Eriksen, J. L., & Janus, C. G. (2007). Plaques, tangles, and memory loss in mouse models of neurodegeneration. *Behavior Genetics, 37*, 79–100.

Galton, F. (1869). *Hereditary genius: An inquiry into its laws and consequences.* London: Macmillan.

Shors, T. J., & Matzel, L. D. (1997). Long-term potentiation: What's learning got to do with it? *The Behavioral and Brain Sciences, 20*, 597–655.

Silva, A. J., Stevens, C. F., Tonegawa, S., & Wang, Y. (1992). Deficient hippocampal long-term potentiation in alpha-calcium- calmodulin kinase II mutant mice. *Science, 257*, 201–206.

Tang, Y. P., Shimizu, E., Dube, G. R., Rampon, C., Kerchner, G. A., Zhuo, M., Liu, G., & Tsien, J. Z. (1999). Genetic enhancement of learning and memory in mice. *Nature, 401*, 63–69.

Learning Networks

► Collective Learning

Learning Nonadjacent Dependencies

Pierre Perruchet[1], Bénédicte Poulin-Charronnat[1], Sébastien Pacton[2]
[1]LEAD/CNRS, Université de Bourgogne, Dijon, France
[2]LPNCog/CNRS, Université Paris-Descartes, Paris, France

Synonyms

Distant associations; Long-distance dependencies; Remote contingencies

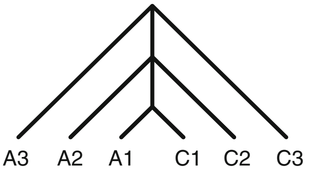
Definition

Nonadjacent dependencies are present whenever a relation exists between two events, A and C, irrespective of the intervening events. This structure is often referred to as A*X*C, where *X* stands for a variable event that is statistically independent from both A and C. An example of nonadjacent dependencies is the relationship between auxiliaries and inflectional morphemes, as in "*is* writ*ing*" in English, which occurs irrespective of the verb stem. The mastery of this kind of structure in the language area has been endowed with major theoretical implications in a Chomskyan perspective, and as a consequence, most research has been carried out in the language domain. However, it is worth pointing out that capturing the relationships between distant events is essential in many other situations. As claimed by Turk-Browne et al. (2005), "people are constantly bombarded with noise in space and time that needs to be segregated in order to extract a coherent representation of the world, and people rarely encounter a sequence of relevant stimuli without any interruptions" (p. 562).

Theoretical Background

By and large, most experimental studies on learning have focused on the human abilities to detect and exploit the relations between *adjacent* elements. For instance, in the traditional literature on associative learning, such as the domain of animal conditioning, or studies on paired-associate learning in humans, the to-be-associated items are displayed in close temporal or spatial proximity. Although in less obvious ways, most complex learning settings also rely on structures defined by adjacent relations. For instance, most of the finite-state grammars that are commonly used in implicit learning studies govern the transitional probability between contiguous elements. It is essential to note that event adjacency is not a fortuitous, accidental property of the materials, something that could be changed without any theoretical consequence. Indeed, from the "theory of contiguity" of Guthrie to the accounts of complex learning relying on the notion of chunks, usually defined as the grouping of a small number of contiguous events, the main theories of learning turn out to be devised for situations in which the relevant events are adjacent. Interestingly, recent studies on language acquisition have shown that learning adjacent relations are far more relevant than has been claimed in the past. For instance, a number of studies have demonstrated that the formation of the lexicon partly relies on statistical relations between adjacent syllables. More surprising, it has been shown that highly local context (the words immediately surrounding a target word) provided a considerable amount of information about the syntactic category of the target word.

However, it is unquestionable that linguistic structures also embed *nonadjacent* dependencies, which are successfully exploited by the learners. Such relations are found at different levels, from the subsyllabic level



Learning Nonadjacent Dependencies. Fig. 1 Schematic diagram of a grammar generating center-embedded sentences (the element *X* in the higher A*X*C structure is another A*X*C structure)

(e.g., the short vs. long pronunciations of vowels according to the presence of a "silent e" ending, irrespective of the intermediary consonant, as in CAP–CAPE, CAR–CARE) to morphosyntactic relationships such as the relation between auxiliaries and inflectional morphemes evoked in the definition above, and complex hierarchical structures. The current literature essentially focuses on situations in which strings are embedded within other strings, thus creating so-called center-embedded structures (Fig. 1). In the utterance "the rat the cat ate stole the cheese," for instance, one relative clause ("the cat ate") is nested within the sentence ("the rat stole the cheese").

