.4 (15.9)	48.1 (23.8)	46.6 (19.1)	47.7 (14.9)	46.5 (17.8)	46.9 (19.0)
18 to 34	30.0	30.0	27.5	26.8	28.6	27.5	26.8	20.0	26.8	25.3
35 to 54	37.5	37.5	37.5	39.0	37.8	40.0	36.6	45.0	36.6	39.5
55 and over	32.5	32.5	35.0	34.2	33.5	32.5	36.6	35.0	36.6	35.2
Education level										
Primary studies	15.0	15.0	15.0	14.6	14.9	12.5	12.2	15.0	21.6	15.0
Secondary studies	50.0	50.0	50.0	48.8	49.7	47.5	56.1	47.5	46.3	50.0
University studies	35.0	35.0	35.0	36.6	35.4	40.0	31.7	37.5	31.7	35.0

Note: SD refers to the standard deviation.

out that immersive technologies offer results with greater external validity and represent a valid alternative to consumer testing in real-life situations compared with conventional laboratory testing.

Our main objective in this paper is to examine whether projection bias is also present in virtual environments. We test the robustness of the VR argument by comparing a virtual supermarket (VS) experiment with a real supermarket (RS) version, where preferences reflect real market data. To the best of our knowledge, only two studies have adapted VR technology to the context of food preferences and have tested the robustness of their results.[2]

A hunger-manipulation-check task was performed under eight different treatments to achieve the main objective of this study. Participants were asked to participate in a shopping task using three food products (healthy and unhealthy) with nutritional claims (NCs) and health claims (HCs).[3] Moreover, to assess the robustness or our results from the virtual settings, we experimented with an RS where the products and their levels, financial incentives and the hunger-manipulation task for each experiment were identical.

We use financial incentives to increase the realism of virtual shopping settings following other studies by Demarque et al., (2015), Muller et al., (2019) and Panzone et al., (2018), who used real incentives to increase the realism of online shopping scenarios. The advantage of using economic incentives in experimental economics is that they mitigate hypothetical bias (Lusk & Shogren, 2007).

2 | METHODOLOGY

2.1 | Data gathering

The experiment was conducted in Zaragoza (Spain) between October 2017 and January 2018. The study followed the guidelines of the Declaration of Helsinki and the Ethics Committee of the host institution. The sample included 323 consumers recruited by a market research agency using a stratified sampling procedure based on gender, age and education level. Targeted respondents were primarily food buyers in the household who consumed the products and were older than 18 years. Participants were randomly assigned to one of the two experiments (VS or RS) and their treatments (Table 1).

2.2 | Identification of products and attributes

Three food product categories (sliced cheese, yoghurt and potato chips) with nutritional and health claims (NC and HC respectively) were chosen for several reasons. First, these products form part of a regular Spanish household's food-shopping basket. Second, the limited number of products prevented distraction of the participants and kept both experiments compatible. Our study also presented respondents with a budget constraint. Other reasons for including these food categories and their NCs and HCs are described in more detail in Appendix SA. The front-of-pack design for all food products was created using a graphic service enterprise. Brand

[2]Specifically, van Herpen et al., (2016) examined whether purchases from a VS with pictorial (2D) stimuli were similar to purchases from a physical store, while Siegrist et al., (2019) investigated whether people standing in front of a shelf in an RS made similar decisions to those in a VS.

[3]This study is part of a more extensive research exploring consumer behaviour regarding nutrition and health claims in Spain, in which multiple experiments were conducted.

names were excluded to avoid brand preference effects.[4] A total of 19 products were shown to participants: 4 types of potato chips, 4 types of sliced cheese and 11 types of yoghurt. A conventional full-fat version for each food category was selected as the baseline product (see Table A1 in Appendix SA).

2.3 | Experimental design

Our experiment included three conditions with a 2×2×2 between-subject factorial design. The first condition consisted of the experimental context in which consumers purchased food (*virtual* vs. *real*). The first experiment was conducted in an RS. It represented the benchmarking condition to externally validate the purchases made in the VS. Following the experimental design of de-Magistris and Gracia (2016), the second condition was the consumption period in which participants were asked to consider when they purchased food (*today* vs. *tomorrow*). The third condition was the appetite level, which varied based on participants' hunger level during the experiment (*hungry* vs. *satiated*) (see Appendix SB, for a detailed explanation of the implemented conditions and treatments).

The combination of these three conditions ([*RS* vs. *VS*], [*today* vs. *tomorrow*] and [*hungry* vs. *satiated*]) resulted in eight treatment groups. The first and second treatment groups, identified as *hungry-today*, denoted *hungry* participants who planned to consume the purchased food immediately after the VS and RS experiments. The third and fourth treatment groups, identified as *satiated-today*, included *satiated* individuals who planned to consume the purchased food immediately after the RS and VS experiments. The fifth and sixth treatment groups, identified as *hungry-tomorrow*, consisted of *hungry* participants who planned to consume the purchased food the next day when expected to be hungry for both experiments. Finally, the seventh and eighth groups, identified as *satiated-tomorrow*, denoted *satiated* respondents who planned to consume the purchased products the next day when expected to be hungry in the RS and VS experiments. Besides the available food alternatives, we also included a *no-buy* option for participants unwilling to buy food in the VS or the RS.

