To the Editor of JoVE:

Thank you for inviting us to submit a revised draft of our manuscript “Novel Object Recognition and Object Location Behavioral Testing in Mice on a Budget” to the Journal of Visualized Experiments. We are grateful for the reviewers’ insightful comments on the manuscript and have edited the manuscript to address their concerns.

Overall, we agree with the general comment from the reviewers regarding our previously restrictive categorization of NORT as completely hippocampus-independent. To address this, we have added more elaboration on the complexities of the brain region-dependency of the NORT, especially in the introduction and discussion sections. We have also accounted for the possible role of the order of the OLT and NORT in encouraging use of spatial strategies in the NORT by adding a note in the protocol itself and a recommendation in the discussion. To address the concern over emphasizing the importance of object selection and validation, we have made several changes to the protocol and discussion. Notably, we have moved the validation step earlier in the protocol and have expanded our discussion section to include the references provided by Reviewer 3. We have also modified Figure 1C and streamlined the discussion of that figure which hopefully provide a clearer explanation of appropriate objects.

In the rest of this letter, we address the reviewers’ specific concerns.

**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.*

We have made the necessary corrections to the manuscript, figures, and table.

*2. 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].”*

We have corrected the citation in the figure legend and uploaded a link to the editorial policy to our Editorial Manager account.

*3. Figure 3: Please define error bars in the figure legend.*

We have added the appropriate definition at the end of the figure legend.

*4. Table 1: Please fix the typos in the table.*

The typos have been corrected.

*5. Please provide an email address for each author.*

We have provided email addresses for every author.

*6. Please use SI abbreviations for all units: L, mL, µL, h, min, s, etc.*

All units in the manuscript, table, and figures are now in SI units.

*7. Please remove the commercial pricing from the manuscript. Though the emphasis of the paper is on the economical method of creating the apparatus, the wholesale listing of prices in the manuscript and in the Tables is not appropriate for publication here. Is there any other way to discuss the cost economies?*

In an effort to preserve the aim of this manuscript while avoiding listing wholesale prices, we have eliminated table 1 and removed the specific companies and cost comparisons in the main text. We hope that this still helps readers make an informed decision about building versus ordering arenas.

*8. Please remove all commercial language from your manuscript and use generic terms instead. All commercial products should be sufficiently referenced in the Table of Materials and Reagents. For example: Logitech Pro Stream Webcam, Belkin, NCH, etc.*

We have omitted the original Table 1 and removed all commercial language from the manuscript and Table 2 (now the new Table 1). They are still referenced in the Table of Materials and Reagents.

*9. Please revise the protocol text to avoid the use of any personal pronouns (e.g., "we", "you", "our" etc.).*

We have revised the protocol text to avoid use of personal pronouns.

*10. Please revise the protocol to contain only action items that direct the reader to do something. 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.”*

The protocol has been streamlined to only contain actions in the imperative tense. Intervening text has been moved to the discussion or simplified as a note. Specifically, the long bodies of text describing statistical methods are now in the discussion section. Notes and additional comments previously in section 1 are now incorporated into the appropriate part of the discussion section.

*11. In the JoVE Protocol format, “Notes” should be concise and used sparingly. They should only be used to provide extraneous details, optional steps, or recommendations that are not critical to a step. Any text that provides details about how to perform a particular step should either be included in the step itself or added as a sub-step. Please consider moving some of the notes about the protocol to the discussion section.*

Several notes have either been moved to the discussion section or modified to be an action (see comment 10).

*12. 1.7: Please specify which two pieces.*

The appropriate 2 pieces are now specified in the text.

*13. 3.1: Please specify the age, gender and strain of mice.*

The age, genders and strain of the mice have been added.

*14. The Protocol should be made up almost entirely of discrete steps without large paragraphs of text between sections. Please simplify the Protocol so that individual steps contain only 2-3 actions per step and a maximum of 4 sentences per step. Please move the discussion about the protocol to the Discussion.*

We have extensively modified the protocol section as detailed in our responses to comments 10 and 11.