The main theoretical issue is: Can this form of acquisition be encompassed within a general theory of associative learning, which would be able to account for both adjacent and nonadjacent dependencies? Undoubtedly, the prevalent models of learning, whether based on the associative principle of contiguity or the formation of chunks, are, as such, inadequate. Moreover, the issue does not seem easy to settle with only minor adjustments. Indeed, irrespective of their differences, most models of associative learning rely on a process of frequency-based selection among a set of possible associations. Only repeated or consistent pairs of events would be strengthened, the others being let aside or suppressed through decay or interference. This selection is a plausible mechanism whenever the number of possible associations remains limited, which is arguably the case when only adjacent events are considered. However, removing the adjacency constraint raises the well-known issue of combinatorial explosion: The number of possible associations becomes unmanageable, making the selection of the relevant associations unrealistic. This kind of theoretical considerations led Chomsky to contend that mastering the recursive structure underlying center-embedded sentences is an innate, language-specific ability, without any relationships with basic associative mechanisms.

A number of empirical observations run against a so extreme standpoint. For instance, humans are also able to master nonadjacent dependencies in other natural domains of high-level knowledge comprising several organizational levels, such as Western music, which uses variations and ornaments. Two structurally important tones, for instance, are often separated by other, less important tones (the ornaments). If the nonadjacent dependency between the two structurally important tones was not captured by the listener, the musical structure would not be perceived. In addition, a number of studies have explored the question of whether animals and humans are able to learn nonadjacent dependencies in artificial experimental settings. While the results on animals remain controversial, the results on humans show consistently that learning arbitrary nonadjacent dependencies is possible, and so not only with linguistic stimuli (Gomez 2002) but also with tones (Creel et al. 2004) and visual patterns (Turk-Browne et al. 2005). However, this form of learning would be dependent on far more restrictive conditions than those required for learning the relations between contiguous events. A number of conditions have been identified, although none of them can be construed as a necessary prerequisite. A nonexhaustive list of potentially cumulative facilitatory conditions includes the following: (1) The fact that the A*X*C structure may be processed as a whole, that is, can be easily isolated from the surrounding sequence of events. (2) The high level of variability of the *X* event. (3) The high level of similarity between A and C. Similarity can be assessed on an acoustic dimension. Using musical tone sequences, Creel et al. (2004) showed that nonadjacent dependencies were not acquired when all elements differed equally one another, whereas learning was successful when A and C were similar in pitch or timbre, and different from *X*. Others have shown that no learning was obtained without some degree of phonological similarity between A and C syllables. (4) The membership of A and C to the same category, itself differing from the category of *X*. For instance, some studies have failed to observe learning with nonadjacent syllables (i.e., A, *X*, and C were syllables), whereas learning occurred when A and C were consonants and *X* was a vowel and, conversely, when A and C were vowels and *X* was a consonant. (5) The occurrence of an earlier training phase during which the to-be-associated pairs have been studied in adjacent conditions. Introducing structural complexity progressively during learning would meet the general learning principle known as the "starting small" effect.

While these results make the initial Chomsky's view no longer defensible, they do not provide evidence that learning nonadjacent dependencies relies on the same associative/statistical mechanisms that are usually considered as responsible for the processing of adjacent events. Some authors contend that associative or statistical learning mechanisms are insufficient for extracting structural information, and that at least two mechanisms are required in order to learn from a complex environment: (1) an associative/statistical mechanism, essentially tracking adjacent associations, and (2) an additional rule-following mechanism dealing with more complex structures (Endress and Bonatti 2007). The issue meets here the great debate about the architecture of cognition, opposing those who argue that statistical learning mechanisms are sufficient for language acquisition and other high-level abilities, to those who advocate for the need of additional, rule-following mechanisms.

Important Scientific Research and Open Questions

A strategy that has been heavily exploited in this research context consists in exploring whether artificial neural networks are able to account for the mastery of complex structures. Given that connectionist networks embed only statistical mechanisms, without any hardwired rules, the fact that their success would provide a strong argument against a dual-mechanism framework is commonly acknowledged. Unfortunately, whether connectionist networks are actually successful in this endeavor is far less consensual. The last episode to date in the domain covered here is the Laakso and Calvo's (2011) demonstration that a standard connectionist network was able to mimic the results collected in a complex experimental setting involving long-distant dependencies, while other authors (Endress and Bonatti 2007) had previously inferred from their own failure to simulate the same data that this situation required a dual mechanism. But considering the past controversies surrounding this kind of demonstration, one may guess that the story is not over, and that connectionist modelers will be faced with further challenges in the near future.