We followed the de-Magistris and Gracia (2016) study design to control participants' hunger levels, where hunger was manipulated during the experiments. To minimise the selection bias that could occur from people wanting to participate at different times (e.g., at or after lunch), we recruited hungry participants. More specifically, participants were asked not to consume any food for about 2 hours before arriving for the experiment. The sessions were set at typical lunch and dinner times when we expected participants to be hungry. For the satiated treatments, the participants ate unrelated food items from a buffet until satiated. A manipulation check question (*translated from Spanish*: Could you indicate the level of your hunger/satiation at this moment?) measured their subjective level of hunger and/or satiety (depending on the treatment) based on a Likert scale from 1 = not hungry/satiated at all to 5 = very hungry/satiated before and after consuming food from the buffet.

Participants were provided with €10 to pay for the purchased food to resemble an actual purchase. Following de-Magistris and Gracia (2016), participants belonging to the

[4]As in other consumer behaviour-related studies (Ballco and Gracia, 2020; Bazzani et al., 2017; Caputo et al., 2018; Marescotti et al., 2020; Sanjuán-López and Resano-Ezcaray, 2020), the brand name was not included as an attribute in our experiments as it did not form part of the main objective. Before the experiment, participants were informed that they would evaluate the products without their brand names. At the end of the experiment, participants were provided with the food packages that included the brand name under the distributors' (Eroski) brand name. According to Kantar Worldpanel, the purchase of food under the distributors' brand name has increased considerably within the last years (Jara, 2020; Segarra, 2019). Further, Spaniards spent €42 out of €100 on distributors' brand products (Press, 2019). Therefore, we believe that hiding the brand might have limited or no implications for either preferences or consumers' deception.

satiated-tomorrow and *hungry-tomorrow* treatments were informed that they would receive the purchased products the next day at the time of their choice (e.g., before lunch/dinner). However, they paid for them at the end of the experiment. Finally, the product category, product attributes, product prices, the buffet manipulation task and the incentive-compatible method were kept identical across both experiments (i.e., virtual and real). Appendix SC, provides more details on the hunger manipulation and control and the treatments used in both experiments (VS and RS).[5]

2.4 | Hypothesis and external validation

The present study tests projection bias[6] by considering the underestimation of the impact of current states on predicted future behaviour. In our study, one of the treatment groups, *hungry-today versus hungry-tomorrow*, was excluded from our analysis in both types of shops since it reflects hyperbolic discount behaviour, which we do not examine here.[7] Based on the evidence that hungry consumers over-purchase and overspend on food (Briz et al., 2015; de-Magistris & Gracia, 2016; Gilbert et al., 2002), the first null hypothesis relates to the presence of projection bias. This is to test whether the number of products bought and the amount of money spent by hungry participants, *hungry-tomorrow*, in the VS and RS is equal to the number of products bought and the amount of money spent by satiated subjects, *satiated-tomorrow*, in both types of shops. Then, the null H_1 is given as:

$$\mathbf{H_1}\text{: P/M (hungry-tomorrow)}_{(VS\ vs.\ RS)} = \text{P/M (satiated-tomorrow)}_{(VS\ vs.\ RS)}$$

where *P* indicates the number of products bought and *M* indicates the amount of money spent when people over-predict hunger for the planned food purchase to consume *tomorrow*. VS indicates that the treatments were implemented in the virtual supermarket, and RS in a real supermarket. If we reject the null H_1, we confirm the existence of projection bias, suggesting that participants in the *hungry-tomorrow* treatment would overbuy and overspend, implying that they expect their current hunger level to persist in the future. If projection bias exists, then the *satiated-tomorrow* participants who consumed food before purchase would behave differently (purchase and spend less on food) compared with the hungry participants, who purchase with an 'empty stomach' (Gilbert et al., 2002).

The second null hypothesis (H_2) refers to the food choices in the *today versus tomorrow* consumption treatments for the *satiated* participants in the VS and RS, type of shop respectively.

$$\mathbf{H_2}\text{: P/M (satiated-today)}_{(VS\ vs.\ RS)} = \text{P/M (satiated-tomorrow)}_{(VS\ vs.\ RS)}$$

In H_2, we expect *satiated-today* participants to buy and spend less on food in both types of shops (i.e., VS and RS), compared with *satiated-tomorrow* participants.

Finally, we test for the robustness of the purchase results between the two types of shops (i.e., VS and RS) and treatments in H_1 and H_2. Statistically insignificant differences between the two types of experiments would provide support for the external validity of using VR settings to study consumer behaviour.

[5] A detailed video of the virtual shopping experience can also be seen at the following link: https://youtu.be/CYljAzM9J6o

[6] Due to budget constraints, we were only able to test the hot-to-cold empathy gap in this study.

[7] The estimations of the hyperbolic discount treatments are available in Appendix SD.

