8.1 Category Learning Research in the Interactive Online Environment *Second Life*

Category Learning Research in the Interactive Online Environment *Second Life*

Jan Andrews
Vassar College Program in Cognitive Science
andrewsj@vassar.edu;

Ken Livingston
Vassar College Program in Cognitive Science
livingst@vassar.edu;

Joshua Sturm
Vassar College Program in Cognitive Science
josturm@vassar.edu;

Daniel Bliss
Vassar College Program in Cognitive Science
daniel.p.bliss@gmail.com;

Daniel Hawthorne
Vassar College Program in Cognitive Science
freedijym@gmail.com;

Abstract. The interactive online environment *Second Life* allows users to create novel three-dimensional stimuli that can be manipulated in a meaningful yet controlled environment. These features suggest *Second Life*'s utility as a powerful tool for investigating how people learn concepts for unfamiliar objects. The first of two studies was designed to establish that cognitive processes elicited in this virtual world are comparable to those tapped in conventional settings by attempting to replicate the established finding that category learning systematically influences perceived similarity. From the perspective of an avatar, participants navigated a course of unfamiliar three-dimensional stimuli and were trained to classify them into two labeled categories based on two visual features. Participants then gave similarity ratings for pairs of stimuli and their responses were compared to those of control participants who did not learn the categories. Results indicated significant compression, whereby objects classified together were judged to be more similar by learning than control participants, thus supporting the validity of using *Second Life* as a laboratory for studying human cognition. A second study used *Second Life* to test the novel hypothesis that effects of learning on perceived similarity do not depend on the presence of verbal labels for categories. We presented the same stimuli but participants classified them by selecting between two complex visual patterns designed to be extremely difficult to label. While learning was more challenging in this condition, those who did learn without labels showed a compression effect identical to that found in the first study using verbal labels. Together these studies establish that at least some forms of human learning in *Second Life* parallel learning in the actual world and thus open the door to future studies that will make greater use of the enriched variety of objects and interactions possible in simulated environments compared to traditional experimental situations.

1. Introduction

The study of how people acquire and represent knowledge of category concepts is a broad area of research in cognitive science and psychology that includes a wide variety of issues and approaches. The human ability to group objects into categories and thereby treat them as equivalent for certain purposes is fundamental to human cognition, providing a foundation for memory, language, and reasoning.

The process by which new category concepts are acquired is also highly relevant to the study

of learning and therefore to the field of education. One method of studying this process is to teach adults (or children) unfamiliar categories and observe resulting changes in perceptual judgments of category instances. Using stimuli of various kinds, both physical objects and computer-generated images, our laboratory and others have demonstrated that certain effects generally occur (see, e.g., [1], [2]): namely, objects classified together are treated as more alike, an effect we call *compression*, and/or objects classified differently are treated as more distinctive, an effect we call *expansion*.

The broad conception of category learning assumed by this approach (and the field more generally) treats potential category instances as consisting of a set of values of various features or dimensions. For example, a particular dog would have values on such dimensions as body size, furriness, color of fur, length of tail, and so forth. We propose that compression and expansion effects caused by learning a new category distinction essentially constitute a change in the way these dimensions are represented, a kind of warping of the psychological dimensional space, such that the learner becomes more sensitive to dimensions important to the category distinction and less sensitive to those irrelevant to the distinction. This results in the psychological clustering of items that are classified together, allowing objects that differ from each other in category-irrelevant ways to cohere with one another and contrast with objects clustered in other categories.

In a typical category learning experiment, participants are shown a series of items, usually from a set of artificial stimuli created by the experimenters, and trained to classify them into two categories by receiving feedback on their classification responses. Training is stopped when classification accuracy reaches a high level or a certain number of runs through the stimuli ("blocks") have occurred. Participants then judge a large number of pairs of stimuli and either rate their similarity or decide whether the objects are identical or not, to determine how alike or confusable objects are within or between categories. A control group of participants judges the same stimulus pairs in the same way without having received classification training.

