Statistical learning using real-world scenes: extracting categorical regularities without conscious intent

Timothy F. Brady and **Aude Oliva**
Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology

Abstract

Recent work has shown that observers can parse streams of syllables, tones, or visual shapes, and learn statistical regularities in them without conscious intent (e.g., learn that A is always followed by B). Here, we demonstrate that these statistical learning mechanisms can operate at an abstract, conceptual level. In Experiments 1 and 2, observers incidentally learned which semantic categories of natural scenes covaried (e.g., Kitchen scenes were always followed by Forest scenes). In Experiments 3 and 4, category learning with images transferred to words that represented the categories. In each experiment the category of the scenes was task-irrelevant. Together, these results suggest that unsupervised statistical learning mechanisms can operate at a categorical level, enabling generalization of learned regularities using existing conceptual knowledge. Such mechanisms may guide learning in domains as disparate as the acquisition of causal knowledge and the development of cognitive maps from environmental exploration.

One of the primary tasks of the brain is to extract regularities from the environment in order to make inferences and guide behavior in novel situations. Indeed, sensitivity to statistical regularities plays an important role in various domains of perception and cognition, from preparing motor actions (Nissen & Bullemer, 1987) to parsing language (Trueswell, 1996) and making higher cognitive judgments, such as predicting the amount of money a movie will earn (Griffiths & Tenenbaum, 2006).

Recent work on statistical learning has demonstrated that observers extract the covariance between syllables, tones or shapes that appear in a predictable order or spatial arrangement (Behrmann, Geng & Baker, 2005; Fiser & Aslin, 2002; Saffran, Aslin & Newport, 1996; Saffran, Johnson, Aslin & Newport, 1999; Turk-Browne, Junge & Scholl, 2005; Chun & Jiang, 1998). In a visual variant of such an experiment, observers are presented with a sequence of shapes, where, unbeknownst to the observers, the sequence consists of arrangements of four temporal 'triplets' in which the same three shapes always appear in the same order (e.g., *ABC*GHI*DEFABC*JKL*DEF*…). After a brief exposure to this stream, observers are able to reliably identify the triplets (e.g., ABC) as more familiar than foil sequences (e.g., AEI), despite the fact that they have seen all of the individual shapes an equal number of times. This indicates that they have learned which shapes appeared together, although their verbal reports indicate no awareness of the structure in the stream.

To date, these experiments have used simple novel shapes to investigate visual statistical learning. In the real world, however, the units over which learning must operate are considerably more information-rich, and we often learn novel statistical regularities that pertain to objects and environments with which we have many prior associations. Furthermore, objects and scenes in the world have semantic knowledge associated with

Please address correspondence to: Timothy Brady, Department of Brain and Cognitive Sciences, 46-4078, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, tfbrady@mit.edu.

them, and important regularities occur at multiple levels of semantic abstraction. For instance, the co-occurrence of different places when moving through an environment allows us to infer that our kitchen will lead to our living room (e.g., cognitive maps; Tolman, 1948), while also allowing us to construct a hierarchy of such associations, moving from specific instances (my kitchen and living room appear together) to basic-level categories (kitchens and living rooms tend to appear together) and more abstract representations (indoor rooms appear together).

The purpose of the present study was to examine the extent to which statistical learning mechanisms can automatically operate at an abstract, conceptual level. Specifically, we used a novel modification of the statistical learning paradigm to test whether observers incidentally learn category-level relationships, even when the task does not require processing information at a categorical level.

We used real-world scenes as a case study for investigating categorical statistical learning and transfer of learning across levels of representation. In Experiment 1, we demonstrate statistical learning with complex real-world scenes. In Experiment 2, we demonstrate that statistical learning mechanisms can extract regularities at a categorical level, by showing a transfer of learning between a particular image and another exemplar of the same category. Finally, in Experiments 3 and 4, we show that these categorical representations can be accessed lexically, which indicates that the regularities are present in an abstract format that is divorced from the visual details of the images.

EXPERIMENT 1: STATISTICAL LEARNING USING VISUAL IMAGES

Previous experiments investigating the extraction of statistics of covariance from the visual world have used simple novel shapes. In Experiment 1, we verify the robustness of the visual statistical learning mechanism to complex stimuli with which subjects had prior associations: visual scenes.

Method

Observers—Ten naïve observers were recruited from the MIT participant pool (age range 18–35) and received 5 dollars for their participation. All gave informed consent.