2.5 | Implementation of the experiment

The VS experiment was conducted in a place near the RS in the centre of Zaragoza, which was easily accessible to all participants. Upon arrival, participants were put into groups of five per session and were informed that they would enter either a VS or an RS to purchase sliced cheese, potato chips and yoghurts. Each participant signed an informed consent form and received €10 to purchase their preferred option from each product category. Participants were also given a no-buy option in the VS, or simply decide not to buy in the RS. Participants paid the corresponding price at the end of the experiment.

All of the products existed in the local market. Participants were informed that they would evaluate the products without their brand name, which was not part of the experiment, though the brand was evident when the purchased food was delivered. The rest of the information (i.e., price, package size, NCs and HCs) corresponded to the real products sold in Spanish stores at the time of the experiment.

To check the participants' hunger level in the *hungry-tomorrow* and *hungry-today* treatments, participants responded to the hunger level manipulation check question (first questionnaire in Appendix SC) and then entered the supermarket (virtual or real). When they finished the purchase task, they were invited to consume unrelated food (excluding from the food products included in the research) from the buffet. After eating, they responded again to the manipulation check question (second questionnaire in Appendix SC). Conversely, participants in the *satiated-today* and *satiated-tomorrow* treatments first responded to the first questionnaire, so we could see that they were indeed hungry at the beginning of the experiment. Second, they were invited to eat unrelated food at the buffet to become satiated for the subsequent choice task. Third, they responded to the second questionnaire manipulation check after the buffet (see Appendix SC), which assured us that they were indeed satiated. Finally, they entered the supermarket (VS or RS) to purchase food.

Each product was placed on a particular shelf in the supermarket. All products were placed at the end of an aisle in a very similar position (near the cashier) to those placed in the VS (see footnote 7 above for a detailed view). Individuals entered the supermarket to select their preferred option(s) and paid the corresponding price(s) at a cash register. Finally, the research assistant delivered the real product(s), which included the brand name. Appendix SE, presents the description of the shopping task specified by individuals.

The virtual experiment was conducted in a laboratory setting, using a virtual experiment shopping simulation software.[8] The 19 products used in the VS were modelled as virtual objects and placed in the virtual world on three shelves corresponding to each food category as they appeared in the RS.

Before entering the virtual experiment, individuals were introduced to a sample virtual purchasing process using chocolate bars to familiarise themselves with the equipment and the VR environment. Appendix SF, presents the description of the virtual purchase process. At the end of the experiment, the research assistant delivered the real products, and the participant paid their corresponding prices.

[8]The TRINIT Association developed the 'Unity Game' engine version 5.4.3.f1 software used to create the virtual experiment. The virtual products' texts were generated by taking photos of the real packages from the RS and were attached to the 3-D virtual objects.

3 | RESULTS

3.1 | Descriptive analysis of sociodemographic characteristics

Table 1 shows the samples' characteristics by experiments (VS vs VR) and treatments (hungry vs. satiated and today vs. tomorrow).

The chi-square test showed that the VS and RS samples did not differ significantly in terms of gender (χ^2 [1] =0.027; $p = 0.87$) or education level (χ^2 [2] =0.017; $p = 0.99$). Likewise, the *t*-test results show that the VS and RS samples do not differ significantly in age groups ($t = 1.44$, $p = 0.15$). Furthermore, the chi-square and *t*-test results indicate that samples do not differ significantly between treatments and gender, education level and age within experiments.[9] Finally, there were no statistically significant differences between experimental contexts in terms of gender, age and education level.[10] These results imply that participants' randomisation successfully randomised their characteristics across treatments and experiments.

3.2 | Manipulation buffet check

A two-way ANOVA was conducted to examine the effect of treatments (*hungry* vs. *satiated* and *today* vs. *tomorrow*) and experiments (VS and RS) on the average self-reported hunger state when individuals participated in the purchase task. The results indicate no significant interaction between the effects of treatments and experiments on the average self-reported hunger state (F [3, 314] =0.58, $p = 0.56$). Likewise, there was no statistically significant difference between experiments (F [1, 314] =2.02, $p = 0.124$), but there were statistically significant differences between treatments (F [3, 314] =71.71, $p < 0.0001$). Since we do not have any interaction, but we do have a treatment effect, a Tukey test was conducted to identify the average differences of self-reported hunger states between treatments (see Appendix SG). The findings show a significant pairwise difference in the average self-reported hunger levels between the individuals who participated in the *hungry*-*today versus satiated*-*today* treatment and those that participated in the *hungry*-*tomorrow versus satiated*-*tomorrow* treatment in the VS and RS. These results indicate that for any experiment participants stated a significantly lower self-reported level of hunger after eating in the *satiated*-*today* and *satiated*-*tomorrow* treatments compared to those in the *hungry*-*today* and *hungry*-*tomorrow* treatments, and that they became satiated afterward.

3.3 | Hypothesis testing

To test the hypotheses related to the presence of projection bias, we considered the total (pooled) expenditure of the purchased food and the number of food products purchased by treatment in the respective category. Additionally, we also included the total (pooled) number of individuals who purchased food independently from each food category (see Appendix SH). First, we tested the hypotheses concerning the projection bias in the VS experiment. A two-sided *t*-test was used to test for differences in average expenditure between treatments. Conversely, a chi-square test was used to test for differences between the number of food products purchased by treatment and to test for differences between the pooled number of individuals who purchased

[9]The results of the tests are available upon request.

