RESPONSE TO REVIEWERS

We would like to thank the editor and the reviewers for taking the time to consider our manuscript and for their constructive comments and suggestions. In light of this feedback, we have made several broad changes to the manuscript, in addition to addressing the point-by-point questions outlined below. First, we have revised the structure of the introduction to bring the focus of the manuscript squarely within the domain of methods for presenting real-world objects under controlled viewing conditions. We emphasize later in the introduction that we use a decision-making task as an example protocol with which to apply our 'real-world' approaches and apparatus. Second, in line with the suggestions of both the Editor and the reviewers, we have substantially revised the structure of the discussion to consider explicitly the critical steps in the protocol, the significance of our approach with respect to existing methods, and possible future applications of the technique. Importantly, we consider carefully the various limitations and challenges of working with real objects in experimental contexts and whether or not the time and resource costs of working with real objects are justified, or if similar results can be obtained using more convenient image displays.

Taken together, the changes we have made to the manuscript have strengthened the paper considerably. Point-by-point responses to each of the reviewer's queries are provided below.

Editorial comments:

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues. The JoVE editor will not copy-edit your manuscript and any errors in the submitted revision may be present in the published version.

Resolved

2. Please revise the title to be more concise and avoid punctuations.

We have revised our title to be more concise, as follows:

"Methods for presenting real-world objects under controlled laboratory conditions"

3. Please define all abbreviations during the first-time use.

Resolved

4. Please remove all commercial language from your manuscript and use generic terms instead. All commercial products should be included in the Table of Materials and Reagents and reference the table wherever applicable.
For example: Canon Rebel T2i DSLR, Adobe Photoshop, PLATO liquid crystal occlusion goggles, MATLAB,

Resolved

5. Please ensure that all text in the protocol section is written in the imperative tense as if telling someone how to do the technique (e.g., “Do this,” “Ensure that,” etc.). The actions should be described in the imperative tense in complete sentences wherever possible. Avoid usage of phrases such as “could be,” “should be,” and “would be” throughout the Protocol. Any text that cannot be written in the imperative tense may be added as a “Note.” However, notes should be concise and used sparingly.

Resolved

6. The Protocol should contain only action items that direct the reader to do something.

Resolved

7. The Protocol should be made up almost entirely of discrete steps without large paragraphs of text between sections.

Resolved

8. Please add more details to your protocol steps. Please ensure you answer the “how” question, i.e., how is the step performed?

Resolved see 9- 15.

9. 1.1.2.4: How is this done?

We have revised this section to detail how adjustments to images are made, as follows:

“1.1.2.4 If necessary, adjust the 2-D images for color, luminance, and visual size using image processing software (see **Table of Materials**). Click on the Hue/Saturation and Brightness/Contrast tabs and move the sliders until the image looks as similar as possible to its real-world counterpart when mounted on the turntable. Fine tune the size of the object in the image by positioning the real object next to the computer monitor and increase/decrease the pixel size until the stimuli are matched exactly for size. Make sure that the image processing software page view (zoom) is set at 100% when editing. Ensure that the monitor used for editing the images is the same (or the same size) monitor that will be used as the participant monitor during the study. Hold constant the resolution, aspect ratio and pixels per inch of the images. Additionally, confirm that the monitor is large enough to display the largest stimulus at its full size.”

10. 2.6.1.1: How is this done? What codes are used?

We have revised this section to include where the code can be found to trigger the liquid crystal occlusion glasses, as follows:

"2.6.1.1 Make a computer command (e.g., a button press) to trigger opening and closing of the glasses allowing for the real food to be visible on the turntable for 3 sec. Once the glasses close, position the participant monitor back in front of the aperture and press a key to open the glasses in order for the participant to make a response (e.g., a bid). Have the glasses automatically close once the participant enters his/her response (see 'run study' script, available at https://snowlab.blogs.unr.edu/)."

11. 2.6.1.1.1.: How?

This section has been combined with the previous section 2.6.1.1.

12. 3.1: Please include the script as a supplemental file.

Because of the nature of the files that are needed to run the study (i.e. .mat and images files) and because all files must be included in a single folder, we believe it will be better to provide a link to a compressed folder on our lab website (https://snowlab.blogs.unr.edu/) than scripts as a supplemental file.