Apparatus and Stimuli—Stimuli were presented using MATLAB with the Psychophysics toolbox extensions (Brainard, 1997; Pelli, 1997). Twelve scene categories were used (see Figure 1): Bathroom, Bedroom, Bridge, Building, Coast, Field, Forest, Kitchen, Living Room, Mountain, Street, and Waterfall. Each category contained 120 different full-color images. For each observer, one picture was drawn from each of the 12 categories at random, resulting in a set of 12 different images. The images were centered and subtended 7.5×7.5 degrees of visual angle. Observers sat 60 cm from a 21" monitor.

Each of the 12 selected images was randomly assigned a position in one of four triplets (e.g., ABC) – sequences of three images which always appeared in the same order. Then a sequence of images was generated by randomly interleaving 100 repetitions of each triplet, with the constraint that the same triplet could never appear twice in a row, and the same set of two triplets could never appear twice in a row (e.g., ABCGHIABCGHI was disallowed). In addition, 100 repeat images were inserted into the stream such that sometimes either the first or third image in a triplet repeated immediately (e.g., ABCCGHI or ABCGGHI). Allowing only the first or third image in a triplet to repeat served to keep the triplet structure intact yet prevented the repeat images from being informative for delineating triplets from each other.

Procedure—Observers watched a 20 minute long sequence of 1000 images, presented one at a time for 300 ms each with a 700 ms inter-stimulus interval. During this sequence, their task was to detect and indicate back-to-back repeats of the same image by hitting the space bar as quickly as possible. This gave observers a cover task to help prevent them from becoming explicitly aware of the structure in the stream (Turk-Browne et al. 2005), as well as avoided having observers simply passively viewing the stream (which would make it unclear what observers were processing). Importantly, they were never informed that there was any structure in the stream of images.

Following this study period, observers were asked if they had recognized any structure in the stream, and then were given a surprise forced-choice familiarity test. On each test trial observers viewed two 3-image test sequences, presented sequentially at the center of the screen with the same ISI as during the study phase, and segmented from each other by an additional 1000 ms pause. One of these test sequences was always a triplet of images that had been seen in the stream (e.g., ABC) and another was a foil constructed from images from three different triplets (e.g., AEI). After the presentation of the two test sequences, observers were told to press either the 1 or 2 key to indicate whether the first or second test sequence seemed more familiar based on the initial study period. Each of the four triplets was tested eight times, paired twice with each of four different foil sequences (AEI, DHL, GKC, JBF) for a total of 32 test trials. Observers' ability to discriminate triplet sequences from foil sequences was used as a measure of statistical learning.

Results and Discussion

All ten of the observers completed the repeat detection task during the study period with few errors, detecting an average of 91% of the repetitions (SD = 5%) with between 1 and 5 false alarms. This demonstrates that observers were attending to the sequence of images. However, when asked no observers reported explicitly noticing that the study stream had any structure.[1] Nonetheless, performance on the familiarity test indicated very robust statistical learning, with triplets being successfully discriminated from foils (86.6% of the test sequences chosen were triplets; 13.4% were foils, $t(9) = 8.72$, $p = 0.00001$; See Figure 3). [2]

These results extend previous demonstrations of visual statistical learning in two ways. First, scene stimuli are the most complicated and information-rich stimuli for which statistical learning has been demonstrated. Second, choosing the correct triplets at test in this experiment required not just forming episodic associations between the correct scene pictures, but also overcoming prior knowledge about how the scenes these pictures represent are associated in the world (e.g bridges are rarely associated with living rooms). In the current experiment, learning likely occurred at the image-level, because identical stimuli were repeated throughout the learning and test phases (and statistical learning has been previously demonstrated for shape and color: e.g. Turk-Browne, Isola, Scholl & Treat, in press). Therefore, to examine the role of category-level semantics on statistical learning we performed a second experiment, in which the same string of images was never presented twice but a pattern occurred at the categorical level.

---

[1]Observers were asked "Did you notice any patterns in the stream of images?" followed by "So, for example, if I asked you what images generally followed Mountains, would you be able to tell me?" All observers responded negatively to both questions.
[2]For all experiments, a Lilliefors test fails to reject the null hypothesis that the data is normally distributed, $p > 0.10$ (Lilliefors, 1967). In addition, using Monte Carlo methods that take into account the actual distribution of the stimuli (the average over 10 subjects of a sum of 32 Bernoulli trials) indicates how unlikely our data is given the null hypothesis (all p's < 0.01).