To test for compression/expansion effects, it is necessary to use unfamiliar categories of objects, since there can be no "control" group for categories that are already known. In addition, for the trained group, only successful learners' data are included because the effects are hypothesized to arise from acquisition of the category concepts. Interestingly, compression and expansion appear to also require multidimensional objects and do not seem to occur for objects consisting of a single dimension of variation (see [3] for a famous counterexample, and [4] for evidence against this counterexample).

In order to ensure that the learning processes tapped in this experimental paradigm are relevant to learning in the real world, it would be very helpful to apply it in a more flexible, dynamic, and interactive environment than that of the standard research laboratory where static stimuli are displayed one at a time on a flat computer display and the participant enters responses via a keyboard press. The research reported in this paper represents an effort to recreate this category learning paradigm in the interactive online 3-D environment *Second Life*. While some scholars have discussed the potential utility of *Second Life* for scientific research (e.g., [5], [6]), we are not aware of its previous use for studying category learning. The purpose of the first study was to determine whether compression/expansion effects would also occur in a category learning task in *Second Life*, in order to establish continuity between the cognitive/perceptual processes being tapped in standard laboratory tasks and in *Second Life*. If successful, this study would support the use of *Second Life* for testing new hypotheses related to category learning.

2. EXPERIMENT 1

2.1 Method

2.1.1 Participants A total of 44 Vassar College undergraduates either volunteered as part of an introductory psychology class requirement or were paid for their participation.

2.1.2 Materials All objects, including stimuli used in the categorization experiment, were built using the *Second Life* (SL) build tools (version 1.19.1 of SL). Categorization stimuli were

inspired by those used in [7]. Each of the 32 different stimuli was constructed by first resizing and then linking four separate objects, or "prims," in the language of SL, to a central sphere: two rectangular, and two pyramidal (see Figure 1). The spherical center of each object was wrapped with the "horizontal stripes" texture. The four separate objects were then attached to the central sphere of each stimulus, with each of these protrusions possessing the default texture. The first protrusion was attached to the top of the sphere and consisted of an inverted red, pyramidal prim, linked to a red rectangular prim. This top two-part protrusion was set to glow at an intensity of 0.10, to enhance its saliency. The second protrusion was a white pyramidal prim attached to the lower, left-hand portion of the central sphere and was the same for all stimuli, as was the third protrusion, a white rectangular prim attached to the lower, right-hand portion of the central sphere.
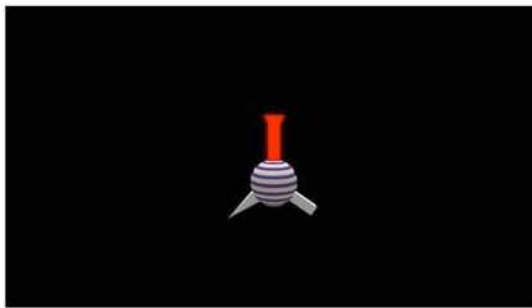

Figure 1. Example of a stimulus used in Experiment 1

The stimuli varied on two dimensions: the height of the top protrusion and the width of the stripes covering the central sphere. There were eight values of each dimension. The top protrusion varied in increments of 0.5m between stimuli and ranged from 2.5m to 6.0m. Stripe width varied from 0.4 to 13.1 repetitions per object. In order to make the increments of this variation perceptually uniform, a set power function of 1.357 was applied. This implies that each increase in stripe width between stimuli corresponded to 2.5 just-noticeable-difference units.

The 32 stimuli were subdivided into 16 Gexes and 16 Zofs (nonsense labels selected for low associability). Members of the Gex category possessed the widest stripes and the longest top

protrusions. Members of the Zof category possessed the narrowest stripes and the shortest top protrusions. The stimuli were displayed on a black rectangular background.