*15. After you have made all the recommended changes to your protocol (listed above), 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.*

We have highlighted the appropriate text for filming.

*16. Please include all relevant details that are required to perform the step in the highlighting. For example: If step 2.5 is highlighted for filming and the details of how to perform the step are given in steps 2.5.1 and 2.5.2, then the sub-steps where the details are provided must be highlighted.*

The appropriate sub-steps for filming have been highlighted.

*17. References: Please do not abbreviate journal titles. Please include volume and issue numbers for all references.*

We have corrected the reference to include full journal titles and volume numbers. We could not include issue numbers for all references as some references do not have issue numbers.

**Reviewers' comments:**

**Reviewer #1:**

*-Line 97- The OLT & NORT are identical until the test phase, when the objects are replaced/moved.*

We have modified the text to say “test phase” instead of “ITI”.

*-Section 1- Very well written. I would suggest presenting an alternative approach using brackets (i.e., Makerbeam) for any scientists that may not feel confident with the cement procedure. You could highlight that your procedure is the cheapest while presenting the alternatives. Also, there is no mention of any safety precautions that should be followed with the cement. Are there any requirements for protective clothing, ventillation, etc.. ?*

We agree that a viable alternative to solvent welding is the more traditional use of brackets. The use of brackets, like Makerbeam, or other fasteners would result in stronger and more durable arenas. However, from our experience, the drilling of pilot holes and use of fasteners with 0.25” thick acrylic is troublesome for multiple reasons. First, the diameter and depth required for drilling pilot holes entails a level of precision that can only be obtained with a drill press. Further, with a thickness of 0.25” overtightening a fastener risks cracking the entire piece as acrylic sheet is both hard and brittle. This can be alleviated by using a thicker 0.375” acrylic sheet which we have previously used for other behavioral test equipment; however, these thicker sheets are only available in clear. Therefore, one would need to sand, prime and paint the arena to eliminate any incidental visual stimuli if one choose to use 0.375” thick acrylic. This method requires familiarity and confidence with power tools which is not always accessible in a biology/neuroscience lab. We therefore restrict our presentation to the solvent welding technique that requires no specialized equipment or skills beyond what a typical lab would already have on hand.

*-Section 3.11- Most animal facilities are now requiring sterilization with unscented bleach germicidal wipes or clidox S (or comparable) in addition to an alcohol cleaning step. I would suggest adding this into the procedure everywhere where there is an alcohol cleaning step.*

We have added “Clean all arenas with animal facility recommended cleaning methods such as wiping with 70% ethanol to minimize olfactory cues before next use.” to all appropriate steps in the protocol to account for different cleaning procedures.

*-Section 3.12.1- Authors state objects should not be too close to the walls, but don't define the desired distance until much later in the manuscript (Fig 2 legend). I would add the 6cm x 6cm recommended spacing here.*

We have added 6 cm x 6 cm to the text.

*-Section 3.12.6- The authors mention the ITI can be varied, which is true. However it should be emphasized that varying this ITI changes the type of memory that is probed by the test. I.e., Short, intermediate, long, and remote memories.*

We agree with this comment and have expanded our discussion of ITIs and the importance of their length in the discussion section.

*-Section 3.14- The authors present the procedure as OLT and then NORT. A variation is NORT and then OLT. It would be good to suggest varying which test is done first to ensure that the order of the testing does not influence performance in either test.*

We agree with this comment and have added a section in the discussion regarding the reversal of testing order or execution of the OLT and NORT separately.

*-Section 4- This is extremely important for the performance of section 3. I would change the order and move Section 4 to Section 3 and vice versa.*

We agree with this comment and have made the appropriate changes to the manuscript (see overall changes above).