13. 3.2: How? Codes used, scripts used to do the same?

A link to the necessary files to perform this task has been included.

14. 8: Include the script/ codes used for this as a supplemental file.

A link to the necessary files to perform this task has been included.

15. 10: How is this done?

We have revised this section to detail how the linear mixed effects model analysis is done, as follows:

"10.1 Use statistical analysis software (see **Table of Materials**) to perform a linear mixed effects modeling analysis. Use a linear mixed effects model to account for the nested responses within participants (i.e., each participant places bids on both real items and items seen as images). Create a dataset with the following variables: Participants, Item, Display Format, Preference, Caloric Density, Estimated Calories, and Bid. Create a model by clicking Analyze, then Mixed Models, then Linear.

10.1.1 Transfer the variable Participants into the Subjects: box then press continue. Transfer Bid into the Dependent Variable: box. Next, transfer Subject and Display Format into the Factor(s): box. Then, transfer Preference into the Covariate(s): box.

10.1.2 Click Fixed… then select and add all of the variables except Participants into the Model: box then click Continue. Click Random… then select and add Participants into the Combinations: box in order to account for variability in responses within and across observers. Click continue.

10.1.3 Click Statistics… then check the Descriptive statistics, Parameter estimates, and Tests for covariance parameters boxes. Click Continue. Click EM Means… then select and transfer all of the Factors and Factor Interactions into the Display Means for: box. Click Continue. Lastly, press OK."

16. Lines 182-191: We cannot have paragraph of texts in the protocol section, please consider making steps in imperative tense or move to the discussion section

This text has been removed.

17. Software steps must be more explicitly explained ('click', 'select', etc.). Please include all the codes used for this experiment as a supplemental file.

Resolved

18. Please do not use personal pronouns throughout the protocol section.

Resolved

19. There is a 10-page limit for the Protocol, but there is a 2.75-page limit for filmable content. Please highlight 2.75 pages or less of the Protocol (including headings and spacing) that identifies the essential steps of the protocol for the video, i.e., the steps that should be visualized to tell the most cohesive story of the Protocol. Remember that non-highlighted Protocol steps will remain in the manuscript, and therefore will still be available to the reader.

Resolved

20. Please include all the Figure Legends together at the end of the Representative Results in the manuscript text.

Resolved

21. Please obtain explicit copyright permission to reuse any figures from a previous publication. Explicit permission can be expressed in the form of a letter from the editor or a link to the editorial policy that allows re-prints. Please upload this information as a .doc or .docx file to your Editorial Manager account. The Figure must be cited appropriately in the Figure Legend, i.e. "This figure has been modified from [citation]."

The appropriate documents have been uploaded to the portal. Elsevier has requested that acknowledgement to the source be made in the following way:

"Reprinted from Publication title, Vol/edition number, Author(s), Title of article/title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]."

22. As we are a methods journal, please revise the Discussion to explicitly cover the following in detail in 3-6 paragraphs with citations:
a) Critical steps within the protocol
b) Any modifications and troubleshooting of the technique
c) Any limitations of the technique
d) The significance with respect to existing methods
e) Any future applications of the technique

The discussion section of the revised draft has been modified substantially to reflect the above criteria. Please see revised manuscript.

23. All figures should be uploaded separately to your Editorial Manager account.

Resolved

24. Please revise the table of the essential supplies, reagents, and equipment. The table should include the name, company, and catalog number of all relevant materials in separate columns. Once done please sort the table alphabetically.

Resolved

Reviewer #1:
Manuscript Summary:
This is a very well-written manuscript describing a procedure for bringing real-world objects into experimental investigations of decision making. The rationale for instituting the procedure is well-formulated.

Major Concerns:
In a few places the authors remark that there has been little or no experiments examining valuation of real food objects versus images of foods. However, at least one of the studies that they cite in another context seems to have done exactly that (Bushong et al., 2010). They should be cited as an exception to statements like the one made at the top of page 4.