EXPERIMENT 2: STATISTICAL LEARNING OF CATEGORICAL REGULARITIES

Recent work has suggested that statistical learning abstracts away at least some of the perceptual information in the input: learning can be preserved in spite of changes in color and even the order of the stimuli at test (Turk-Browne et al., 2007; Turk-Browne & Scholl, 2006). In the real world, however, it may be useful to learn at an even more abstract level: the conceptual rather than perceptual level. Categorization allows us to compress information, since rather than storing the same information about each possible object or scene, we can store it only one time and apply it to many different stimuli (Mervis & Rosch, 1981). This suggests that for statistical learning to be most useful (e.g., to maximize the number of stimuli to which a given regularity will apply), we should learn statistical relationships between semantic categories as well as between individual objects.

In order to learn information at the categorical level during our everyday visual experience, we would need to extract the category of a given stimulus even though such information is task-irrelevant in most of our everyday interactions. Experiment 2 therefore sought to determine (1) whether observers automatically extract the basic-level category of real-world scenes even when doing so is task-irrelevant, and (2) whether observers learn the statistics of covariation at this categorical level. We used a design much like Experiment 1, except that rather than repeating the same 12 pictures throughout the study and test period, a new picture from the same category was drawn each time a particular triplet was shown (see Figure 2). Thus, with the exception of the repeat detection task, no images were repeated throughout the entire experiment, and the statistical structure of the stream was defined only by the basic level category of the images. Observers still performed an image-level repeat detection task during the study period, and were unaware that the category of the images was relevant or that there was a statistical structure in the stream of images.

Method

Eleven naïve observers participated in this experiment. The apparatus and stimuli were identical to Experiment 1, with one exception: every time a triplet appeared in the stream, new images from the same category were used. Thus, the triplet ABC would be a particular mountain image, bathroom image, and street image the first time it was presented, and a different mountain image, bathroom image and street image the next time (e.g., $A_1B_1C_1G_1H_1H_1I_1D_1E_1F_1A_2B_2C_2...$). Repeats (for the repeat-detection task) were still of the exact same image, and observers were not informed that the images were drawn from a particular set of categories or that there was a structure in the stream. The test sequences consisted of entirely new images drawn from the same categories as the training images.

Results and Discussion

Ten of the observers completed the repeat detection task with few errors, detecting an average of 94% of the repetitions (SD = 3%) with between 1 and 8 false alarms. One other observer was excluded from further analysis for a high error rate on this task (having missed 22% of the repetitions). No observers reported noticing that the study stream had any structure.

Nonetheless, performance on the familiarity test indicated robust categorical statistical learning, with triplets successfully discriminated from foils (61.3% triplets; 38.8% foils; $t(9)$ = 2.61, p=0.028 - see Figure 3). This indicates that observers were able to learn that images from particular categories co-occurred with images from other categories, even when the categorization was entirely unrelated to the task being performed.

The result also demonstrates that people extract category information automatically when confronted with real-world images (e.g., Grill-Spector & Kanwisher, 2005). Although this corroborates intuition in the literature on scene gist (Oliva, 2005; Potter, 1976) and fits well with the literature on semantic priming (where pictures presented outside of conscious awareness result in priming that transfers to similar words, e.g. Dell'Acqua & Grainger, 1999), to our knowledge observers' ability to incidentally categorize visual scenes at the basic level while performing an unrelated task has not been previously reported.

EXPERIMENT 3: LEXICAL RECALL OF VISUALLY-LEARNED REGULARITIES

Experiment 2 demonstrated that statistical learning operates at a categorical level, forming associations based on the basic-level category of scenes. However, since the images used are real-world scenes, many of the images within categories share low-level properties with each other that images across categories do not (e.g., Oliva & Schyns, 2000; Oliva & Torralba, 2001). Thus, the statistical learning observed in Experiment 2 could be operating over low level regularities (color, orientations) that happen to covary strongly with image category, rather than directly on a more conceptual representation. This distinction is tied with debates about the representation of conceptual knowledge and whether this representation is modality-dependent (Barsalou, 1999) or more abstract (Pylyshyn, 1984), and furthermore is related to the degree to which categories are actually defined by visual similarity in the first place (Mervis & Rosch, 1981). However, for the purposes of a useful representation of knowledge, we suggest that if the representations formed by statistical learning can be accessed in the absence of the low-level image regularities, they should be thought of as conceptual rather than simply low-level (although low-level regularities are likely being learned as well; this probably accounts for why learning is significantly stronger in Experiment 1 than Experiment 2).