The (at least provisional) success of connectionist approaches to complex learning supplies a feasibility proof that rule-based mechanisms are not necessary to learn nonadjacent dependencies, but it does not provide, as such, a psychologically relevant theory of learning. In their tentative elaboration of an integrative theory, several authors have put forth the concept of attention as the possible cornerstone of a unified framework. Emphasizing the role of attention in associative learning is far from new, but empirical data suggesting that learning does not occur without a minimal level of attention have accumulated in recent years. Moreover, a few authors have suggested that associative learning is an automatic process that links together all of the components that are present in the attentional focus at a given point. In other words, the joint attention given to a pair of events would be a *necessary*, but also a *sufficient* condition for the emergence of associative learning and memory. Pacton and Perruchet (2008) have emphasized the potential of this proposal to serve as a foundational principle for an integrative, unified theory. Indeed, this account is fully compatible with the conventional focus on the condition of contiguity, because the mental content composing the attentional focus at a given moment has a high chance of representing events that are close on spatial and/or temporal dimensions in the environment. However, the attentional content may also encompass events that are not adjacent in the environment, provided that some specific reasons lead to pay joint attention to those events. Thus, an attention-based view accounts for both the easy formation of associations between contiguous events and the more limited ability to build associations between nonadjacent events (the joint attentional processing of those events requires some special conditions, such as those described in the empirical studies outlined above). Although an attention-based framework may be to date the best alternative to a dual view positing the need for rule-based mechanisms, the debate remains open. An attention-based framework trades the concept of contiguity in the environment for the concept of contiguity at the level of attentional, internal representations. This amounts to trade a directly measurable variable for another that may only be inferred from overt behavior, hence making the model harder to falsify. Further studies are clearly needed to assert whether both adjacent and nonadjacent dependencies can be accounted for by general, all-purpose associative mechanisms.

Cross-References

- ▶ Associative Learning
- ▶ Attention and Implicit Learning
- ▶ Behaviorism and Behaviorist Learning Theories
- ▶ Conditioning
- ▶ Connectionist Theories of Learning
- ▶ Human Contingency Learning
- ▶ Language and Learning
- ▶ Learning by Chunking
- ▶ Paired-Associate Learning
- ▶ Statistical Learning in Perception

References

Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: Statistical learning of non-adjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 1119–1130.

Endress, A., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition, 105*, 247–299.

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13*, 431–436.

Laakso, A., & Calvo, P. (2011). How many mechanisms are needed to analyze speech?: A connectionist simulation of structural rule learning in artificial language acquisition. *Cognitive Science* (in press).

Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 80–96.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134*, 552–564.

Learning Not to Fear

Clas Linnman, Mohamed A. Zeidan,
Mohammed R. Milad
Department of Psychiatry, Harvard Medical School &
Massachusetts General Hospital, Charlestown,
MA, USA

Synonyms

Fear extinction

Definition

Fear extinction is the process in which a previously conditioned fear response to a cue is reduced when the cue is presented in the absence of a previously paired aversive stimulus.

Theoretical Background

Humans, rodents, even nematode worms can learn to fear and avoid threatening situations. Fear learning is an essential skill that helps us adapt to a dangerous world, but learning *not* to fear is almost equally important, as this allows us to update our predictions in the light of new information. The ability to overcome a learned fear is critical as it allows us to go on with our daily lives even after suffering traumatic events. Imagine yourself in a bad car accident when driving through an intersection. This single event may condition you to fear intersections or cues that remind you of the accident, such as the song that was playing on the radio just before the accident, or the smell of burning rubber. Each time you pass through an intersection, hear that song, or smell burning rubber, your palms start to sweat, your heart rate increases, and you will find yourself taking shorter breaths – all physiological signs of an elevated fear response. Learning to fear these cues is called *fear conditioning*. For most of us, a few trips through intersections with no accidents will gradually remind us that the accident was an isolated incident and that intersections are predominantly safe. Our heart rates come back down to normal, and we breathe easy, and the subsequent learning not to fear the intersection is called *fear extinction* – the active process by which we learn not to fear through repetitive presentations of a previously feared cue (the intersection, the song, the smell) in absence of the once-paired aversive stimuli (the accident).

Importantly, fear extinction does not erase the initial fear memory. Learning not to fear does not occur by forgetting a traumatic event, but instead happens by creating new safety memories that compete with, and eventually overpower, the fear memory. This process forms the basis behind existing treatments for anxiety disorders such as posttraumatic stress disorder (PTSD), generalized anxiety disorder and many others. The treatment method is called Exposure Therapy and falls under the treatment umbrella of Cognitive Behavioral Therapy (CBT). During exposure therapy, a patient reads with a psychologist and discusses the