Spatial Constraints on Visual Statistical Learning of Multi-Element Scenes

Christopher M. Conway (cmconway@indiana.edu)
Department of Psychological & Brain Sciences, Indiana University, Bloomington, IN 47405 USA

Robert L. Goldstone (rgoldsto@indiana.edu)
Department of Psychological & Brain Sciences, Indiana University, Bloomington, IN 47405 USA

Morten H. Christiansen (mhc27@cornell.edu)
Department of Psychology, Cornell University, Ithaca, NY 14853 USA

Abstract

Visual statistical learning allows observers to extract high-level structure from visual scenes (Fiser & Aslin, 2001). Previous work has explored the types of statistical computations afforded but has not addressed to what extent learning results in unbound versus spatially bound representations of element co-occurrences. We explored these two possibilities using an unsupervised learning task with adult participants who observed complex multi-element scenes embedded with consistently paired elements. If learning is mediated by unconstrained associative learning mechanisms, then learning the element pairings may depend only on the co-occurrence of the elements in the scenes, without regard to their specific spatial arrangements. If learning is perceptually constrained, co-occurring elements ought to form perceptual units specific to their observed spatial arrangements. Results showed that participants learned the statistical structure of element co-occurrences in a spatial-specific manner, showing that visual statistical learning is perceptually constrained by spatial grouping principles.

Keywords: Visual Statistical Learning, Associative Learning, Perceptual Learning, Spatial Constraints.

Introduction

Structure abounds in the environment. The sounds, objects, and events that we perceive are not random in nature but rather are coherent and regular. Consider spoken language: phonemes, syllables, and words adhere to a semi-regular structure that can be defined in terms of statistical or probabilistic relationships. The same holds true for almost all aspects of our interaction with the world, whether it be speaking, listening to music, learning a tennis swing, or perceiving complex scenes.

How the mind, brain, and body encode and use structure that exists in time and space remains one of the deep mysteries of cognitive science. This issue has begun to be elucidated through the study of "implicit" or "statistical" learning[1] (Cleeremans, Destrebecqz, & Boyer, 1998; Conway & Christiansen, 2006; Reber, 1993; Perruchet & Pacton, 2006; Saffran, Aslin, & Newport, 1996). Statistical learning (SL) involves relatively automatic learning mechanisms that are used to extract regularities and patterns distributed across a set of exemplars in time and/or space, typically without conscious awareness of what regularities are being learned. SL has been demonstrated across a number of sense modalities and input domains, including speech-like stimuli (Saffran et al., 1996), visual scenes (Fiser & Aslin, 2001), and tactile patterns (Conway & Christiansen, 2005). Because SL appears to make contact with many aspects of perceptual and cognitive processing, understanding the underlying cognitive mechanisms, limitations, and constraints affecting SL is an important research goal.

Initial work in SL emphasized its unconstrained, associative nature (e.g., see Frensch, 1998; Olson & Chun, 2002, for discussion). That is, a common assumption has been that statistical relations can be learned between any two or more stimuli regardless of their perceptual characteristics or identity; under this view, there is no reason to believe that learning a pattern involving items A, B, and C should be any easier or harder than learning the relations among A, D, and E. However, recent research has shown that this kind of unconstrained, unselective associative learning process may not be the best characterization of SL (Bonatti, Peña, Nespor, & Mehler, 2005; Conway & Christiansen, 2005; Saffran, 2002; Turk-Browne, Junge, & Scholl, 2005). Instead, factors related to how the sensory and perceptual systems engage SL processes appear to provide important constraints on the learning of environmental structure.

In this paper we examine a largely unexplored constraint on visual statistical learning (VSL): the relative spatial arrangement of objects. If VSL operates via unconstrained associative learning mechanisms, we ought to expect that it is the co-occurrence of two objects that is important, not the relative spatial arrangement of those objects. However, another possibility is that VSL is akin to perceptual learning, in which two frequently co-occurring objects can form a new perceptual "unit" (Goldstone, 1998). Such unitization would be highly specific to not only the individual items but to their relative spatial arrangement as well. Before describing the empirical study in full, we first briefly review other work that points toward spatial constraints affecting visual processing.