*-Section 5- Manual scoring is cheaper than an automated system, but the authors should present at least a high-level analysis of time spent analyzing data manually vs. automated systems, especially considering >=2 scientists must score the videos. I agree the up-front costs of implementing manual scoring are much less than using an automated system, but once validated, an automated system can save a significant amount of time (and money) especially when conducting large studies. Is there an inflection point where once an experiment hits a certain size and/or if a lab performs a certain number of experiments per month/year, it is more cost effective to implement an automated scoring system?*

The reviewer brings up a valid point regarding upfront cost savings of manual scoring versus the relative time-savings as this cost is spread over several experiments. This is a difficult issue to address as each lab considering behavioral testing will have different needs in terms of number of experiments, number of mice, and number of people involved in execution and analysis. This difficulty is compounded by the complicated nature of defining the cost of a researcher’s time, especially since this also depends on each researcher’s efficiency and training level. For example, undergraduate student researchers receiving academic credit can frequently be trained to score videos and the “cost” of their time is different than the time of a postdoctoral fellow or research scientists. To address this concern without delving too far in to the many variations entailed, we have added a statement regarding the potentially time-consuming nature of manual analysis and a recommendation that each lab should evaluate the cost of purchasing software versus the cost of the researcher’s time.

*-Section 5.2- How is the video calibrated such that the scientists scoring the data know when a mouse is <= 2cm from an object?*

Thank you for bringing this omission to our attention. Typically, the area can be demarcated on the screen during video analysis by placing a clear sheet with a circle around each object that denotes a 2 cm boundary. We have added the appropriate text to the protocol to clarify this process.

*-Section 5.3.1- While in theory, >50% indicates preference for an object, what is the typical variability of the novelty preference? Can the authors also provide guidance for a practical score indicating novelty preference that would be above noise? Can the authors perform a power analysis of the data they present in the manuscript to inform signal:noise and minimum number of subjects?*

The reviewer brings up an important discussion point here that we did not address in the original manuscript. The OLT and NORT can only reveal memory in the aggregate. Any individual mouse’s score cannot be interpreted as memory or lack of memory because any individual mouse could show a percent time investigating an object that varied from 50% for a variety of reasons. Memory can only be concluded by looking at a group of multiple mice and performing some kind of statistical analysis to determine whether the difference that emerged is likely to occur from chance variation. We have modified some of the data analysis section (now in the discussion primarily) to reflect this distinction more clearly.

We have also added a power analysis of the data presented in figure 3A and B to the discussion of data analysis, though we also note that the sample sizes and power will all depend heavily on effect size and variability which can vary in different sets of mice (of different ages, sexes or exposed to different treatments, for example).

As for what constitutes noise, in our re-framing of the data analysis, we emphasize that no individual data point can be interpreted alone, but rather grouped data must be used and compared statistically to determine whether memory is detected or whether differences are detected between groups. We are not aware of an agreed upon fixed level of object preference that is considered a threshold for being meaningful. We therefore restrict ourselves to providing the guidelines for finding statistically significant differences by relatively standard methods within behavioral neuroscience.

*-Section 5.4.4- I agree with the point the authors are making here, but it is statistically incorrect to predicate the argument around p-values. I would stick to means and confidence intervals here and remove the references to p-values.*

While we agree that reliance on p values is statistically invalid, it is the measure most commonly reported for these tasks and is even frequently required by journals to publish results of tests like these. To make our recommendations as useful and practical as possible, we would like to retain the p-value discussion, but with some added discussion of the reviewer’s point.

*-Line 409- Same as above- provide more guidance of what is a real signal, rather than referencing the theoretical >50% as novelty preference.*

This concern is discussed above.

*-Line 421- Emphasize that one cannot combine the 2 statistical approaches suggested in the figure legend.*

We have added this emphasis to the figure legend as well as the results section to specify that the two statistical approaches are separate methods of analysis.

*-Line 424- Authors state NORT is highly sensitive to intrinsic value of objects. Authors should indicate that OLT is also highly sensitive to objects.*

This statement has been added to the discussion section where we have also expanded our discussion of object selection and validation.

*-Line 511- Dimensions indicated in text are," 16m x 16m x 16 m". I think the authors meant, "16 in x 16 in x 16 in", which is closer to the size of the acrylic panels indicated in the methods.*

The original text is actually correct because those dimensions refer to the size of the testing room itself, not the arenas. To clarify this, we have modified the sentence to state that the 16m refers to each dimension of the room.