[10]The results of the *t*-test are available upon request.

TABLE 2 Average expenditure and number of purchased food between the *hungry-tomorrow* and *satiated-tomorrow* treatments in the VS experiment.

	Sliced cheese	**Potato chips**	**Yoghurt**	
Average expenditure by treatment (standard deviation)				Pooled average expenditure[a]
hungry-tomorrow	€1.37 (1.34)	€0.69 (0.69)	€1.06 (0.79)	€3.12 (2.04)
satiated-tomorrow	€0.85 (1.27)	€0.33 (0.58)	€0.74 (0.88)	€1.93 (2.03)
$M(\textit{hungry-tomorrow})_{VS}$ $=M(\textit{satiated-tomorrow})_{VS}$	$t = 1.79$*	$t = 2.49$**	$t = 1.72$*	$t = 2.64$**
Number of food products purchased by treatment (%)				Pooled purchased food[b]
hungry-tomorrow	21 (52.50)	21 (52.50)	27 (67.50)	35 (87.50)
satiated-tomorrow	13 (31.71)	11 (26.83)	18 (43.90)	26 (63.41)
$P(\textit{hungry-tomorrow})_{VS} = P(\textit{satiated-tomorrow})_{VS}$	$\chi^2(1)=3.59$*	$\chi^2(1)=5.58$**	$\chi^2(1)=4.57$**	$\chi^2(1)=6.32$**

***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

[a]Indicates participants' pooled average expenditure on food independently of whether they purchased a product from each food category or not (i.e., 3, 2, or just one food product).

[b]Indicates the number of individuals who purchased food independently, of whether they purchased a product from each food category or not.

food by treatments. Second, we conducted a robustness test to check whether the projection bias results were also confirmed in the RS experiment. Hence, in this section, the results from the VS, RS and their treatments will be shown separately. A two-way ANOVA model was conducted to examine the effect of treatments and experiments on all products' total expenditure and each product category. Finally, the Tukey test was used to identify the differences in average expenditures between experiments.

3.3.1 | Virtual reality

Table 2 presents the average expenditure and the number of purchased food across the *hungry-tomorrow* and *satiated-tomorrow* treatments in the VS experiment.

Based on the results, we reject the null H_1 for the VS experiment (H_1: [(P/M (*hungry-tomorrow*)$_{VS}$ =P/M (*satiated-tomorrow*)$_{VS}$] which implies some projection bias. Specifically, the total (pooled) average expenditure of hungry consumers who purchased food was €1.19 greater than the satiated subjects' purchases.

Similar behaviours were observed for all food categories, with differences of €0.52 for sliced cheese, €0.36 for potato chips, and €0.22 for yoghurt. The same trend was observed for the number of food products purchased by participants in each treatment. The number of products bought for future consumption by *hungry-tomorrow* subjects is significantly higher at 21 (52.5%) than the number of products bought by the *satiated-tomorrow* participants at 13 (31.71%). These results reject the null H_1 and demonstrate the presence of a projection bias in virtual environments.

Regarding null H_2, Table 3 presents the average expenditure and the number of purchased food across the *satiated-today* and *satiated-tomorrow* treatments. The results reject the null H_2 [(P/M (*satiated-today*)$_{VS}$ =P/M (*satiated-tomorrow*)$_{VS}$] in the VS experiment for total (pooled) average expenditure and total (pooled) purchased food. Specifically, the total (pooled) average expenditure of the *satiated-tomorrow* consumers differs by €0.72 than the *satiated-today* subjects. However, there were no significant differences in the average expenditure for the specific

TABLE 3 Average expenditure and number of purchased food between the *satiated-today* and *satiated-tomorrow* in the VS experiment.

	Sliced cheese	Potato chips	Yoghurt	
Average expenditure by treatment (standard deviation)				Pooled average expenditure[a]
satiated-today	€0.48 (1.06)	€0.24 (0.48)	€0.48 (0.82)	€1.21 (1.85)
satiated-tomorrow	€0.85 (1.27)	€0.33 (0.58)	€0.74 (0.88)	€1.93 (2.03)
$M(\textit{satiated-today})_{VS} = M(\textit{satiated-tomorrow})_{VS}$	$t = -1.42$	$t = -0.72$	$t = -1.41$	$t = -1.68$*
Number of food product purchased by treatment (%)				Pooled purchased food[b]
satiated-today	7 (17.50)	9 (22.50)	11 (27.50)	17 (42.50)
satiated-tomorrow	13 (31.71)	11 (26.83)	18 (43.90)	26 (63.41)
$P(\textit{satiated-today})_{VS} = P(\textit{satiated-tomorrow})_{VS}$	$\chi^2(1)=2.20$	$\chi^2(1)=0.20$	$\chi^2(1)=2.37$*	$\chi^2(1)=3.55$*

***, ** and *indicate statistical significance at the 1%, 5% and 10% levels.

[a]Indicates participants' pooled average expenditure on food independently of whether they purchased a product from each food category or not (i.e., 3, 2, or just one food product).