[1] We consider implicit and statistical learning to refer to the same learning ability, which we hereafter refer to simply as statistical learning.

The Role of Space in Visual Processing

Intuitively, each sensory modality seems biased to handle particular aspects of environmental input. For instance, vision and audition appear to be most adept at processing spatial and temporal input, respectively (Kubovy, 1988). For instance, whereas the auditory system must compute the location of sounds through differences in intensity and time of arrival at each ear, the location of visual stimuli is directly mapped onto the retina and then projected topographically into cortical areas. In general, empirical work in perception and memory suggests that in visual cognition, the dimensions of space weigh most heavily, whereas for audition, the temporal dimension is most prominent (Friedes, 1974; Kubovy, 1988; Penney, 1989).

In the area of VSL, the ways in which time and space constrain learning have only recently begun to be explored. Although VSL can occur both with items displayed in a spatial layout (Fiser & Aslin, 2001, 2005), as well as with objects appearing in a temporal sequence (Conway & Christiansen, 2006; Fiser & Aslin, 2002; Turk-Browne et al., 2005), some evidence suggests that it is the former that occurs most naturally and efficiently. For instance, Gomez (1997) suggested that visual learning of artificial grammars proceeds better when the stimulus elements are presented simultaneously – that is, spatially arrayed – rather than sequentially, presumably because a simultaneous format permits better chunking of the stimulus elements. Likewise, Saffran (2002) found that participants learned predictive relationships well with a visual-simultaneous presentation, but did poorly in a visual-sequential condition. Finally, Conway and Christiansen (2007) further explored spatial constraints on VSL by creating structured patterns that contained statistical relations among temporally-distributed, spatially-distributed, or spatiotemporally-distributed elements. The results revealed that participants had difficulty acquiring the statistical patterns of the temporal and spatiotemporal stimuli, but easily learned the spatial patterns.

These data suggest that VSL occurs most easily for spatial layouts. However, a separate and hitherto unanswered question is whether VSL for spatially-distributed patterns necessarily leads to knowledge that is specific to the relative positions of the stimuli. For instance, suppose object A consistently is paired with object B, with A always occurring above B. After exposure to such pairs of items in a multi-element display, will participants learn that A and B co-occur, without regard to their arrangement, or that A and B co-occur in a specific spatial position (A above B)? If SL produces knowledge that is specific to the spatial arrangement of the co-occurring items, then this would suggest that VSL rather than being an unconstrained associative learning mechanism, may be more similar to perceptual learning processes which lead to highly specific forms of knowledge (e.g., Fahle & Poggio, 2002).

In the following two experiments, we build upon the work pioneered by Fiser and Aslin (2001; 2005), who investigated VSL for complex, multi-element displays. We used their paradigm to investigate to what extent VSL results in spatially bound versus unbound representations of object co-occurrences. Following the presentation of the experiments, we discuss the results in terms of how to best characterize the mechanisms underlying VSL

Experiment 1

Experiment 1 uses Fiser and Aslin's (2001) methodology in which participants are exposed to complex, multi-element scenes under passive, unsupervised viewing conditions. The scenes are composed of "base-pairs", which are two shapes that are consistently paired together in a particular spatial arrangement. Following presentation of the scenes, we tested participants' knowledge of the base-pairs in a forced-choice familiarity task. Unlike Fiser and Aslin (2001) who provided only one kind of test comparison (base-pairs vs. infrequent pairs), we also tested participants' familiarity of "switched" pairs. Switched pairs are two shapes of a base-pair that have had their spatial arrangements reversed. By including additional foil type, we can investigate to what extent participants' knowledge of the co-occurrence statistics is bound by the relative spatial arrangements in which the shapes had consistently been presented.

Method

Participants Seventeen undergraduate students at Indiana University participated and received course credit. All subjects were native speakers of English.

