



A common probabilistic framework for perceptual and statistical learning

József Fiser^{1,2} and Gábor Lengyel^{1,2}

System-level learning of sensory information is traditionally divided into two domains: perceptual learning that focuses on acquiring knowledge suitable for fine discrimination between similar sensory inputs, and statistical learning that explores the mechanisms that develop complex representations of unfamiliar sensory experiences. The two domains have been typically treated in complete separation both in terms of the underlying computational mechanisms and the brain areas and processes implementing those computations. However, a number of recent findings in both domains call in question this strict separation. We interpret classical and more recent results in the general framework of probabilistic computation, provide a unifying view of how various aspects of the two domains are interlinked, and suggest how the probabilistic approach can also alleviate the problem of dealing with widely different types of neural correlates of learning. Finally, we outline several directions along which our proposed approach fosters new types of experiments that can promote investigations of natural learning in humans and other species.

Addresses

¹ Department of Cognitive Science, Central European University, Nador utca 9, 1051 Budapest, Hungary

² Center for Cognitive Computation, Central European University, Oktober 6 utca 7, 1051 Budapest, Hungary

Corresponding author: Fiser, József (fiserj@ceu.edu)

Current Opinion in Neurobiology 2019, **58**:218–228

This review comes from a themed issue on **Computational Neuroscience**

Edited by **Máté Lengyel** and **Brent Doiron**

<https://doi.org/10.1016/j.conb.2019.09.007>
<https://doi.org/10.1016/j.conb.2019.09.007>

0959-4388/© 2019 Elsevier Ltd. All rights reserved.

Introduction

System-level characterizations of human learning of sensory information have been keeping a systematic distinction between low-level “perceptual learning” and high-level representational or “statistical learning”. According to this distinction, these two kinds of learning differ not only in their

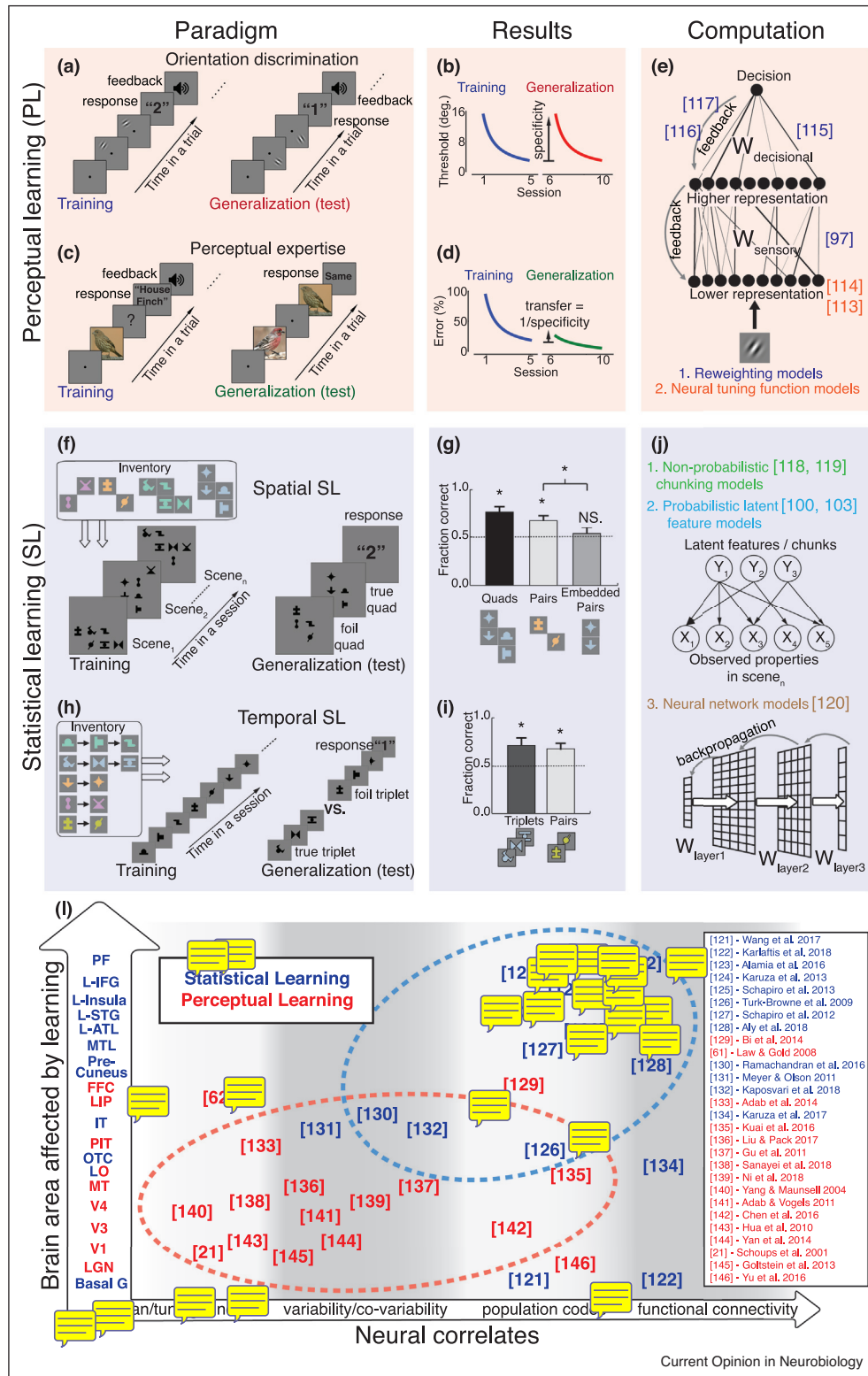
testing paradigms and stimuli, but also by their main characteristics, the presumed underlying mechanisms of learning, and the locus where the changes due to learning occur in the brain (Figure 1).

