Sensory Learning: Rapid Extraction of Meaning from Noise

Recent studies show that humans can rapidly learn to differentiate originally meaningless sounds into long-lasting memories, illustrating the flexibility of sensory processes and raising important questions about how sensory memories are formed.

Aaron R. Seitz

As we interact with the world, we are confronted with a rich array of sensory stimuli; some are known to us and are associated with meaningful labels, whereas others are ignored or considered as noise. For example, if you pause and listen you might hear people talking, birds singing, music playing, and so on. These are all labeled sounds that have meaning to us. You might also hear noise, for example from a ventilation system, static on the radio or from a car engine. While noise is often ignored, it can have diverse structures; for example, if you listen carefully to the noise made by your car-engine you can begin to pick out patterns. These patterns to an expert can be diagnostic of how the engine is working. Once meaningful structure has been found, the sounds are no longer just noise. Different engine sounds will fall under different labels, each associated with a sense of the operation of the engine. This raises an important question: how do sensory stimuli become differentiated and meaningful to us? If you think about it, this question is rather profound: it seems that most, if not all, stimuli that we experience go from a state of meaninglessness to a state of meaningfulness. Take, for example, the case of language. The sounds, or written characters, of an unknown language may have no clear framework until one learns how to differentiate and label them. Clearly, the process by which noise becomes meaningful is foundational to how we interact with the world.

Agus *et al.* [1] recently addressed this question directly by studying how people form memories of auditory noise patterns. They did this by creating a large set of unique segments of auditory noise (see Figure 1A for a couple of examples) and then asking observers to discriminate between regular noise patterns (one second of noise) *versus* noise patterns that contain an internal repeat (half-second noise segment played twice in succession). To a naïve observer all the stimuli sound like hisses and are very difficult to differentiate. When the sounds are played multiple times, however, observers become much better at discriminating the stimuli. In some cases, observers went from chance performance to perfect performance after just four trials of exposure. Notably, this learning is specific to the particular sounds that were presented multiple times and the improvements in performance do not transfer to novel sounds. Thus, the authors concluded that the performance improvements are signs that observers developed memories of specific sounds. These memories were long lasting (for weeks or longer) and were robust to the sounds being sped up or played backwards. Furthermore, observers were able to explicitly identify which sounds had been repeated multiple times. Thus, long-lasting memories of originally meaningless sounds can develop in a handful of trials, without any direct instructions to the observers that the stimuli should be memorized, or any feedback regarding task performance. In essence, it seems that people are predisposed to pick out and make sense of the consistently presented noise patterns and that this can be accomplished extremely rapidly.

These results fit well with other studies of sensory learning and memory formation. For example, a single exposure to an originally meaningless visual pattern (for example, the dog in Figure 1B) can result in a long-term memory of that stimulus [2]. Likewise, single-shot learning for auditory stimuli has been found in the context of learning language [3]. The new results also relate to studies of statistical learning, which show that both children and adults pick up statistical regularities of meaningless auditory or visual stimulus sets after a small number of presentations and in the absence of any task other than attending to the stimuli [4,5]. Similarly, research on contextual cueing has shown that, in tasks where observers are looking for a visual target among distractors, with just a handful of presentations of a stimulus array, the observers learn the configurations of distractors (these can be considered as a type of visual noise) and use this information to better their performance in locating the search target [6]. Both statistical learning and contextual cueing have been shown to be long lasting and implicit [7,8]. Thus, the results of Agus *et al.* [1] fit in well with a burgeoning literature demonstrating that regularities in the sensory

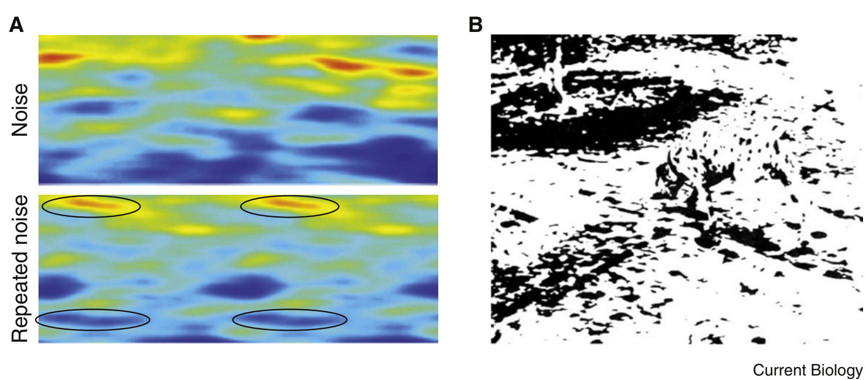

Figure 1. Examples of rapidly learned stimuli that initially are sensed as noise.

(A) Example of stimuli from Agus *et al.* [1]. Top shows noise stimulus and bottom repeated noise stimulus. Circled regions indicate modulations in some of the frequency bands that could be diagnostic to discriminating the sound when played forward or backwards (note, other frequency bands could also be used and the most prominent features likely differ across observers). (B) Naïve observers often fail to identify the dog, but once the dog has been identified it is ever after always seen (photograph by Ronald C. James).

environment are often quickly picked up and utilized to improve performance on tasks involving those stimuli.