Stimuli Twelve arbitrary complex shapes, used by Fiser and Aslin (2001), were displayed in a 3 x 3 grid. The experiment consisted of two types of phases: exposure and test. During the exposure phases, the twelve shapes were organized into six base pairs. Each base pair consisted of two shapes that always occurred together in a specific spatial arrangement. As in Fiser and Aslin (2001), the six base pairs were organized into three orientations, two of each type: horizontal, vertical, and oblique. Scenes were created by randomly selecting 1 base pair of each orientation, and placing them on the 3 x 3 grid so that each base-pair touched at least one other base-pair. This method produces a total of 144 distinct scenes (see Figure 1 for examples). Given this method of scene creation, the probability of occurrence of a given individual shape is the same for all shapes; additionally, the joint probability of two shapes of a base-pair occurring in any given scene is 0.5.

Two other types of shape pairs were created to be used during the test phases: non-pairs and switched pairs. A non-pair was a pair of shapes that originated from two different base-pairs in the exposure phase. The probability of any given non-pair occurring together in the exposure phase was very low, less than 0.02. A switched pair was a base-pair that had the position of its two shapes reversed; that is, if a particular base-pair consisted of shape A always occurring above shape B, the switched pair contained shape B occurring above shape A. Thus, the joint probability of the two shapes of a switched pair occurring together (independent of their relative spatial arrangement) was 0.5, the same as the probability of a base-pair. However, the probability of the shapes of a switched pair occurring in that particular spatial arrangement was 0. Thus, in this way, the use of switched pairs allows us to pit spatial-independent statistics against spatial-specific statistics.

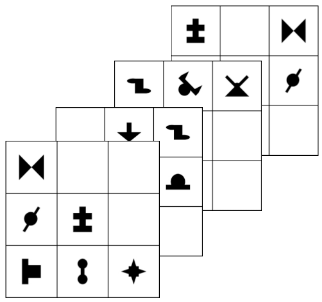

Figure 1: Illustration of scene presentation during exposure phases of Experiment 1. Scenes were shown 1 at a time.

Procedure Participants were instructed that they would view complex scenes one at a time. They were told to pay attention to what they saw because they would later be asked some questions. In the first exposure phase, participants saw each of the 144 scenes twice, presented in random order. Each scene was displayed for 2 s, with a 1 s pause inserted between scenes. Halfway through, participants were given a chance to take a voluntary rest break. The entire duration of this exposure phase was about 15 minutes. Note that at no point were participants told anything about the scenes having any kind of invariant structure.

Following the first exposure phase, participants were then given a series of temporal two-alternative forced-choice (2AFC) tests, in which two different pairs of shapes were shown on the grid, one at a time (see Figure 2). Participants were instructed to choose the pair that looked "most familiar" relative to the scenes they viewed in the exposure phase, by pressing the "1" or "2" keys. There were three types of comparisons: base-pair vs. non-pair; base-pair vs. switched pair; switched pair vs. non-pair[2]. For all cases, the two options had the same spatial arrangement (horizontal, vertical, or oblique) and absolute spatial position on the grid. There were 12 different 2AFC tests for each type of comparison, giving a total of 36 test trials. Each pair in a test was presented for 2 s with 1 s pause inserted in between. After the participant made a response, the next 2AFC test was initiated.

Following Test 1, participants engaged in a second exposure phase, which was identical in all respects to the first exposure phase except that each scene was viewed only once, in random order, for a total of 144 scene presentations. After the second exposure phase, participants were given Test 2, which consisted of the same 36 2AFC tests that they had received in Test 1.

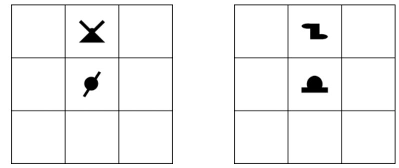

Figure 2: Illustration of sample 2AFC. Note that the two scenes are shown 1 at a time. The correct response in this case is the base-pair, on the right.