We thank Reviewer 1 for highlighting this point. We have modified the introduction to address this concern, as follows:

Introduction, p.3: "However, few studies have examined in adults whether the format in which a snack food is displayed influences decision-making or valuation[12,32,33] and only one study to date, from our laboratory, has tested this question when stimulus parameters and environmental factors are matched across formats[7]."

On page 4, starting on line 114 the authors note that their procedure provides advantages over previous Beck-DeGroot-Marschk bidding task set-ups allowing greater control over stimulus parameters, presentation and timing. It would be beneficial to indicate some of the specific features of the classic bidding task formats and how they are limited. It would also be helpful to indicate which features of the current method overlap with the existing ones and what specific features are different?

We agree with Reviewer 1 that it would be beneficial to describe what advantages our procedures provide over previous studies that have used the Becker-DeGroot-Marschk bidding task. We have substantially modified the introduction to highlight the key similarities and differences between our approach and that of Bushong et al., (2010):

Introduction, pp.3-4: "Here, we describe innovative techniques and apparatus for investigating whether decision-making in healthy human observers is influenced by the format in which the stimuli are displayed. Our study[7] was motivated by a previous experiment conducted by Bushong and colleagues[12] in which college-aged students were asked to place monetary bids on a range of everyday snack foods using a Becker-DeGroot-Marschak (BDM) bidding task[34]. Using a between-subjects design, Bushong and colleagues[12] presented the snack foods in one of three formats: text descriptors (i.e., 'Snickers bar"), 2-D colored images, or real foods. Average bids for the snacks (in dollars) were contrasted across the three participant groups. Surprisingly, students who viewed real foods were willing to pay 61% more for the items than those who viewed the same stimuli as images or text descriptors –a phenomenon the authors termed the 'real-exposure effect'[12]. Critically, however, participants in the text and image conditions completed the bidding task in a group setting and entered their responses via individual computer terminals; conversely, those assigned to the real food condition performed the task one-on-one with the experimenter. The appearance of the stimuli in the real and image conditions was also different. In the real food condition, the foods were presented to the observer on a silver tray, whereas in the image condition the stimuli were presented as scaled cropped images on a black background. Thus it is possible that participant differences, environmental conditions, or stimulus-related differences, could have led to inflated bids for the real foods. Following from Bushong, et al.[12] we examined whether the real foods are valued more than 2-D images of food, but critically, we used a within-subjects design in which environmental and stimulus-related factors were carefully controlled. We developed a custom-designed turntable in which the stimuli in each display format could be interleaved randomly from trial to trial. Stimulus presentation and timing were identical across the real object and image trials, thus reducing the likelihood that participants could use different strategies to perform the task in the different display conditions. Finally, we controlled carefully the appearance of the stimuli in the real object and image conditions so that the real foods and images were matched for closely for apparent size, distance, viewpoint, and background."

The authors note that one of the advantages of the set-up is that it allows for random, interleaved presentations of stimuli. They relate this to enhancing the ability to do within-subjects studies which increases power. I did not find this argument as compelling for 2 reasons. First, one of the drawbacks of a within-subjects design is that the purpose of the experiment would be pretty transparent and therefore the likelihood of experimenter demand effects would be high. A between-subjects design eliminates this considerable concern. Second, it is not clear that the suggested procedure is the only one that allows such a design or allows for randomized presentations. I suggest touching on these points in the paper for clarity.

We agree with Reviewer 1. These points have been clarified in the introduction section of the revised manuscript, as indicated below:

Introduction, p.4: "There are likely to be other procedures or mechanisms that could allow for randomizing stimulus formats across trials but our method allows for a large number of objects (and images) to be presented in relatively rapid interleaved succession. From a statistical standpoint, this design maximizes power to detect significant effects more so than is possible using between-subjects designs. Similarly, the effects cannot be attributed to a-priori differences in willingness-to-pay (WTP) between observers. It is, of course, the case that in within-subjects designs open the possibility for demand characteristics. However, in our study participants understood that they could 'win' a food item at the end of the experiment regardless of the display format in which it appeared in the bidding task. Participants were also informed that arbitrarily reducing bids (i.e., for the images) would reduce their chances of winning, and that the best strategy for wining a desired item is to bid one's true value[34-36]."