A mobile seat (the pink square in Figure 2) was also created and then programmed using the Linden Scripting Language (LSL) to respond to the click of a participant's mouse. The participant navigated this world of Gexes and Zofs as an androgynous avatar.
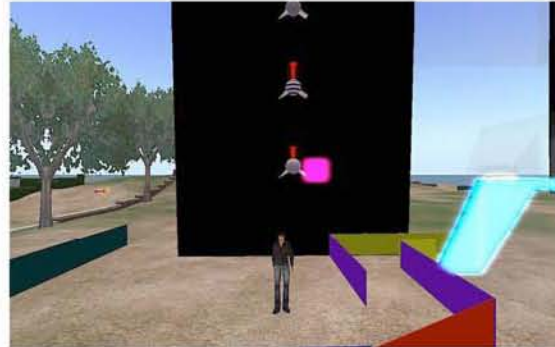

Figure 2. The participant avatar and experimental setting used in Experiment 1

2.1.3 Procedure Participants were randomly assigned to the learning or control condition. They were tested individually and entered SL using a Macintosh computer running OSX 10.4.11. Positioned in front of the screen, participants were told that they would first use the arrow keys on the keyboard to navigate the avatar through a short orientation path, in order to become familiar with the skills needed to complete the task. Once through the path, the participants were asked to click on the virtual seat in front of the avatar to fix the avatar's view in "mouse look" mode (first-person perspective).

In the classification task, all 32 stimuli were arrayed vertically in a different random order against each of eight black backgrounds, one behind the other (see Figure 2). The same random orders were used across participants. Participants viewed the stimuli one-by-one while moving upward on the seat, moving past the objects displayed against the black background. The seat was programmed to stop at each stimulus location. After each stimulus was shown, participants indicated which category they thought the object belonged to by clicking either the "Gex" or "Zof" button that

975

appeared in the upper right-hand corner of the screen. Immediately following the button press they received auditory feedback concerning the correctness of the response. Arrows directed participants to each successive set of stimuli (with a new black background and seat). Participants completed all eight blocks unless they achieved a total of at most one incorrect response in two consecutive blocks, at which point training was stopped immediately.

After the classification task, participants rode on a seat to the similarity task area. There they viewed the 32 stimuli in pairs, seeing the objects in a given pair one at a time for 3 seconds each. Participants rated the similarity between members of each pair on a scale from 1 to 9 (1 being least similar and 9 being most similar) by clicking one of nine buttons that appeared in the upper, right-hand corner of the screen. A total of 90 pairs of stimuli (30 Gex-Gex, 30 Zof-Zof, 30 Gex-Zof) were rated, and the pairs were presented in the same random order across participants. The 90 pairs were split into six blocks of 15 pairs; as above, blocks were separated onto different black backgrounds, and participants followed arrows to each successive block.

Control condition participants completed only the similarity judgment task, without any prior category learning.

2.2 Results

Four of the 24 participants in the learning condition failed to pass the learning criterion – they did not complete two consecutive blocks with a total of at most one incorrect response in the classification task – so their data were not included in the analysis.

A 2 (group: learning vs. control) by 2 (pair-type: members within the same category vs. members from separate categories) analysis of variance with repeated measures on the second factor was performed on the similarity ratings. This yielded a significant main effect of pair-type ($F(1,38) = 501.724$, $p < .001$) and a significant interaction between group and pair-type ($F(1,38) = 7.847$, $p = .008$). As shown in Figure 3, within-category pairs were judged to be more similar than between-category pairs by both groups, but the learning group judged within-category pairs to be more similar than did the control group and the between-category pairs to be less similar

than did the control group, corresponding to compression and expansion, respectively. However, planned one-tailed t-tests revealed that only the compression effect was significant ($t(38) = 1.748$, $p < .05$).
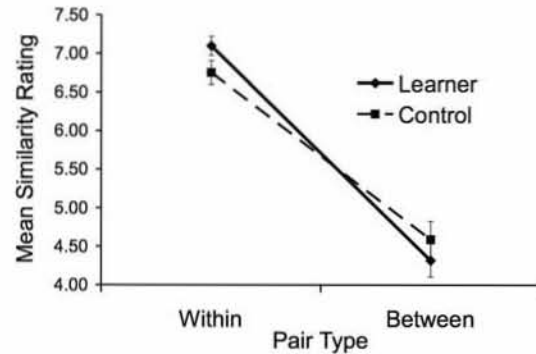