Results and Discussion

Test 1 and Test 2 results are reported for the three types of forced-choice comparisons, shown in Figure 3. In Test 1, only one comparison type, base-pair vs. switched pair, had performance significantly above 50% (M = 6) chance levels [M = 7.8; $t(16)$ = 4.3, p = .001]. Neither performance on base-pair vs. non-pair [M = 6.6; $t(16)$ = .98, p = .34] nor switch vs. non-pair comparisons [M = 4.9; $t(16)$ = -1.6, p = .12] reached significance. These results indicate that in Test 1, participants were able to distinguish a base-pair from its spatially-inverted arrangement, but could not distinguish a base-pair from a non-pair nor a switched pair from a non-pair. Thus, participants' knowledge following the first unsupervised learning phase was relatively fragile, limited only to the spatial-specific positions of base-pairs.

In contrast, Test 2 results indicate that both base-pair vs. switched pair [M = 10.1; $t(16)$ = 6.5, $p < .001$] and base-pair vs. non-pair [M = 10.2; $t(16)$ = 8.7, $p < .001$] comparisons were significantly greater than chance, whereas the switch vs. non-pair comparison was not [M = 6.7; $t(16)$ = .99, p = .34]. These results indicate that by Test 2, participants had learned the shape co-occurrence patterns and could not only distinguish a base-pair from its spatially-inverted foil, but could also reliably pick base-pairs over non-pairs.

In sum, the results from Experiment 1 strongly suggest that visual statistical learning is constrained such that co-occurrence patterns are learned in a spatially-specific manner. Incorporating three different types of test comparisons allowed us to closely examine the nature of knowledge gained from exposure to the structured scenes. On the switched pair vs. non-pair comparison, participants did not reliably choose one of the pairs over the other as being most familiar. If participants tended to choose the switched pair, this would have been strong evidence for a "spatial-independent" aspect of visual statistical learning. This result would have indicated that even though the shapes' spatial positions were inverted, the fact that the two shapes had consistently occurred together was enough for participants to learn their co-occurrence, independent of the actual relative positioning of the items. However, this was not what was found. The results instead showed that participants treated the switched pair no different than a non-pair, suggesting that the knowledge regarding the co-

[2] Note that for scoring purposes, for the switched vs. non-pair comparison, we arbitrarily chose the switched pair as being the correct response.

occurrence patterns was highly inflexible and constrained by the specific relative spatial arrangements of the objects.

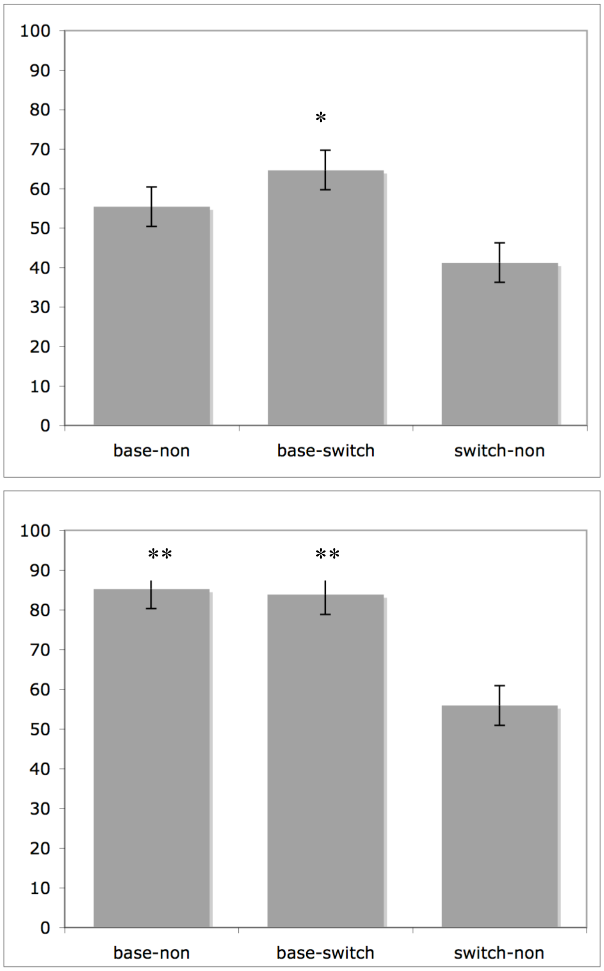

Figure 3: Experiment 1 performance (% correct) on each of the three comparison types for Test 1 (top) and Test 2 (bottom).