Page 11: A question related to the food auction. "if the computer makes a bid . . . less than or equal to the participant's bid, then the participant wins the item for consumption and pays the price of the computer's bid." Why does the consumer not pay the price that they bid as opposed to what the computer bid? If the participant knows they could end up paying less than their bid, then could this introduce different strategies that might complicate interpreting the results? This also lacks ecological validity. Perhaps the rationale for this could be explained somewhere.

We agree with Reviewer 1 that it seems counterintuitive that the participant should pay for the food item the price that was generated by the computer rather than their own bid value. However, first, it is important to point out that the focus of the current paper is on methods for presenting real objects under controlled viewing conditions, more so than the particulars of the Becker-DeGroot-Marschak bidding task. Second, the BDM task is a relatively well-established method that has been applied in a large number of previous studies in the decision-making literature. Finally, given that one of our major goals was to replicate the 'real exposure effect' described originally by Bushong et al., (2010) – but under more controlled conditions, we kept the bidding task in line with the original study. If Reviewer 1 would like to follow up on this issue further, the following papers provide an in-depth consideration of the rationale for the BDM procedure:

34 Becker, G. M., DeGroot, M. H. & Marschak, J. Stochastic models of choice behavior. *Behav Sci.* **8** (1), 41-55 (1963).

35 Plassmann, H., O'Doherty, J. & Rangel, A. Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions. *The Journal of Neuroscience.* **27** (37), 9984-9988, doi:10.1523/jneurosci.2131-07.2007, (2007).

36 Becker, G. M., DeGroot, M. H. & Marschak, J. Measuring utility by a single-response sequential method. *Behav Sci.* **9** (3), 226-232 (1964).

38 Johnson, E. J., Haubl, G. & Keinan, A. Aspects of endowment: A query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* **33** (3), 461-474, doi:10.1037/0278-7393.33.3.461, (2007).

In anticipation that other readers may have similar questions about the BDM task, we have added a sentence to p.11 of the Methods section to direct readers to relevant background material about the BDM bidding task, as follows:

"A number of previous studies have provided an in-depth consideration of the rationale for the BDM bidding task[34,36,38].

Page 13: It is noted that although it is tempting to use photographs in experiments rather than repeating the set-up for each participant, it is "imperative to ensure that empirical findings generalize to real-world contexts . . ." One of the issues that would be important to touch on in the discussion is how the differences between how 2D and 3D stimuli were valuated might lead to erroneous conclusions in experiments (and what kind of errors). For example, is the reduction in willingness to pay for 2D images versus real objects is just characterized by a linear transformation? If this is the case, the issue may not be a particularly troubling one if there is still a considerable range. It just means that the translation to real-life is simply a linear and predictable one. This is an important issue to address in the paper as it speaks to evaluating the criticality of the manipulation (in the face of its time and resource costs). Note, however, that regardless of the relationship observed here, it does not dismiss the utility of the application in other settings, but it is something that should be considered when deciding on whether the extra resource cost of this procedure is necessary for a given application.

We thank Reviewer 1 for highlighting this point. A similar concern was also raised by Reviewer 2. Given that both reviewers asked about this issue, we have revised the discussion quite substantially to consider this point in detail, as follows:

Discussion, pp.15-16 (last paragraph): "The above considerations raise the critical question of whether or not the time and resource costs of working with real objects are justified, or whether similar results can be obtained using more convenient image displays. The results from our decision-making paradigm[7] indicate that real food displays elicit a constant increase in valuation (i.e., a linear effect) that does not interact with other factors such as preference or caloric density. These results from decision-making dovetail with findings from other domains