Figure 3. Results of Experiment 1

2.3 Discussion

Experiment 1's instantiation of the category learning paradigm described in the Introduction within the environment of *Second Life* produced a significant compression effect of the sort typically found in standard laboratory versions, suggesting that the category learning processes tapped are substantially the same. While our ultimate goal is to develop more innovative uses of *Second Life* for research on category learning, we next took advantage of some of the more straightforward features of the laboratory we had already constructed there to explore a long standing question about the role of verbal labels in category learning.

Normally when we learn new categories we simultaneously learn words for those categories. In fact, using a single word to refer to a set of objects that vary on several dimensions has been hypothesized to be central to the category learning process in humans (e.g., [8]). The single label provides an explicit feature common to all category instances and may support the learning of categories in important ways. To rigorously test this claim requires the study of category learning in the absence of verbal labels, and that poses a methodological challenge. We made creative use of *Second Life's* unique stimulus-building tools to meet this challenge.

976

In Experiment 2 we designed a novel response system to allow participants to learn the same categories of objects used in Experiment 1, but without any verbal labels. This allowed us to test two interesting questions: (1) Will category learning be more difficult under these conditions compared to those of Experiment 1? And, even if this is the case, (2) will those participants who do successfully learn the category distinction exhibit compression effects similar to those found in Experiment 1? That is, will the underlying mechanism of category learning be the same in the absence of labels?

3. EXPERIMENT 2

3.1 Method

3.1.1 Participants A total of 39 Vassar College undergraduates either volunteered as part of an introductory psychology class requirement or were paid for their participation.

3.1.2 Materials The stimuli and categories were identical to those used in Experiment 1. However, two visual patterns were used in place of the verbal labels "Gex" and "Zof." These patterns were colorful designs similar to tie-dyed fabric and slightly different from each other (see Figure 4). They were chosen because they are extremely difficult to describe in any simple way and thus do not easily lend themselves to description by a single label.
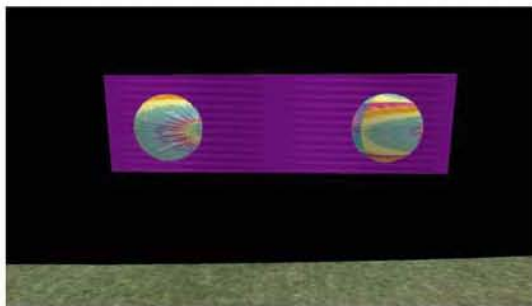

Figure 4. Nonverbal response buttons used in Experiment 2

3.1.3 Procedure The procedure was identical to that used in Experiment 1 except that the labels "Gex" and "Zof" were excluded from the classification judgments and feedback. Participants pressed one of the two buttons shown in Figure 4 to indicate which category they thought a stimulus belonged to and were simply told whether they were correct. The right-left positions of the two response buttons relative to each other were randomly varied to prevent participants from using the labels "right" and "left" for the two patterns.

3.2 Results

Nineteen of the 39 participants failed to meet the learning criterion used in Experiment 1, and their data were not included in analyses. A chi-square test of independence showed that learning success was related to the presence (Experiment 1) or absence (Experiment 2) of verbal labels (χ^2(1, N = 63) = 6.58, p = .01).

In order to determine whether compression or expansion occurred, the data from this experiment were combined with the control group data from Experiment 1 and analyzed in the same way. The resulting 2 (group: nonverbal vs. control) by 2 (pair-type: within-category vs. between-category) analysis of variance with repeated measures on the second factor yielded a significant main effect of pair-type (F(1,38) = 682.612, p < .001) and a significant interaction between group and pair-type (F(1,38) = 8.852, p = .005). As shown in Figure 5, the pattern of the means is identical to that of Experiment 1 and, once again, planned one-tailed t-tests revealed that only the compression effect was statistically significant (t(38) = 1.850, p < .05).