Experiment 2

Although the results of Experiment 1 are highly suggestive, one possible limitation is that participants received the identical test in both test phases. It is possible that the first test biased participants' performance on the second test. Thus, to eliminate this potential confound, we conducted Experiment 2 which incorporated only one test phase. Additionally, in order to encourage participants to better attend to the scenes in the exposure phase, we used a same-different task (Conway & Christiansen, 2005), rather than passive exposure.

Method

Participants An additional seventeen undergraduate students at Indiana University participated and received course credit. All subjects were native speakers of English.

Stimuli The shapes, scenes, and test pairs were identical to those used in Experiment 1.

Procedure The procedure was identical to Experiment 1 except in the following respects. Instead of having multiple exposure and test phases, there was only one exposure phase and one test phase. In the exposure phase, participants were told that they would see pairs of scenes, one scene at a time. For each pair of scenes, they were to decide whether they were the same or different, and press "S" or "D", respectively. The pairs of scenes consisted of the 144 multi-element scenes previously described. Each of the 144 scenes was paired with another scene, with half of all pairs being identical and half being different. The pairs that were different differed only in terms of 1 base-pair; and in almost all cases the absolute position of shapes on the 3 x 3 grid was the same. In this way, participants could not do the same-different task merely by noting that, for instance, the first scene had a shape in the upper left-hand location but the second scene did not. Doing this task successfully requires participants to pay attention to the actual identity of shapes in the scenes, in addition to their spatial positioning. Participants completed 144 same-different pairs (i.e., they viewed each of the 144 scenes two times). As before, each scene was shown for 2 s and there was a 1 s pause in between exposures.

Following the exposure phase, participants completed a familiarity test phase, which was identical to the tests used in Experiment 1.

Results and Discussion

The mean performance on the same-different task in the exposure phase was $M = 122.3$ out of a possible total of 144, with a range of (99, 138).

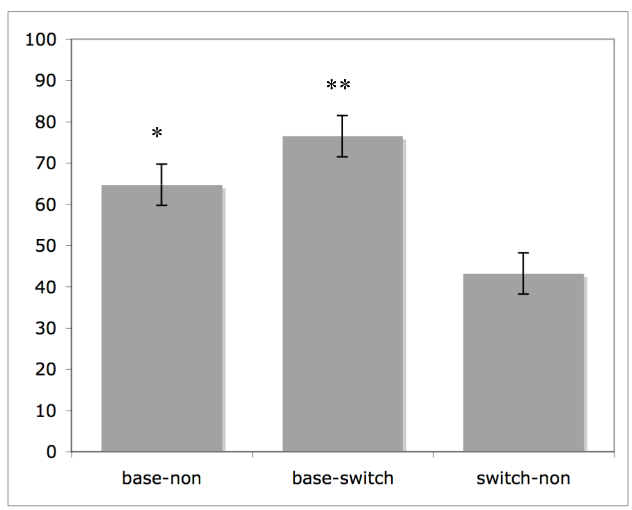

Figure 4: Experiment 2 test performance (% correct) on each of the three comparison types.

The results for the test phase are shown in Figure 4. As can bee seen, both the base-pair vs. switched pair [$M = 9.2$;

$t(16) = 6.4$, $p < .001$] and base-pair vs. non-pair [$M = 7.8$; $t(16) = 2.7$, $p < .02$] comparisons were significantly greater than chance, whereas the switch vs. non-pair comparison was not [$M = 5.2$; $t(16) = -1.5$, $p = .16$]. Performance for base-pair vs. switch pair was marginally greater than performance for base-pair vs. non-pair [$t(16) = 1.4$, $p = .09$].

The marginal difference indicates that on average, participants were slightly better at distinguishing base-pairs from switched pairs than they were at distinguishing base-pairs from non-pairs. That is, having positional information involved in the forced-choice task appears to aid performance, providing further support that VSL intimately relies on relative spatial position information.