In order to probe the degree of abstraction present in the representations learned in our task, we changed the test sequences to words representing the categories rather than to images from the categories (e.g., in the triplet ABC, if A represented mountain images throughout training, A was the word "Mountain" at test). Expression of learning in this context would indicate that the regularities learned in this task result in the association of more than just low-level perceptual features of the images.

Method

Ten naïve observers participated in this experiment. The apparatus and stimuli were identical to Experiment 2, with one exception: at test words rather than images were presented. Thus, during training the triplet ABC would be a particular mountain image, bathroom image, and street image the first time it was presented, and a different mountain image, bathroom image and street image the next time. At test, this triplet would be represented by the words "Mountain", "Bathroom" and "Street".

Results and Discussion

All ten of the observers completed the repeat detection task with few errors, detecting an average of 92% of the repetitions (SD = 4%) with between 1 and 4 false alarms. No observers reported noticing that the study stream had any structure.

Nonetheless, performance on the familiarity test indicated robust statistical learning, with triplets of words being successfully discriminated from foils (61.1% triplets; 38.9% foils; $t(9) = 2.51$, $p=0.036$ - see Figure 3). This indicates that observers were able to incidentally

learn that images from particular categories co-occurred with images from other categories, and, surprisingly, were able to transfer this knowledge to a lexical task at test.

This suggests that the associations formed in this task are not solely based on low-level image regularities, and at the least can be expressed in the absence of these regularities. Thus, the learned information is automatically abstracted from the perceptual input into a high-level conceptual representation.

EXPERIMENT 4: AUTOMATIC ABSTRACTION OF CATEGORICAL REGULARITIES

Experiments 2 and 3 demonstrated that statistical learning operates at a categorical level, forming associations based on the basic-level category of scenes. However, in both of these experiments individual images were never repeated, meaning that the only regularity present in the stream was at the categorical level. Does categorical learning happens even when other regularities are present in the stream? Experiment 4 combined elements of Experiment 1 and Experiment 3 as follows: During the study stream, twelve identical images were repeated over and over again, but at test we examined whether observers could transfer their learning to a lexical test. Successful transfer would indicate that observers extracted the regularities at the categorical level even when regularities existed at the level of the individual items.

Method

Fifteen naïve observers participated in this experiment. The apparatus and stimuli were identical to Experiment 1, with one exception: during the test trials we presented words rather than images.

Results and Discussion

All fifteen of the observers completed the repeat detection task during the study period with few errors, detecting an average of 93% of the repetitions (SD = 5%) with between 1 and 3 false alarms. No observers reported noticing that the study stream had any structure.

Nonetheless, performance on the familiarity test indicated robust statistical learning, with triplets of words being successfully discriminated from foils (61.5% triplets; 38.5% foils; $t(14) = 2.14$, p = 0.050 - see Figure 3). This indicates that observers were able to incidentally learn at a categorical level.

This suggests that the categorical information is abstracted from the perceptual input into a high-level categorical representation, even when regularities exist at the perceptual level as well as at the categorical level. This provides further evidence for the automaticity of learning at the categorical level. The difference between the learning in this experiment and Experiment 1 is significant (86.6% triplets vs 61.5% triplets; $p = 0.003$) and indicates that observers are better able to identify which triplets they saw when the exact same images are used at test than when words are used at test. We suggest this is because observers are simultaneously learning at both the perceptual and categorical levels, and in Experiment 1 both of those sources of knowledge help choose the correct triplet at test, whereas in the present experiment only the conceptual regularities inform observers' choices at test.

GENERAL DISCUSSION

We examined for the first time whether statistical learning can operate at an abstract conceptual level, to discover the co-occurrence of semantic categories. In Experiment 1, we

showed that statistical learning operates over complex real-world scenes, extending the domain of visual statistical learning to semantically meaningful stimuli. In Experiment 2, we demonstrated that statistical learning can successfully extract categorical regularities in a stream of pictures even when the semantic category of the scene is task-irrelevant. In Experiment 3, we demonstrated that these learned regularities can be accessed lexically, suggesting that they are accessible in an abstract format. Finally, in Experiment 4, we show that observers learned the categorical structure of the stream even when they could have learned regularities at the perceptual level only. Importantly, in all experiments, observers were unaware of the presence of a structure in the stream and were unaware that category was relevant.