Additional t-tests showed that the mean similarity ratings for the learning groups in Experiment 1 (with verbal labels) and Experiment 2 (without verbal labels) did not differ significantly for either the within-category pairs or the between-category pairs.

3.3 Discussion

The use of complex visual patterns in place of verbal labels in Experiment 2 made the category learning task, otherwise identical to that used in Experiment 1, much more difficult. This is consistent with results reported by Lupyan, Rakison, and McClelland [8] using a similar task. As they note, while it is impossible to be certain that participants were not surreptitiously using invented verbal labels, the fact that learning was significantly more difficult using nonverbal category responses suggests that explicit verbal labels do facilitate category learning. It is interesting that half of our participants in the

977

nonverbal condition were nonetheless able to learn the categories to a very high level of accuracy.
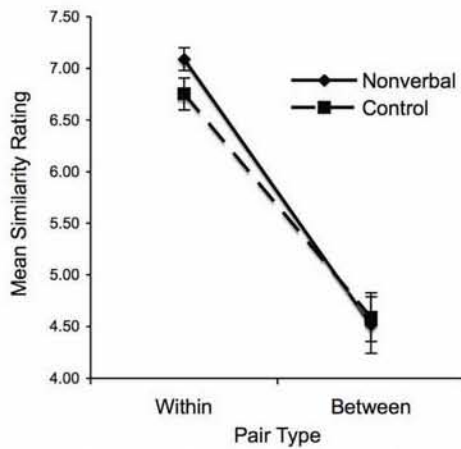

Figure 5. Results of Experiment 2

However, our main interest was in determining whether, for those participants who were able to master the category distinction, the compression result obtained in Experiment 1 would still occur. In fact, the compression effect for successful learners was virtually identical in the two experiments, suggesting strongly that while verbal labels may make category learning easier, they are completely unrelated to the compression/expansion processes associated with category learning. This is consistent with the idea that these processes are indeed fundamental to the formation of new category concepts.

4. CONCLUSION

These experiments on category learning in *Second Life* demonstrate both its continuity with standard laboratory research and its utility for testing new hypotheses. Our evidence that *Second Life* taps the same cognitive processes observed in laboratory research supports its validity for further research on learning and cognition.

Our next study, currently underway, makes much greater use of *Second Life*'s potential for creating dynamic, interactive stimulus features and engaging, goal-oriented tasks. We expect this will allow us to explore category learning processes in ways that are not possible in a standard laboratory situation but that are actually significantly more like the real world.

REFERENCES

1. Goldstone, R. (1994). "Influences of categorization on perceptual discrimination," *Journal of Experimental Psychology: General,* 123, pp. 178-200.

2. Livingston, K.; Andrews, J. (2005). "Evidence for an age-independent process in category learning," *Developmental Science,* 8(4), pp. 319-325.

3. Tajfel, H.; Wilkes, A. (1963). "Classification and quantitative judgement, " *British Journal of Psychology,* 54, pp. 101-114.

4. Andrews, J.; Livingston, K. (2000). "Accentuation of category differences: Revisiting a classic study." In L. Gleitman & A. Joshi (Eds.), Proceedings of the 22nd Annual Conference of the Cognitive Science Society. Mahwah, NJ: Erlbaum.

5. Boellstorff, T. (2008). Coming of Age in Second Life: An Anthropologist Explores the Virtually Human. Princeton: Princeton University Press.

6. Bainbridge, W. S. (2007). "The Scientific Research Potential of Virtual Worlds," *Science,* 317, pp. 472-476.

7. Cooke,T; Jakel, F.; Wallraven, C.; Bulthoff, H. (2007). "Multimodal similarity and categorization of novel three-dimensional objects," *Neuropsychologia,* 45, pp. 484-495.

8. Lupyan, G.; Rakison, D.; McClelland, J. (2007). "Language is not just for talking: Redundant labels facilitate learning of novel categories," *Psychological Science,* 18(12), pp. 1077-1083.