[b]Indicates the number of individuals who purchased food independently of whether they purchased a product from each food category or not.

TABLE 4 Average expenditure and number of purchased food between the *hungry-tomorrow* and *satiated-tomorrow* treatments in the RS experiment.

	Sliced cheese	Potato chips	Yoghurt	
Average expenditure by treatment (standard deviation)				Pooled average expenditure[a]
hungry-tomorrow	€1.87 (1.18)	€1.04 (0.54)	€1.55 (0.73)	€4.46 (2.03)
satiated-tomorrow	€1.24 (1.31)	€0.59 (0.67)	€0.97 (0.95)	€2.80 (2.35)
$M(\textit{hungry-tomorrow})_{RS} = M(\textit{satiated-tomorrow})_{RS}$	$t = 2.29$**	$t = 3.33$***	$t = 3.08$***	$t = 3.41$***
Number of food product purchased by treatment (%)				Pooled purchased food[b]
hungry-tomorrow	30 (73.17)	34 (82.93)	36 (87.80)	36 (87.80)
satiated-tomorrow	20 (48.78)	19 (46.34)	23 (56.10)	29 (70.73)
$P(\textit{hungry-tomorrow})_{RS} = P(\textit{satiated-tomorrow})_{RS}$	$\chi^2(1)=5.13$**	$\chi^2(1)=12.0$***	$\chi^2(1)=10.21$***	$\chi^2(1)=3.63$*

***, **and * indicate statistical significance at the 1%, 5% and 10% levels.

[a]Indicates participants' pooled average expenditure on food independently of whether they purchased a product from each food category or not (i.e., 3, 2, or just one food product).

[b]Indicates the number of individuals who purchased food independently of whether they purchased a product from each food category or not.

food categories between the two treatments.[11] Regarding the total (pooled) purchased food by treatment, results indicate that the number of satiated participants who purchased food products when they expected to consume them immediately (i.e., *satiated-today*) is significantly

[11]The average expenditures in the *satiated-today* treatment for potato chips, sliced cheese and yoghurt differ compared to the average expenditures in the *satiated-tomorrow* treatment (€0.09, €0.37 and €0.26 respectively). However, these differences are not statistically significant.

TABLE 5 Average expenditure and number of purchased food between the *satiated-today* and *satiated-tomorrow* in the RS experiment.

	Sliced cheese	Potato chips	Yoghurt	
Average expenditure by treatment (standard deviation)				Pooled average expenditure[a]
satiated-today	€0.53 (1.09)	€0.19 (0.42)	€0.61 (0.83)	€1.33 (1.76)
satiated-tomorrow	€1.24 (1.31)	€0.59 (0.67)	€0.97 (0.95)	€2.80 (2.35)
$M(satiated\text{-}today)_{RS} = M(satiated\text{-}tomorrow)_{RS}$	$t = -2.65$**	$t = -3.22$***	$t = -1.83$**	$t = -3.20$***
Number of food product purchased by treatment (%)				Pooled purchased food[b]
satiated-today	8 (20.00)	7 (17.50)	15 (37.50)	19 (47.50)
satiated-tomorrow	20 (48.78)	19 (46.34)	23 (56.10)	29 (70.73)
$P(satiated\text{-}today)_{RS} = P(satiated\text{-}tomorrow)_{RS}$	$\chi^2(1)=7.45$***	$\chi^2(1)=7.73$***	$\chi^2(1)=2.82$*	$\chi^2(1)=4.53$**

***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

[a]Indicates participants' pooled average expenditure on food independently of whether they purchased a product from each food category or not (i.e., 3, 2, or just one food product).

[b]Indicates the number of individuals who purchased food independently of whether they purchased a product from each food category or not.

lower [17 (42.50%)] than the number of satiated participants who purchased food to consume in the future hunger state (i.e., *satiated-tomorrow*) [26 (63.41%)].

3.3.2 | Real supermarket: robustness test

Table 4 presents the results from the average expenditure and the number of purchased food between the *hungry-tomorrow* and *satiated-tomorrow* treatments in the RS experiment.

The RS findings also indicate the existence of projection bias. In particular, the difference in total (pooled) expenditure is €1.66, while the differences in average expenditure for sliced cheese, potato chips and yoghurt are €0.63, €0.45 and €0.58, respectively. Regarding the number of purchased food by participants, the same trend is observed in the RS experiment as in the VS experiment. The number of products purchased by category from the participants in the *hungry-tomorrow* treatment is significantly higher than the number of food purchased by category from the individuals in the *satiated-tomorrow* treatment. More precisely, the differences in sliced cheese, potato chips and yoghurts between the *hungry-tomorrow* and *satiated-tomorrow* are 10 (24.4%), 15 (36.6%) and 13 (31.70%), respectively. Likewise, the total (pooled) number of hungry individuals who purchased food products in *hungry-tomorrow* treatment is significantly higher than the total (pooled) number of satiated individuals in the *satiated-tomorrow* treatment. Based on these results, we rejected the null H_1 for the RS experiment [$(P/M\ (hungry\text{-}tomorrow)_{RS} = P/M\ (satiated\text{-}tomorrow)_{RS}$] and confirmed the existence of projection bias in the RS experiment.