of human cognition. For example, real-world objects are more easily recognized[10,40,41], enhance memory[42], and capture attention[43,44] more so than images do. Compared to 2-D images, fMRI repetition suppression effects are reduced for real objects (Snow et al., 2011). Similarly, fine-grained examination of the temporal dynamics of brain responses to real objects as measured by high-density EEG reveals that real objects (vs. images) elicit stronger and more prolonged desynchronization of the mu rhythm –a signature of activation in visuo-motor networks involved in automatic planning of motor actions[9]. The amplification in mu desynchronization for real objects is independent of early signal differences related to stereopsis[9]. Taken together, these findings suggest that the pattern of results that could be obtained using image displays may be broadly consistent, but just less compelling, than what might otherwise have been observed had real-world objects been used. In other words, if findings from studies of image vision transfer predictably to real object vision, then the translational value of basic research studies of image vision is preserved. Although there is currently insufficient data to make firm conclusions on this issue, recent evidence for dissociations in the effects of real objects across motor areas in the left versus right hemispheres[9] and across egocentric distances[6] raise concerns about this assumption. For example, the effect of real objects on attentional capture fall to the levels observed for 2-D and 3-D images when the objects are positioned outside of reach of the observer, or when they are within reach but behind a transparent barrier [6], suggesting that the potential for manual interaction with a real object (but not an image) determines how it is processed. Future studies could use the protocols described here to investigate whether similar underlying causal mechanisms modulate 'real-food exposure effects' on willingness-to-pay. For example, a distance or barrier manipulation[6] could be employed to determine whether real snack foods that are reachable or graspable are processed differently to those that are not (and to determine whether the same manipulation has any effect on processing of food images). Future studies using ecologically-valid real-object stimuli are required to make definitive conclusions on this issue. Importantly, it may not be the case that similar mechanisms are at play in different cognitive domains, or in different tasks. Nevertheless, our approach to working with real-world objects promises to provide important new insights into the underlying processes and mechanisms that drive naturalistic vision.
."

Minor Concerns:
I was not clear on the positioning of the subject's monitor relative to the aperture. I gather that this is on some kind of slide and moves back and forth depending on the stimulus condition. A few more details about the slide itself early on would be helpful.

We have revised the location of the sliding platform image and description in **Figure 3** from **D** to **C** and placed the entire figure after the description of the sliding platform in the text for clarity.

Reviewer #2:
Manuscript Summary:
The authors present a method and apparatus that can be used to assess responses to real-world stimuli, and importantly, directly compare these responses with 2-D displayed versions. The authors make a compelling case for the utility of this approach for this particular topic, and

the apparatus is nicely designed.

Major Concerns:
The manuscript is clear and for the most part well-written (see minor comments below). One general comment is that the multiple short sections (some consisting of only one sentence) throughout the paper are jarring. I think some of these could be consolidated to help the flow of the paper. To give just one example, Section 2 could be divided into two parts A setups for a test session, and B procedure during test session, with fewer subsections.

We agree that in many cases the multiple short sections are jarring to the reader, but the Editor has indicated that this is the format that is required for JOVE. However, we have done our best to combine some of the sections to reduce the number of short sentences within the limits of the journal requirements.

I realize that the results shown here are from another, already published, paper but this study provides an excellent opportunity to directly assess the impact of stereopsis by including real-world tests with monocular viewing (easy to control with the test goggles). This would form a nice comparison with the monitor-based trials in which motion parallax and stereopsis provide conflicting depth information.

This point has been outlined in the discussion section of the revised manuscript:

Discussion, p.14: "The methods presented here open up new avenues to examine the underlying mechanisms for the observed effects. For example, future studies could assess directly the impact of stereopsis by presenting real-world stimuli under monocular viewing conditions (which could, for example, be tested easily using monocular vs. binocular states of the computer-controlled glasses described here). This would form a nice comparison with the image-based trials in which both motion parallax and stereopsis provide conflicting depth information."

I assume there is some reason for using the awkward term 'likeingness' instead of preference or appreciation. Similarly, the rating scale could be called a preference or appreciation scale.

We have changed the term 'likingness' to 'preference' throughout the manuscript, and in the figures/captions.

Throughout the paper it would be useful for the reader if the authors used 'estimated' caloric density rather than simply caloric density, since the participants' estimates were not likely veridical.

To clarify, we included both actual caloric density, and participants' estimates of the caloric density of the foods, in our linear mixed effects modeling analysis. We have clarified this point in the manuscript, as follows:

Protocol Section 10.1: "Create a dataset with the following variables: Participants, Item, Display Format, Preference, Caloric Density, Estimated Calories, and Bid."

The effect of estimated caloric density is weak; I assume that this is discussed in the original paper.