Perceptual learning (PL) is classically defined as improvement in simple sensory tasks with extensive practice [1,2] (Figure 1a–e). Early visual studies established that perceptual learning led to enhanced performance, among other tasks, in contrast [3,4] and motion detection [5], orientation [6] and texture discrimination [7,8], hyperacuity [9] and stereoscopic vision [10]. Extensive practice typically amounts to 5–14 days of repetitive exposure over 1–2 h [11]. Sleeping across the days is necessary for PL since it significantly alters the amount of learning due to consolidation [12,13], and the changes remain in effect for days, months, even years [14]. While in a few studies, feedback on the correctness of the observer’s response during trials was not provided [15], typically, there is feedback, and it is crucial for improving [16] or even permitting learning [17]. The amount of learning is usually measured in improvements of a threshold indicating a change in sensitivity [2]. There are several hallmarks of perceptual learning that cast this type of learning as a low-level phenomenon. The first is the specificity of learning: the acquired improvement in performance does not hold when conditions are altered (Figure 1b–d). Examples of such alterations are the stimulus being presented at a different location [18,19], orientation [20], spatial frequency [6], paired with different background [20] or seen through a different eye [18]. Especially eye-specificity has been used to argue for a low-level origin of PL: since merging of monocular representations happens in V1, eye-specific differences require learning also to occur in the primary visual cortex [21]. PL has also been associated with tasks using higher-level stimuli in a number of expertise-learning studies that depended on fine discriminations of sensory input [22,23] (Figure 1c).

Statistical learning (SL) refers to the type of representational learning that is purely observational without any task or feedback, which automatically and implicitly represents repeatedly appearing spatial and temporal patterns in the sensory input [24,25] (Figure 1f–k). Originally introduced in the domain of language learning for solving the problem of word segmentation [26], statistical learning has been later predominantly investigated in the domain of vision [27,28]³. Initial results established that adults and infant alike demonstrate spatial and temporal

³ A large body of statistical learning studies not discussed in the present paper is focused on the domain of language development (see [25,29,30]).