In general, the pattern of results of Experiment 2 is essentially identical to that of Experiment 1 (Test 2). Experiment 2 thus serves to replicate the finding in Experiment 1 of spatial-specific learning mediating VSL.

General Discussion

In this paper, we attempted to investigate the nature of spatial constraints affecting VSL. Following exposure to structured multi-element scenes that contained pairs of invariantly arranged shapes, participants' knowledge of the co-occurrence pairs was tested. We created test comparisons that allowed us to determine to what extent learning was either independent of, or specific to, relative spatial position. The results were quite clear: participants' knowledge of the shape co-occurrence statistics was specific to the spatial arrangements in which they had occurred.

Note that this was not an inevitable result. From a purely unselective associative standpoint, it might have been expected that participants would treat the switched pair as being familiar because it was composed of elements that had co-occurred frequently. However, participants treated the switched pairs no differently than the non-pairs; in their eyes, the switched pairs were just as unfamiliar as two shapes that had never or rarely occurred together in the exposure phase.

That VSL is constrained by relative spatial position is consistent with other work showing the importance of the dimension of space to vision (Friedes, 1974; Penney, 1989). For example, results from experiments using the contextual-cueing paradigm (Chun, 2000) have shown that the visual system picks up invariant spatial relationships and uses this context to guide attention; furthermore, spatial features appear to play a more important cueing role than surface features such as color (Olson & Chun, 2002). The current data also complement our knowledge regarding the nature of constraints affecting statistical learning more generally. For instance, Turk-Browne et al. (2005) have illustrated attentional constraints on VSL. They presented participants with two streams of statistically-structured visual materials; only the stream to which participants were asked to attend resulted in learning. Bonatti et al. (2005) have shown that the presence of linguistic constraints affect statistical learning. In an auditory SL task, they found that participants preferentially learned statistics among consonants but not among vowels. Finally, Conway and Christiansen (2005, 2007) have revealed the presence of modality constraints affecting SL. They have shown that each sensory modality not only is particularly attuned to either spatial or temporal patterns, but also that each is differentially biased to pick up statistics at the beginning or ending of elements in a temporal stream.

Coupled with the results of Conway and Christiansen (2005, 2007), the present finding of spatial-specificity in VSL suggests that limitations in perceptual processing constrain what statistics are learned. There are at least two possible interpretations of these data. One possibility is that VSL is an associative learning mechanism in which particular perceptual, attentional, and cognitive constraints affect how and what types of statistics are learned. A second possibility, which we will entertain here, is that VSL may be more closely related to perceptual processing – specifically, perceptual learning – than to associative learning.

Although associative and perceptual learning are not necessarily mutually incompatible (e.g., see Hall, 1991), they do stress two different aspects of learning. Associative learning theories have to do with the linking of two or more stimuli or concepts such that the presence or excitement of one activates the other. Perceptual learning, on the other hand, emphasizes improvement in the perception or discrimination of stimuli following exposure. That is, the former theory has to do with cognitive "enrichment" whereas the latter has to do with perceptual "differentiation" and "specificity" (e.g., Gibson & Gibson, 1955; Pick, 1992; Postman, 1955).

Not surprisingly, many researchers have stressed the associative nature of SL (e.g., Fiser & Aslin, 2001; Frensch & Runger, 2003); at least superficially, learning the statistical relations between two co-occurring items appears to involve forming an association between them. However, our results show that VSL involves more than merely learning the association between two unbound elements; spatial position is also encoded. It is true that an associationist perspective could account for these results by assuming that associations are learned not just between two shapes but also between each shape and its spatial position. Even so, to be consistent with our data, the learned associations must involve *relative* spatial position, not just *absolute* position. One advantage of a perceptual learning account is that it predicts *a priori* that learning would be specific to the relative spatial position of the items (see Goldstone, 2000).