The present results extend our understanding of statistical learning mechanisms, demonstrating that they are powerful mechanisms for organizing the input we receive from the world, capable of extracting its underlying structure without conscious intent or awareness. Statistical learning mechanisms are capable of extracting many different regularities with only minutes of exposure (joint probabilities, Fiser & Aslin, 2002; non-adjacent dependencies, Newport & Aslin, 2004; etc.) and appear to be relatively ubiquitous, occurring in the auditory, tactile and visual domains, and in infants, adults, and monkeys (Conway & Christiansen, 2005; Kirkham, Slemmer & Johnson, 2002; Hauser, Newport & Aslin, 2001). They have been used to explain effects as diverse as how we segment a speech stream and how we create visual objects from low-level perceptual information (Saffran et al. 1996; Turk-Browne et al., 2007).

The present study demonstrates that statistical learning mechanisms operate at multiple levels of abstraction, including the level of semantic categories. Making use of categories should limit the amount of information that must be extracted and stored by statistical learning mechanisms, where computational limitations are particularly acute because of the large number of possible statistics and units over which such statistics could be computed (Brady & Chun, 2007; Turk-Browne et al. 2005). [3] In other words, in the real world where we need to track relationships between huge numbers of objects and locations, it would make sense to learn these statistics at a level where there is not a huge amount of redundancy, like the level of the semantic category.

The present results also demonstrate a surprising degree of abstraction in statistical learning, with transfer from regularities learned with images to a lexical test. There has been some debate over how closely statistical learning mechanisms are tied to the perceptual characteristics of the input, particularly in the literature on artificial grammar learning (e.g., Conway & Christiansen, 2006; Marcus et al, 1999). The transfer observed in the present study is between highly related stimuli (pictures from a category and the word representing that category), and therefore does not directly bear on the question of whether people learn amodal 'algebraic rules' that can be applied to any new stimuli (as has been argued by Marcus et al., 1999). However, the abstraction observed here does indicate that statistical learning need not be tied to the *perceptual* features of the input, as has been suggested previously (Conway & Christiansen, 2006).

Overall, our results suggest that statistical learning is about more than just organizing low-level perceptual inputs – it might be a useful tool for organizing our conceptual knowledge as well. The learning demonstrated here maybe a useful mechanism for encoding new

---

[3]Even the earliest statistical learning studies presumably operated over categories, although ones with no semantic meaning associated with them. For example, the results of Saffran et al. (1996) are interesting only if infants can learn not only that the particular sound waveform used in the experiment to represent 'pa' is often followed by the particular waveform used for 'bi', but that the syllable 'pa' is followed by the syllable 'bi'.

relationships between visual episodes, tied together in a particular context. For example, learning new cognitive maps can be thought of as extracting sequences of places that often occur together from a continuous stream of inputs (Tolman, 1948). In general, when confronted with a new environment with a novel distribution of stimuli, statistical learning may provide observers with the ability to implicitly learn the new distribution of regularities (or lack of regularities) in a context-dependent fashion, informed by, but not constrained by, the regularities found in other similar environments they have encountered.

More broadly, inferences about causality are often based on the observed covariance of causes (e.g. a block touches an apparatus) and effects (e.g. a noise is produced), and statistical learning is often cited as responsible for learning such covariance information, particularly in infants (Gopnik & Schultz, 2004). If infants possess categorical statistical learning mechanisms like those revealed by the present study in adults, they could have available the covariance information necessary to learn not just that a particular object creates a particular effect, but that all objects of a known category would create the same effect (e.g., intuitive theories: Carey, 1985). At this point, however, the utility of statistical learning mechanisms for building such high-level knowledge remains an open question. Assessing the scope and depth of knowledge that can be acquired via statistical learning mechanisms presents a challenge for future work.

Acknowledgments

REFERENCES

Barsalou LW. Perceptual symbol systems. Behavioral and Brain Sciences. 1999; 22:577–609. [PubMed: 11301525]

Behrmann, M.; Geng, JJ.; Baker, CI. Acquisition of long-term visual representations: Psychological and neural mechanisms. In: Uttl, B.; Ohta, N.; Macleod, C., editors. Dynamic cognitive processes: The Fifth Tsukuba International Conference; Tokyo: Springer Verlag; 2005. p. 1-26.

Brady TF, Chun MM. Spatial constraints on learning in visual search: Modeling contextual cueing. Journal of Experimental Psychology: Human Perception and Performance. 2007; 33(4):798–815. [PubMed: 17683229]

Brainard DH. The Psychophysics Toolbox. Spatial Vision. 1997; 10:433–436. [PubMed: 9176952]

Carey, S. Conceptual change in childhood. MIT Press; 1985.