Figure 1



Classical perceptual and statistical learning and their neural correlates. The paradigms (left), typical behavioral results (middle), and computational frameworks (right) of perceptual learning (a–e, pink background) and statistical learning (f–j, blue background). Bracketed numbers in panels e, j, l indicate references. (a, b) Classical orientation discrimination task with the corresponding performance improvement in the trained condition (drop in blue curve) and specificity (i.e. a lack of transfer of performance to a different condition, initial jump in red curve). (c, d)

statistical learning based on both joint and conditional probabilities as well as higher-order embedded structures of previously unknown inputs [31–34] (Figure 1g,i). These results were extended to various modalities (visual, auditory, tactile) [35–38], to different stimulus complexities [39,40], and to other animals species [41–45]. This ubiquitousness fueled the proposal that statistical learning is a domain-general process that might serve as the fundamental learning method for acquiring internal representations of the environment [38,46] even though some auxiliary domain-specific constraints might exist [47]. SL is automatic and persists for a long time [48], sleep does not improve it [49,50] and while attention can influence SL [51], it is not required for successful learning [52]. Statistical learning has also been linked to or contrasted with higher level abstract concept learning [53] and rule learning [54,55]³

Accumulating recent evidence suggests a vanishing distinction between PL and SL

While earlier studies have already found evidence indicating an overlap between the neural substrates and computational features of PL and SL [15,56,57*], more recent reports greatly accelerated this convergence due to the increasing similarity in stimulus complexity and task specificity between experiments conducted in the two domains (Figure 2).

In the domain of PL, it has been firmly established by now that PL induces changes not only in V1 but in a large set of brain regions and influencing post- sensory processes as well [58,59]. PL is task- and context-specific [60], it appears to share common neural mechanisms with decision making processes in monkeys [61,62] and humans [63], and both exogenous and endogenous spatial attention affect it [64,65]. Even pure mental imagery without any sensory input can induce PL [66]. Using a “double-training” learning paradigm, various studies reported enhanced or complete transfer of the learned ability to a new condition [67,68] not only across different locations but across different physical properties that share “conceptual level” similarities [69**]. Transfer was enhanced when trials from multiple versions of the same task were delivered in a fixed order [70], transfer depended on the precision of the transfer test, not only of

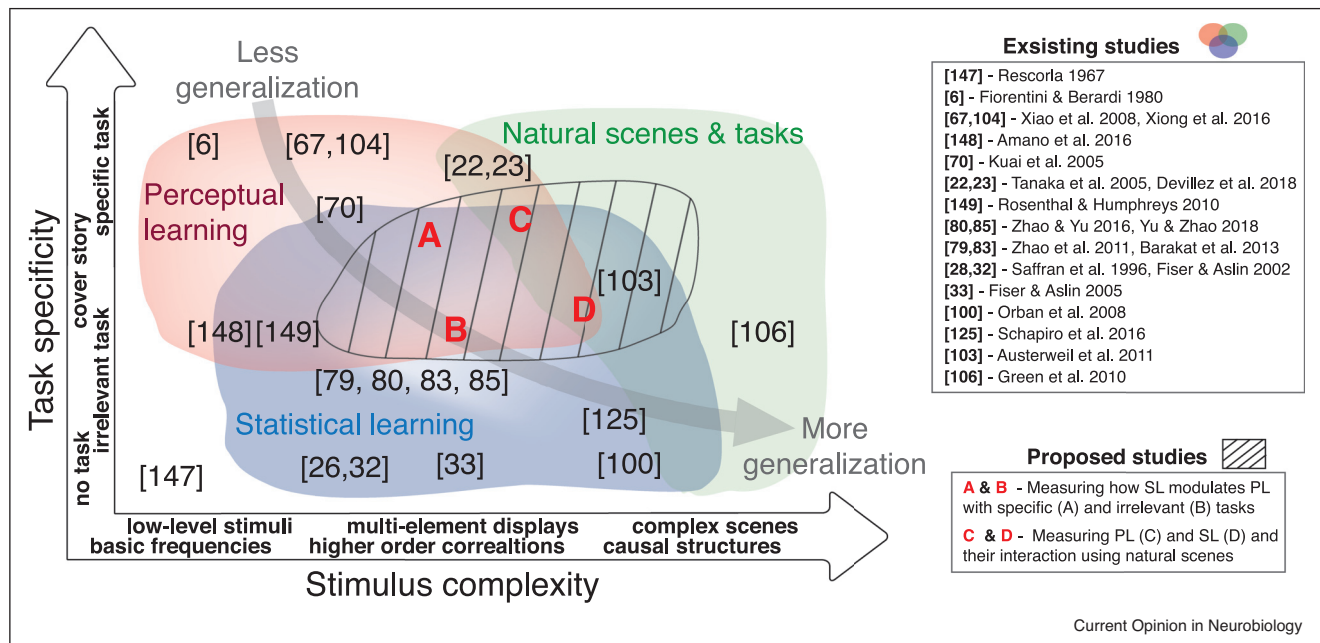
the original training task [71], and in general, the relationship between the type of the training and test tasks determined the success of generalization [72–74]. In addition, increasing stimulus complexity also facilitates generalization [75]. Higher-level generalization in PL has been investigated with training to play video games and learning was manifested not by simply having better attention, but by improved ability to generate templates for task learning [76,77]. Such structure-learning revealed by a faster learning rate could occur independently from the traditional immediate transfer in performance during PL [78*].