While it is clear that the literature as a whole demonstrates that this type of fast automatic learning of sensory signals is the norm rather than the exception, it seems unlikely that all regularly presented sensory stimuli are automatically learned. We experience an astronomical number of sensory stimuli in our lifetimes and the number of stimuli that we experience multiple times is extremely large. We simply do not have sufficient memory capacity to encode every repeated stimulus into long-term memory [9]. The results of numerous studies indicate that sensory learning is not completely automatic, but instead requires attention and/or reinforcement (for reviews see [10,11]). Attention plays a key role in determining which stimuli are learned and which not in studies of statistical learning [12] and perceptual learning [13]. Along these lines, car mechanics, who study engines and their sounds, are better at ascribing meaning to engine noise than the typical individual. Notably, in Agus *et al.* [1], observers were required to attend to the sounds and repeatedly respond to these stimuli. Thus, it seems that some combination of the repeated presentation of the same sensory stimuli (to repeatedly stimulate a population of neurons), consistent attention to those stimuli (to enhance the responses of those neurons), and repeated discrimination of those sensory stimuli (this releases reinforcement signals [10]) are key to promote rapid memory formation.

Agus *et al.* [1] demonstrate how fast auditory learning can occur, but a number of questions remain unanswered. For example, what exactly did the observers actually learn? The authors suggest that it is unlikely that the observers encoded an exact representation of their auditory stimuli, instead it seems likely that the observers learned to discover patterns in the noise. For example, if the observers learned to identify a periodic pattern in one or a few frequency bands (Figure 1A), then this could be equally diagnostic of the sound when played forward or backward (as was observed by the authors). Another question regards whether the results of short-term memory (minutes) and long-term memory (days) are due to the same processes. Agus *et al.* [1] found performance benefits after a handful of trials; however, a full 50 trials per stimulus were completed before the long-term memory was tested. It will be relevant to examine whether these additional trials were necessary to trigger the development of the long-term memories. Finally, how does the brain encode a previously unknown stimulus from such limited exposure? Does this involve plasticity of sensory structures or of specialized memory structures, or both? Also, how does repeated stimulation of the given neural population interact with attentional and reinforcement factors to produce such plasticity (for a review of possible mechanisms see [14])? Answering these, and related, questions will remain a challenge for future studies.

References

1. Agus, T.R., Thorpe, S.T., and Pressnitzer, D. (2010). Rapid formation of robust auditory memories: insights from noise. Neuron *66*, 610–618.
2. Rubin, N., Nakayama, K., and Shapley, R. (1997). Abrupt learning and retinal size specificity in illusory-contour perception. Curr. Biol. *7*, 461–467.
3. Carey, S., and Bartlett, E. (1978). Acquiring a single new word. Proc. Stanf. Child Lang. Conf. *15*, 17–29.
4. Saffran, J.R., Johnson, E.K., Aslin, R.N., and Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. Cognition *70*, 27–52.
5. Fiser, J., and Aslin, R.N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. Psychol. Sci. *12*, 499–504.
6. Chun, M.M. (2000). Contextual cueing of visual attention. Trends Cogn. Sci. *4*, 170–178.
7. Jiang, Y., Song, J.H., and Rigas, A. (2005). High-capacity spatial contextual memory. Psychon. Bull. Rev. *12*, 524–529.
8. Kim, R., Seitz, A., Feenstra, H., and Shams, L. (2009). Testing assumptions of statistical learning: is it long-term and implicit? Neurosci. Lett. *461*, 145–149.
9. Von Neumann, J. (1958). The Computer and the Brain. Mrs. Hepsa Ely Silliman memorial lectures. (New Haven: Yale University Press.)
10. Seitz, A.R., and Watanabe, T. (2009). The phenomenon of task-irrelevant perceptual learning. Vision Res. *49*, 2604–2610.
11. Ahissar, M., and Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. Trends Cogn. Sci. *8*, 457–464.
12. Turk-Browne, N.B., Junge, J., and Scholl, B.J. (2005). The automaticity of visual statistical learning. J. Exp. Psychol. Gen. *134*, 552–564.
13. Ahissar, M., and Hochstein, S. (1993). Attentional control of early perceptual learning. Proc. Natl. Acad. Sci. USA *90*, 5718–5722.
14. Seitz, A.R., and Dinse, H.R. (2007). A common framework for perceptual learning. Curr. Opin. Neurobiol. *17*, 148–153.

Department of Psychology, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA.
E-mail: aseitz@ucr.edu

DOI: 10.1016/j.cub.2010.06.017

Visual Consciousness: The Binocular Rivalry Explosion

A new behavioural technique solves a long-standing puzzle of binocular suppression, demonstrating that adapting reciprocal inhibition governs visual sensitivity and raising key questions about visual awareness.

Daniel H. Baker

Usually our two eyes receive similar views of the world, and the brain is able to combine, or ‘fuse’, these into a single, stable percept. But when the eyes report very different images, the brain is faced with a paradox: which image is correct? Like a canny investor, the brain chooses to hedge its bets. Instead of choosing just one image, or combining the two, we experience alternations between them, typically every few seconds. This phenomenon (illustrated in Figure 1A) is known as binocular rivalry, because it is as though the neural representations of the two images are competing against each other in a continuous ‘tug of war’. At a given point in time, one image is dominant (perceived) and the other is suppressed entirely from awareness, yet both remain present at the retina.

Aside from some early investigations, rivalry alternations were largely treated as a curiosity until, in 1965, W.J.M. Levelt’s doctoral research [1] brought a quantitative rigour to study of the phenomenon.