Chun MM, Jiang Y. Contextual cueing: implicit learning and memory of visual context guides spatial attention. Cognitive Psychology. 1998; 36:28–71. [PubMed: 9679076]

Conway CM, Christiansen MH. Modality-constrained statistical learning of tactile, visual, and auditory sequences. Journal of Experimental Psychology. Learning, Memory, and Cognition. 2005; 31(1):24–39.

Conway CM, Christiansen MH. Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. Psychological Science. 2006; 17(10):905–912. [PubMed: 17100792]

Dell'Acqua R, Grainger J. Unconscious semantic priming from pictures. Cognition. 1999; 73:B1–B15. [PubMed: 10536225]

Fiser JZ, Aslin RN. Statistical learning of higher-order temporal structure from visual shape sequences. Journal of Experimental Psychology: Learning, Memory, & Cognition. 2002; 28(3):458–467.

Gopnik A, Schultz L. Mechanisms of theory formation in young children. Trends in Cognitive Science. 2004; 8(8):371–377.

Griffiths TL, Tenenbaum JB. Optimal Predictions in Everyday Cognition. Psychological Science. 2006; 17(9):767–773. [PubMed: 16984293]

Grill-Spector K, Kanwisher N. Visual recognition: as soon as you know it is there, you know what it is. Psychological Science. 2005; 16(2):152–160. [PubMed: 15686582]

Hauser MD, Newport EL, Aslin RN. Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. Cognition. 2001; 78(3):B53–B64. [PubMed: 11124355]

Kirkham NZ, Slemmer JA, Johnson SP. Visual statistical learning in infancy: evidence for a domain general learning mechanism. Cognition. 2002; Vol. 83:35–42.

Lilliefors H. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. Journal of the American Statistical Association. 1967; 62:399–402.

Marcus GF, Vijayan S, Bandi Rao S, Vishton PM. Rule-learning in seven-month-old infants. Science. 1999; 283:77–80. [PubMed: 9872745]

Mervis CB, Rosch E. Categorization of Natural Objects. Annual Review of Psychology. 1981; 32:89–115.

Newport EL, Aslin RN. Learning at a distance I. Statistical learning of non-adjacent dependencies. Cognitive Psychology. 2004; 48:127–162. [PubMed: 14732409]

Nissen MJ, Bullemer P. Attentional requirements of learning: Evidence from performance measures. Cognitive Psychology. 1987; 19:1–32.

Oliva, A. Gist of the scene. In: Itti, L.; Rees, G.; Tsotsos, JK., editors. Neurobiology of attention. San Diego, CA: Elsevier; 2005. p. 251-256.

Oliva A, Schyns PG. Diagnostic colors mediate scene recognition. Cognitive Psychology. 2000; 41:176–210. [PubMed: 10968925]

Oliva A, Torralba A. Modeling the shape of the scene: A holistic representation of the spatial envelope. International Journal of Computer Vision. 2001; 42:145–175.

Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision. 1997; 10:437–442. [PubMed: 9176953]

Potter MC. Short-term conceptual memory for pictures. Journal of Experimental Psychology: Human Learning and Memory. 1976; 2:509–522. [PubMed: 1003124]

Pylyshyn, ZW. Computation and cognition. MIT Press; 1984.

Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. Science. 1996; Vol. 274:1926–1928. [PubMed: 8943209]

Saffran JR, Johnson EK, Aslin RN, Newport EL. Statistical learning of tone sequences by human infants and adults. Cognition. 1999; 70(1):27–52. [PubMed: 10193055]

Tolman EC. Cognitive maps in rats and men. Psychological Review. 1948; 55:189–208. [PubMed: 18870876]

Trueswell JC. The Role of Lexical Frequency in Syntactic Ambiguity Resolution. Journal of Memory and Language. 1996; 35:566–585.

Turk-Browne NB, Scholl BJ. The space-time continuum: Spatial visual statistical learning produces temporal processing advantages. Journal of Vision. 2006; 6(6):676–676a. [Abstract] doi: 10.1167/6.6.676 http://journalofvision.org/6/6/676/.

Turk-Browne NB, Isola PJ, Scholl BJ, Treat TA. Multidimensional visual statistical learning. Journal of Experimental Psychology: Learning, Memory, and Cognition. (in press).