*-Figure 3- Any data on performance of OLT with objects that do not have equivalent intrinsic values to mice?*

While we recognize that providing results for an OLT with objects of differential intrinsic values could further demonstrate the importance of object validation and selection, we have not tested objects with inherent value differences in an OLT.

*-Figure 3F- Are n=4 animals enough to validate an object given the variance observed in the other panels?*

We originally recommended a minimum of 4 animals for validation, requiring that any hint of object preference is carefully heeded. Our test with n = 4 revealed a subtle aversion to one object that became significant with n = 16. In retrospect, it would be more prudent test at least 8 mice to make subtle preferences/aversions more evident.

*-Table 1- I would recommend adding the CleverSys system to the comparison table.*

In the revised manuscript, we have omitted this table per the Editor’s request to remove all specific vendors from the manuscript.

*-Table 2- Stanley (SAE/metric comb square) is misspelled.*

We have corrected the spelling.

**Reviewer #2:**

*The authors are overly conservative about the brain regions involved in NLT and NOR. PFC and Hippocampus are both involved.*

We agree with this comment and have made broad changes to the manuscript (see overall changes above).

*Add to the discussion of the criteria to choose the test objects. Heavy, non-displaceable objects are best for NOR and NLT tasks. Avoid Wooden objects which mice give preference to. Objects that are made of Plexiglas, glass, or metals are preferred*

We appreciate this suggestion and have added a section in our discussion that addresses this issue in greater detail.

*The authors have should discuss criteria for excluding data, e.g. when exploration time is <5s .*

We agree that this was an omission in our original manuscript. We have added criteria for excluding data in the discussion and protocol sections of the revised manuscript.

*The problems with using an arena to study more than one animal at a time need further consideration. The animals can be distracted by a conspecific or if studying multiple animals at the same time adds to the noise from multiple lab personnel. Adding animals to a multianimal arena at different times can affect data collections*

The reviewer raises an important point regarding multi-animal arenas. We have added this to the discussion.

**Reviewer #3:**

*While the concept of offering an affordable alternative to commercially available behavior products is wonderful, and the authors offer ingenious method for doing so, the actual protocol for running OLT/ORT misses a few important points, especially in terms of troubleshooting or adapting the protocol for different cohorts of mice. Additionally, the authors should more explicitly refer to other protocols that go into greater detail, so that readers can figure out how to adapt the protocol to their specific needs and still be successful in running the assay (such as Ennaceur 2010 for more specifics on object selection; Ennaceur and Delacour 1988 for the original protocol, as well as the minimum exploration criterion; also Leger et al 2010 or Lueptow 2017 for further troubleshooting and overall experimental design).*

We thank the reviewer for the additional references. We have expanded our discussion to be more specific with regard to troubleshooting or modifying the protocol.

*The authors repeatedly state that ORT is "hippocampal-independent", which isn't exclusively true. Some researchers have found hippocampal lesions to interfere with ORT, especially when the task is shifted to encourage a more spatial strategy, such as use of spatial cues. This is an important point to make, as it is possible that in their combined protocol, which first calls for OLT and then ORT, may be encouraging a more spatial strategy, as the OLT is conducted first, and spatial cues are used. Specifically line 100-101 should address the possible role of hippocampus in ORT (as well as in the discussion-line 482).*

We agree with this comment and have made appropriate changes to the manuscript (see overall changes above).

*Lines 65ff: Be careful with the use of "positive" and "negative" here in the context of reinforcement and punishment, as these terms have specific connotations within the field (i.e., positive refers to adding a motivating factor that alters the likelihood of a specific behavior, while negative removes an unpleasant factor to influence the likelihood of the behavior.) Specifically, I would try to reword parts of the paragraph to avoid the confusion (especially line 71- here MWM is relying on negative reinforcement for motivation, essentially, but say this instead of "negative stressor"; also line 76- negative and positive reinforcement)*

We thank the reviewer for this suggestion and have replaced negative or positive “stressor” with “reinforcement” as appropriate.