A perceptual learning account leads to an additional prediction. One of the primary mechanisms of perceptual learning is a "unitization" process in which two frequently co-occurring items become perceptually fused if a single image can be formed that integrates the two items (Goldstone, 1998). In the context of VSL, this would mean that the two individual shapes of a base-pair would, after sufficient exposure, be formed into a single functional unit. The prediction that follows is that VSL should lead to new units that are more easily perceived than combinations of items that did not co-occur frequently. We are currently testing this prediction. If an improvement is found in perception following statistical learning, this would be additional evidence supporting the idea that VSL may be akin to perceptual learning. Of course, as already stated, associationist theories can also be crafted to be consistent with such data, as long as they take into account the bidirectional effects between perception and learning, especially those involving relative spatial position.

To summarize, this paper investigated how spatial grouping principles constrain VSL. Consistent with previous work, VSL does not appear to involve spatially-insensitive associative learning processes, but instead is constrained by the relative spatial arrangement of the elements of a scene, limiting what kinds of patterns are readily learned. Based on this evidence, we suggest that it may be fruitful to explore possible links between VSL and perceptual learning to investigate the extent to which these two learning phenomena may ultimately be relying on common mechanisms.

Acknowledgments

We wish to thank Luis Hernandez, Jamie Lubov, and Maksim Sayenko for their help on this project. We also wish to thank József Fiser and Richard Aslin for providing us with the twelve shape stimuli used in these experiments. This work was supported in part by NIH DC00012, Department of Education, Institute of Education Sciences grant R305H050116, and NSF REC grant 0527920.

References

Bonatti, L.L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations. *Psychological Science, 16*, 451-459.

Chun, M.M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4*, 170-177.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences, 2*, 406-416.

Conway, C.M. & Christiansen, M.H. (2007). *Seeing and hearing in space and time: Effects of modality and presentation rate on implicit statistical learning*. Unpublished manuscript.

Conway, C.M. & Christiansen, M.H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science*, *17*, 905-912.

Conway, C.M. & Christiansen, M.H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. Jour*nal of Experimental Psychology, 31*, 24-39.

Fahle, M. & Poggio, T. (Eds.) (2002). *Perceptual learning*. Cambridge, MA: MIT Press.

Fiser, J. & Aslin, R.N. (2005). Encoding multielement scenes: Statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General*, 134, 521-537.

Fiser, J. & Aslin, R.N. (2002). Statistical learning of higher order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*, 458-467.

Fiser, J. & Aslin, R.N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*, 499-504.

Frensch, P.A. (1998). Once concept, multiple meanings: On how to define the concept of implicit learning. In M.A. Stadler & P.A. Frensch (Eds.), *The handbook of implicit learning* (pp. 47-104). London: Sage Publications.

Frensch, P. A., & Runger, D. (2003). Implicit learning. *Current Directions in Psychological Science, 12*, 13-18.

Freides, D. (1974). Human information processing and sensory modality: Cross-modal functions, information complexity, memory, and deficit. *Psychological Bulletin, 81*, 284-310.

Gibson, J.J. & Gibson, E.J. (1955). Perceptual learning: Differentiation or enrichment? *Psych Review, 62*, 32-41.

Goldstone, R.L. (2000). Unitization during category learning. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 86-112.

Goldstone, R.L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585-612.

Gomez, R.L. (1997). Transfer and complexity in artificial grammar learning. *Cognitive Psychology, 33*, 154-207.

Hall, G. (1991). *Perceptual and associative learning*. Oxford University Press.

Kubovy, M. (1988). Should we resist the seductiveness of the space:time::vision:audition analogy? *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 318-320.

Olson, I.R. & Chun, M.M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition, 9*, 273-302.

Penney, C.G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition, 17*, 398-422.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: Two approaches, one phenomenon. *Trends in Cognitive Sciences, 10*, 233-238.

Pick, H.L., Jr. (1992). Eleanor J. Gibson: Learning to perceive and perceiving to learn. *Developmental Psychology, 28*, 787-794.

Postman, L. (1955). Association theory and perceptual learning. *Psychological Review, 6*, 438-446.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford, England: Oxford University Press.

Saffran, J.R. (2002). Constraints on statistical language learning. *Journal of Memory and Language, 47*, 172-196.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926-1928.

Turk-Browne, N. B., Junge, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134*, 522-564.