Turk-Browne NB, Jungé J, Scholl BJ. The automaticity of visual statistical learning. Journal of Experimental Psychology: General. 2005; 134:552–564. [PubMed: 16316291]
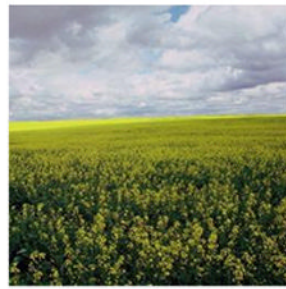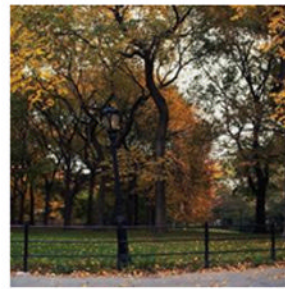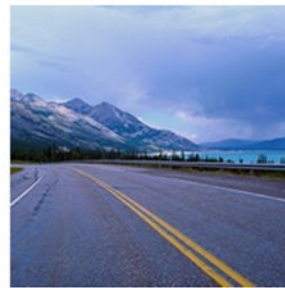
Figure 1.
Examples from each of the categories of real-world scenes used in all three experiments.
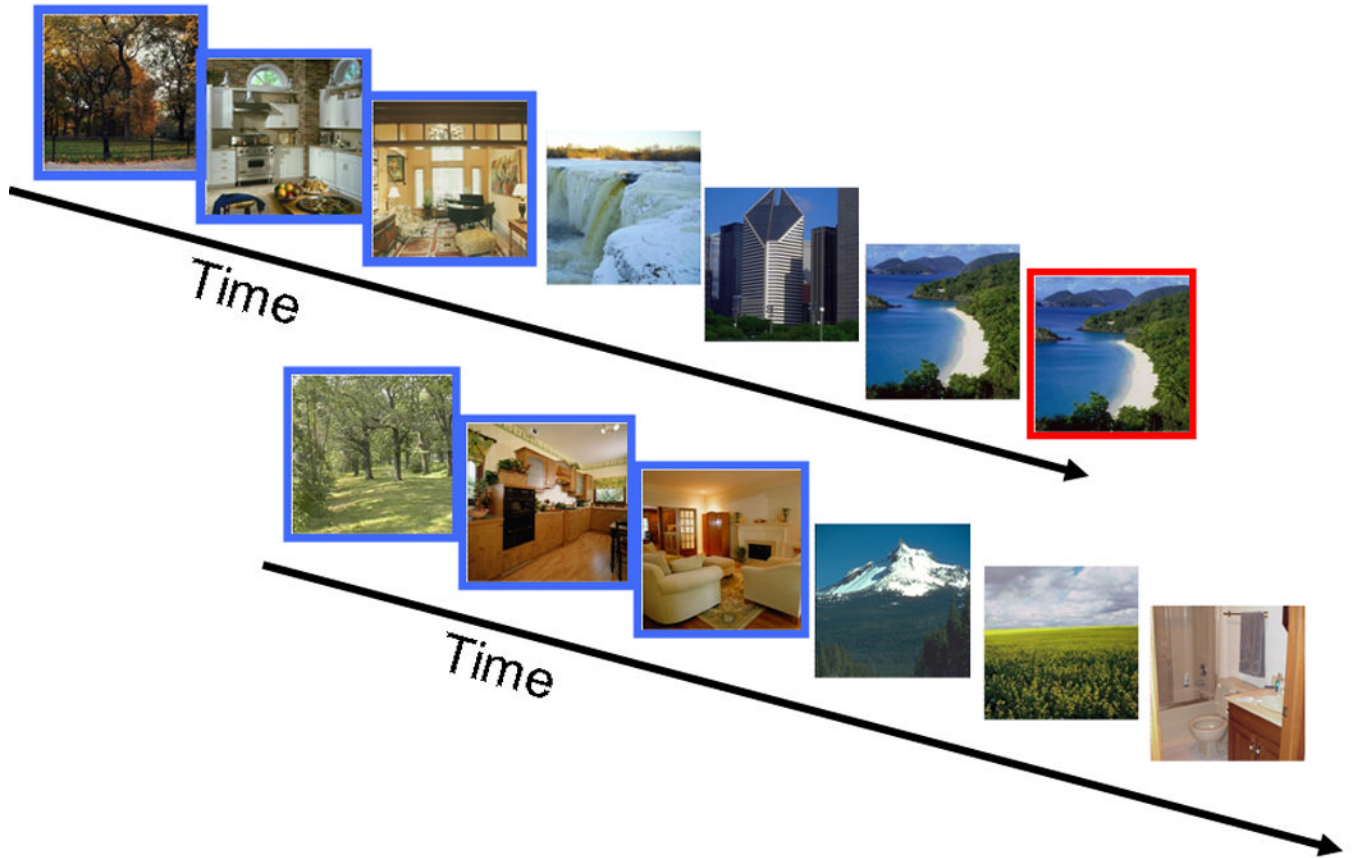Overall, 120 different images of each category were selected.