We did discuss this result in the original manuscript, and in fact in the original paper we ran a second experiment to address this question in more detail. For the purpose of the current manuscript, we have clarified this point, as follows:

Representative Results, p.11: "Although we observed an effect of estimated calories on Bids, the effect was relatively weak. This result may be explained by the fact that participants performed the estimation task in response to text prompts after the main experiment, rather than while looking at the foods at the time of stimulus presentation. Furthermore, estimating the number of calories in a given food item is not necessarily an intuitive task; many observers are unaware (or do not pay attention to) the caloric density of the foods they consume."

In all of the data presented here the difference between the real-world and displayed objects conditions is a constant shift, with no interaction effects. That is the pattern of the results is consistent. I assume that in the original paper the authors discuss that depending on the aims of the study the results obtained using displays may be valid, just less compelling.

We thank Reviewer 2 for highlighting this point. A similar concern was also raised by Reviewer 1. Given that both reviewers asked about this issue, we have revised the discussion quite substantially to consider this point in detail, as follows:

Discussion, pp.15-16 (last paragraph): "The above considerations raise the critical question of whether or not the time and resource costs of working with real objects are justified, or whether similar results can be obtained using more convenient image displays. The results from our decision-making paradigm[7] indicate that real food displays elicit a constant increase in valuation (i.e., a linear effect) that does not interact with other factors such as preference or caloric density. These results from decision-making dovetail with findings from other domains of human cognition. For example, real-world objects are more easily recognized[10,40,41], enhance memory[42], and capture attention[43,44] more so than images do. Compared to 2-D images, fMRI repetition suppression effects are reduced for real objects (Snow et al., 2011). Similarly, fine-grained examination of the temporal dynamics of brain responses to real objects as measured by high-density EEG reveals that real objects (vs. images) elicit stronger and more prolonged desynchronization of the mu rhythm –a signature of activation in visuo-motor networks involved in automatic planning of motor actions[9]. The amplification in mu desynchronization for real objects is independent of early signal differences related to stereopsis[9]. Taken together, these findings suggest that the pattern of results that could be obtained using image displays may be broadly consistent, but just less compelling, than what might otherwise have been observed had real-world objects been used. In other words, if findings from studies of image vision transfer predictably to real object vision, then the translational value of basic research studies of image vision is preserved. Although there is currently insufficient data to make firm conclusions on this issue, recent evidence for dissociations in the effects of real objects across motor areas in the left versus right hemispheres[9] and across egocentric distances[6] raise concerns about this assumption. For example, the effect of real objects on attentional capture fall to the levels observed for 2-D and 3-D images when the objects are positioned outside of reach of the observer, or when they are within reach but behind a transparent barrier [6], suggesting that the potential for manual interaction with a real object (but not an image) determines how it is processed. Future studies could use the protocols described here to investigate whether similar underlying causal mechanisms modulate 'real-food exposure effects' on willingness-to-pay. For example, a distance or barrier manipulation[6] could be employed to determine whether real snack foods that are reachable or graspable are processed differently to those that are not (and to determine whether the same manipulation has any effect on processing of food images). Future studies using ecologically-valid real-object stimuli are required to make definitive conclusions on this issue. Importantly, it may not be the case that similar mechanisms are at play in different cognitive domains, or in different tasks. Nevertheless, our approach to working with real-world objects promises to provide important new insights into the underlying processes and mechanisms that drive naturalistic vision."

Section 5.1 is redundant with the text that precedes it.

This section has been removed.

The organization of the sections describing the apparatus is a bit confusing. I suggest describing the sliding platform prior (1.2.5.1 and to Figure 3 where this piece of the set-up is referred to.

We have revised the location of the sliding platform image and description in **Figure 3** from **D** to **C** and placed the entire figure after the description of the sliding platform in the text for clarity.

Minor Concerns:
40 tense shift (this is one example, there are many other instances throughout)

Resolved

97 text should be modified to refer to the position and timing of object presentation

Resolved

163 The statement that the 50cm distance is typical in visual perception experiments is too broad/sweeping. I would just remove that clause.

The statement has been removed

263 The reference to the updated occluding glasses is unclear - a bit more detail would be useful here.

This reference has been removed.

342 would view -> viewed

Resolved

435 the participant loses the item -> does not win the item

Resolved