*1. Are there any cheap alternatives for a round arena? Some labs prefer to use a round arena to encourage locomotor behavior in mice that may be high anxiety, in order to discourage sitting in the corners of a square arena.*

The reviewer asks an interesting question. A circular arena would have certain advantages over a rectangular open field in terms of discouraging sitting in the corners. Also, its construction would be much simpler. Unfortunately, all commercially available materials are cost prohibitive as a cast acrylic tube 18” in diameter is $630 per foot, making a 16” high round arena just under $1000 in total materials. Additionally, acrylic tubing of that size would only be available in clear. Therefore, one would need to sand, prime and paint the arena to eliminate any incidental visual stimuli, requiring specific experience with materials and construction that is not a prerequisite for the current protocol.

*2. (from part 4 or line 309ff) As the authors state, running the pilot studies to discover potential object preference/aversion is very necessary. The object selection is critical to success of this assay and should be done prior to any other testing. Therefore, the section should be moved before the section on habituation, and further emphasis should be given to the importance of this step. Also, it would be helpful to either further discuss necessary considerations in object selection, or explicitly direct readers to a paper that does. One important consideration is whether or not mice can climb on the object. Using more mouse-sized objects that allow for climbing may encourage exploration, but it should be noted that time sitting on the object is NOT included in the overall exploration time (I see it's briefly mentioned in 5.2, but is worth further emphasis)*

We agree with this comment and have made broad changes to the manuscript (see overall changes above). In addition, we have added exclusion criteria in the protocol and discussion for when mice climb on top of objects.

*3. Objects and locations should be counterbalanced (this should be mentioned in 3.12.1)*

We have added this statement to the text.

*4. I'm not completely convinced that the pictured objects are the best representative objects. They are greatly different in size, and the shorter one might more greatly encourage climbing. It also has more textural features, which could affect preference. (see Ennaceur 2010 for further discussion)*

We agree with this comment and have made necessary changes to the protocol and discussion sections of the manuscript as well as Figure 1C. While we recognize the reviewer’s concern regarding this object, in our experience with our mice, they have not been able to successfully climb and perch atop this object. However, this does not preclude the possibility that other labs with different types and sizes of mice may experience mice climbing atop shorter objects. To prevent this from affecting behavioral experiments, we recommend in the protocol and discussion sections thorough validation of objects with mice that are representative of the age, strain and gender to be used in actual experimentation.

*5. There is no mention of a minimum criterion level for exploration of the objects. If a mouse does not sufficiently explore the object, how can you verify that any "learning" has been achieved? Generally 20s has been used as a minimum level of necessary exploration, though somewhat arbitrarily. However, some value should be defined.*

We agree that this was an omission in our original manuscript. We have added a statement regarding minimum criterion for investigation of objects in the discussion and protocol sections of the revised manuscript.

*6. Is there a reason to run OLT and then ORT? As mentioned earlier, perhaps the use of cues and spatial memory to train and test prior to the ORT could bias the learning or memory strategy for the ORT to be more hippocampal-dependent. Have the authors considered running ORT first and then OLT?*

We agree with this concern and have added a section in the discussion regarding the reversal of testing order or execution of the OLT and NORT separately.

*7. Data analysis: It should be noted that while both methods for data analysis are acceptable, if there is great variability in the exploration time between mice, it is perhaps better to control for those differences by using the discrimination index.*

The data analysis section has been moved to the discussion section and greatly expanded to account for all of the reviewers’ comments. Furthermore, a statement regarding the use of a discrimination index to better depict OLT and NORT data when total object investigation times vary greatly between individual mice has been added to the results section.

*8. Discussion: one additional critical component of successfully running this assay is in choosing an appropriate ITI. The authors mention it briefly, but I think a slightly expanded discussion (or reference to another paper that discusses it) would be important. Someone running the test for the first time may not understand the importance of the ITI and miss the correct window to see an effect, especially if they only try one ITI in their cohort.*

We agree with this comment and have expanded our discussion of ITIs and the importance of their length in the discussion section.