Figure 2.
A sample stream from Experiment 2. Images were presented in a predictable order, but this predictability was present only at the categorical level. Thus, the first time a triplet was presented one set of images was used, and the next time another set of images was used, etc. (images outlined in blue are an example). Repeated images were still physically identical (image outlined in red is an example).
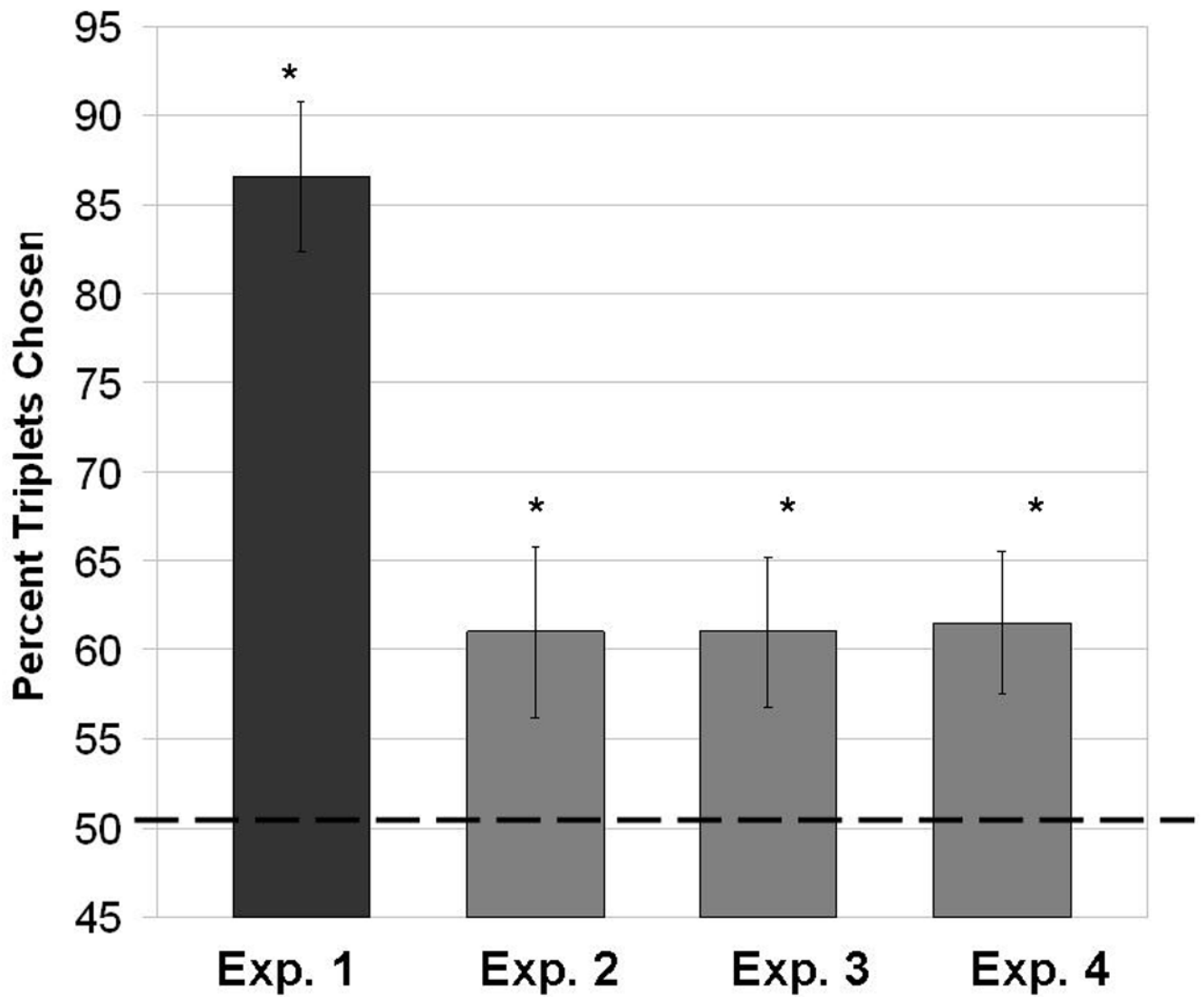
Figure 3.
Percentage of familiar triplets chosen in each experiment. Error bars represent one standard error of the mean. Chance level is 50%.