Regarding H_2, Table 5 presents the average expenditure and the number of food purchased by participants across the *satiated-today* and *satiated-tomorrow* treatments in the RS experiment.

The results indicate that we reject the null H_2 [$P/M\ (satiated\text{-}today)_{RS} = P/M\ (satiated\text{-}tomorrow)_{RS}$]. The total (pooled) average expenditure of the satiated consumers who planned a food purchase for immediate consumption differed by €1.47 compared to the satiated subjects who purchased food for an advanced/future consumption when expected to be hungry. The average expenditure on sliced cheese, potato chips and yoghurt for the participants in the

TABLE 6 Significant pairwise differences from Tukey test on average expenditure.

	Average expenditure	Contrast	Tukey *t*-test
Pooled	*satiated-tomorrow* vs. *hungry-tomorrow*	−1.422	− 4.70***
	satiated-today vs. *satiated-tomorrow*	− 1.099	− 3.62***
	RS vs. *VS*	0.895	4.16***
Sliced cheese	*satiated-tomorrow* vs. *hungry-tomorrow*	− 0.574	− 3.07**
	satiated-today vs. *satiated-tomorrow*	− 0.540	− 2.88**
	RS vs. *VS*	0.384	2.89
Potato chips	*satiated-tomorrow* vs. *hungry-tomorrow*	− 0.400	− 4.48***
	satiated-today vs. *satiated-tomorrow*	− 0.244	− 2.73**
	RS vs. *VS*	0.248	3.91
Yoghurt	*satiated-tomorrow* vs. *hungry-tomorrow*	− 0.314	− 2.42**
	satiated-today vs. *satiated-tomorrow*	− 0.447	− 3.46***
	RS vs. *VS*	0.262	2.85**

Note: ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

satiated-today treatment is significantly lower than the participants' average expenditures in the *satiated-today* treatment with a difference of €0.71, €0.40 and €0.36, respectively. Similarly, the chi-square results show that the number of purchased sliced cheese, potato chips and yoghurts is statistically significantly lower in the *satiated-today* treatment compared to those of the *satiated-tomorrow* treatment (with a difference of 12 [28.78%], 12 [28.84%] and 8 [18.60%], respectively). Likewise, the total (pooled) number of purchases from the satiated individuals in the *satiated-today* treatment is lower than the number of totals (pooled) purchased food of the participants in the *satiated-tomorrow* treatment.

Additionally, we estimated a three-way ANOVA model to examine the effect of treatments and experiments on the pooled average expenditures for all products and each product category. We found no significant three-way interactions for sliced cheese ($F[6, 291] = 1.56$, $p = 0.1572$), potato chips ($F[3, 291] = 0.28$, $p = 0.8417$) or yoghurt ($F[3, 291] = 0.76$, $p = 0.5185$). We considered two-way interactions since we found no statistically significant interaction between experiments, treatments and products. The results demonstrated no statistically significant interactions between the effects of treatments and experiments on the pooled average expenditure ($F[3, 315] = 1.64$, $p = 0.1810$), the average expenditure for sliced cheese ($F[3, 315] = 1.10$, $p = 0.5096$) or yoghurt ($F[3, 315] = 0.70$, $p = 0.5532$). However, we found a statistically significant two-way interaction for the total average amount of money spent on potato chips ($F[3, 315] = 2.95$, $p = 0.0329$). Conversely, the results of the two-way ANOVA model showed a statistically significant treatment effect ($p < 0.0001$) and experimental effect ($p < 0.0001$) for both the pooled average expenditure for all products and each product category. Hence, we conducted the Tukey post-hoc test to identify these differences. The results of the Tukey test are shown in Table 6.

These results confirm the hypotheses concerning the existence of projection bias in both the VS and RS settings, also suggesting that the VS setting, at least in this case, produces similar results to the RS setting. However, the total amount of money spent in the RS is higher than in the VS when considering both the total expenditure on products and expenditure for each product category. This suggests that the virtual environmental setting might reduce the projection bias of consumers.

4 | DISCUSSION AND CONCLUSIONS

There has been a growth in integrating behavioural evidence from psychological research with economic research. We have tried to improve the realism of a VR shopping environment and examine whether projection bias is present in both a virtual environment and in a real-world setting. In so doing, we also compare the validity of virtual reality experiments with their more real counterpart.

We find that hungry participants spent more money and purchased more food products to consume the next day (i.e., *hungry-tomorrow*) compared to their satiated counterparts (i.e., *satiated-tomorrow*). Moreover, satiated participants purchasing food for immediate consumption (i.e., *satiated-today*) spent and purchased less food than the satiated participants purchasing food for later consumption (i.e., *satiated-tomorrow*). These findings are consistent with the presence of projection bias, and with previous research findings, which include both laboratory and field experiments (Briz et al., 2015; de-Magistris & Gracia, 2016; Gilbert et al., 2002; Fisher & Ranger, 2014; Read & van Leeuwen, 1998). Specifically, we find evidence of projection bias with participants tending to over-project their hunger and under-project their satiation.

We also find that our participants in the virtual experiment behave similarly to those in a real setting. This is also in line with previous research, which identifies similarities in behaviour between virtual and real worlds (van Herpen et al., 2016; Siegrist et al., 2019; Spann et al., 2010; Waterlander et al., 2011, 2012, 2013). VR experiments are promising tools and, in our case at least, produce externally valid results. This suggests that virtual reality technology is capable of eliciting robust responses that mimic real purchase behaviours. Our results also suggest that the degree of projection bias (measured by the total amount of money spent for all (pooled) and each food) is reduced in the virtual setting.

The limitations of our study also suggest opportunities for future research. Our study examined consumer behaviour for only three product categories that differ in the extent to which they are habitually purchased. More product categories might be considered—to extend a VS to include all the product categories available in the RS—though would be challenging to design and implement. We have specifically focused here on a particular version of projection bias—the effects of current hunger/satiation condition on peoples' purchasing behaviour for future consumption. Our results reflect a more general problem—that behaviour is context dependent, and that isolating behavioural responses from their contexts is likely to produce biased results.

Virtual reality technologies have considerable promise in developing more sophisticated controlled experiments to incorporate different contextual conditions that better reflect the real world. Although the literature has shown several advantages in the use of VR in behavioural economics, it is limited so far, and we believe that the significance of these studies on consumer behaviour requires further assessment and extension. Researchers may use this novel approach to better understand and predict behaviour under hypothetical and future economic scenarios.

ACKNOWLEDGEMENTS

Tiziana de-Magistris acknowledges the financial support of the Spanish National Institute of Agricultural and Food Research and Technology (INIA) project RTA 2013-0092-00-00 entitled 'Consumer purchase behaviour of food with nutritional and/or health claims'. The authors thank the editor, David Harvey, and anonymous referees for their helpful comments and suggestions.

REFERENCES

Acland, D. & Levy, M.R. (2015) Naivete, Projection Bias, and Habit Formation in Gym Attendance. *Management Science*, 61(1), 146–160.

Ballco, P. & Gracia, A. (2020) An extended approach combining sensory and real choice experiments to examine new product attributes. *Food Quality and Preference*, 80, 103830.

Bazzani, C., Caputo, V., Nayga, R.M. & Canavari, M. (2017) Revisiting consumers' valuation for local versus organic food using a non-hypothetical choice experiment: Does personality matter? *Food Quality and Preference*, 62, 144–154.

Bienenstock, S. & Ropaul, M. (2018) On the benefits of being naive: the choice of contract duration with projection bias, *European Journal of Law and Economics*, 45(3), 469–496.

Briz, T., Drichoutis, A.C. & House, L. (2015) Examining projection bias in experimental auctions: the role of hunger and immediate gratification. *Agricultural and Food Economics*, 3(1), 22.

Buchheim, L. & Kolaska, T. (2015) *Projection Bias with Salient State-Dependent Utility*, Working paper, Available at https://lukasbuchheim.files.wordpress.com/2014/07/pbias.pdf. (accessed 26 November 2020).

Burke, R.R. (2018) Virtual Reality for Marketing Research. In: Moutinho, L. & Sokele, M. (Eds.) *Innovative Research Methodologies in Management: Volume II:* Futures. Biometrics and Neuroscience Research: Springer International Publishing, Cham, pp. 63–82.

Caputo, V., Van Loo, E.J., Scarpa, R., Nayga, R.M. & Verbeke, W. (2018) Comparing Serial, and Choice Task Stated and Inferred Attribute Non-Attendance Methods in Food Choice Experiments. *Journal of Agricultural Economics*, 69(1), 35–57.

DellaVigna, S. (2009) Psychology and Economics: Evidence from the Field. *Journal of Economic Literature*, 47(2), 315–372.

de-Magistris, T. & Gracia, A. (2016) Assessing projection bias in consumers' food preferences, *PLoS One*, 11(2), 146308

Demarque, C., Charalambides, L., Hilton, D.J. & Waroquier, L. (2015) Nudging sustainable consumption: The use of descriptive norms to promote a minority behavior in a realistic online shopping environment. *Journal of Environmental Psychology*, 43, 166–174.

Fisher, G. & Rangel, A. (2014) Symmetry in cold-to-hot and hot-to-cold valuation gaps, *Psychological Science*, 25(1), 120–127.

Gilbert, D.T., Gill, M.J. & Wilson, T.D. (2002) The Future Is Now: Temporal Correction in Affective Forecastin. *Organizational Behavior and Human Decision Processes*, 88(1), 430–444.

Grable, J., Lytton, R.H., Oneill, B., Joo, S.-H. & Klock, D. (2006) Risk tolerance, projection bias, vividness, and equity prices, *The Journal of Investing*, 15(2), 68–74.

Gürerk, Ö. & Kasulke, A. (2018) Does virtual reality increase charitable giving? An experimental study, *Organizations & Markets: Formal & Informal Structures eJournal*, Available at: https://doi.org/10.2139/ssrn.3072002.

Harrison, G.W. & List, J.A. (2004) Field Experiments, *Journal of Economic Literature*, 42(4), 1009–1055.

van Herpen, E., van den Broek, E., van Trijp, H.C.M. & Yu, T. (2016) Can a virtual supermarket bring realism into the lab? Comparing shopping behavior using virtual and pictorial store representations to behavior in a physical store. *Appetite*, 107, 196–207.

Levitt, S.D. & List, J.A. (2007) What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? *Journal of Economic Perspectives*, 21(2), 153–174.

Loewenstein, G., O'Donoghue, T. & Rabin, M. (2003) (2003), 'Projection Bias in Predicting Future Utility'. *The Quarterly Journal of Economics*, 118(4), 1209–1248.

Loewenstein, G. & Schkade, D. (1999) Wouldn't it be nice? Predicting future feelings', In: *Well-Being: The Foundations of Hedonic Psychology*. New York, NY: Russell Sage Foundation, pp. 85–105.

Lusk, J.L. & Shogren, J.F. (2007) *Experimental auctions: Methods and applications in economic and marketing research*. Cambridge: Cambridge University Press.

Marescotti, M.E., Caputo, V., Demartini, E. & Gaviglio, A. (2020) Consumer preferences for wild game cured meat label: do attitudes towards animal welfare matter? *International Food and Agribusiness Management Review*, 23(4), 599–618.

Mol, J.M. (2019) Goggles in the lab: Economic experiments in immersive virtual environments. *Journal of Behavioral and Experimental Economics*, 79, 155–164.

Muller, L., Lacroix, A. & Ruffieux, B. (2019) Environmental Labelling and Consumption Changes: A Food Choice Experiment, *Environmental and Resource Economics, Springer Netherlands*, Available at: https://agris.fao.org/agris-search/search.do?recordID=US201900321401 (accessed 25 June 2020).

Panzone, L.A., Ulph, A., Zizzo, D.J., Hilton, D. & Clear, A. (2018) The impact of environmental recall and carbon taxation on the carbon footprint of supermarket shopping. *Journal of Environmental Economics and Management*, 102–137. https://doi.org/10.1016/j.jeem.2018.06.002

Press, E. (2019) Aumenta el número de españoles que consume productos de marca blanca, *Vozpópuli*, Available at: https://www.vozpopuli.com/economia-y-finanzas/productos-marca-blanca-Mercadona-Lidl-Espana_0_1249375589.html (accessed 26 November 2020).

Read, D. & van Leeuwen, B. (1998) Predicting Hunger: The Effects of Appetite and Delay on Choice. *Organizational Behavior and Human Decision Processes*, 76(2), 189–205.

Sanjuán-López, A.I. & Resano-Ezcaray, H. (2020) Labels for a Local Food Speciality Product: The Case of Saffron. *Journal of Agricultural Economics*, 71(3), 778–797.

Segarra, P. (2019) La marca blanca, que alcanza el 40% de las ventas, se reinventa y apuesta también por productos prémium, www.20minutos.es - Últimas Noticias, (2019), 31 May, available at: https://www.20minutos.es/noticia/3654075/0/marca-blanca-mercadona-premium-distribucion-supermercados/ (accessed 26 November 2020).

Siegrist, M., Ung, C.-Y., Zank, M., Marinello, M., Kunz, A., Hartmann, C. & et al (2019) Consumers' food selection behaviors in three-dimensional (3D) virtual reality, *Special Issue on* "*Virtual Reality and Food: Applications in Sensory and Consumer Science"*, 117, 50–59.

Sinesio, F., Moneta, E., Porcherot, C., Abbà, S., Dreyfuss, L., Guillamet, K. et al (2019) Do immersive techniques help to capture consumer reality? *Food Quality and Preference*, 77, 123–134.

Spann, M., Hinz, O., Hann, I.-H. & Skiera, B. (2010) *Decision making in virtual worlds: An experimental test of altruism, fairness and presence*, presented at the Proccedings of the European Conference on Information Systems (ECIS), p. 36.

Waterlander, W.E., Scarpa, M., Lentz, D. & Steenhuis, I.H. (2011) The virtual supermarket: An innovative research tool to study consumer food purchasing behaviour. *BMC Public Health*, 11(1), 589.

Waterlander, W.E., Steenhuis, I.H.M., de Boer, M.R., Schuit, A.J. & Seidell, J.C. (2012) Introducing taxes, subsidies or both: The effects of various food pricing strategies in a web-based supermarket randomized trial. *Preventive Medicine*, 54(5), 323–330.

Waterlander, W.E., Steenhuis, I.H., de Boer, M.R., Schuit, A.J. & Seidell, J.C. (2013) Effects of different discount levels on healthy products coupled with a healthy choice label, special offer label or both: results from a web-based supermarket experiment. *International Journal of Behavioral Nutrition and Physical Activity*, 10(1), 59.

Weinberg, P. & Gottwald, W. (1982) Impulsive consumer buying as a result of emotions, *Journal of Business Research*, 10(1), 43–57.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: de-Magistris T, López-Galán B, Ballco P. Do virtual reality experiments replicate projection bias phenomena? Examining the external validity of a virtual supermarket. *J Agric Econ*. 2021;00:1–15. https://doi.org/10.1111/1477-9552.12443