In the domain of SL, there has also been a steady progress of expanding and concretizing the areas and the extent to which SL influences or changes perceptual processes (Figure 2). SL interferes with the process of extracting summary statistics of scenes [79], attention is spontaneously biased to structures identified implicitly by SL [79–81], and SL reduces perceived numerosity [79,80]. SL enhances memory for element of learned triplets and reduces memory for inserted distractors [82], alters the internal representation of pair elements based on their predictability [83], and it can create novel object associations based on transitive relations [84*]. Importantly, these kinds of associations do not only establish novel links between the identity of elements, but also influence perception of features across elements. For example, after learning that two elements belong to the same pair, seeing one of them at a different size will influence the observer’s perception of the size of the other element [85**]. These effects have been typically conceptualized as top-down influences reaching down to even the most basic attributes, such as motion perception [86] or rivalry [87], and they can be manifested neurally at the lowest level of cortical representations [88] similarly to findings in PL.

The above summary suggest that in contrast to their original conceptualization, PL and SL share characteristics in almost every domain. Both of them can influence various neural metrics at multiple levels of the cortical hierarchy from primary sensory to high-level areas, both of them involve strong top-down effects, and show flexible generalization depending on context.

(Figure 1 Legend Continued) Perceptual expertise task of bird species discrimination showing both improving performance with trained birds (blue curve) and generalization to previously unseen birds (transfer i.e. no initial jump in green curve). **(e)** Structure and references of the dominant computational models in PL assuming tuning changes in the representational units (orange) or re-weighting of representation-to-decision connections (blue). **(f–g)** Classical spatial visual SL task with the inventory, the composed set of training scenes, the segmented substructures of the training scenes (“chunks”) vs. random shape combinations used as test scenes, and the corresponding familiarity performance with the tests scenes indicating generalization of learning. **(h–i)** Same as f–g but with classical temporal visual SL task using a long temporal chain of shape images as a training sequence and shape triplets presented consecutively as test stimuli. **(j)** Structure and references of the dominant computational models in SL based on non-probabilistic (green) and probabilistic (turquoise) latent chunk learning, and on biologically and computationally motivated connectionists learning (brown). **(l)** Reports on neural correlates of PL (red) and SL (blue) ordered along two relevant dimensions: the complexity of the reported neural correlate modulated by learning (x axis), and the rough position of the investigated brain area within the cortical hierarchy (y axis) colored in red/blue according to which learning was found to influence the area predominantly. Dashed areas indicate typical combinations of neural correlates and involved areas of PL (red) and SL (blue) ([113,114,115,116,117,118,119,120,121,122,123,124,125,126,127,128,129,130*,131,132,133,134,135,136*,137,138,139,140,141,142*,143,144,145,146]).

Figure 2



Vanishing differences between perceptual (PL) and statistical learning (SL). The relationships between PL (pink area) and SL (blue area) mapped onto the two dimensions of stimulus complexity (x axis) and task specificity (y axis). In recent studies [70,80,85**] using more complex stimuli and a larger variability in the selected task that can create more natural conditions (green area), the classical separation between PL and SL waned. However, a systematic exploration on the integration of PL and SL (striped area) with specific new paradigms (A,B,C & D) still awaits. Bracketed numbers indicate references for previous studies [6,22,23,26,28,32,33,67,70,80,83,85,100,103,104,106,147–149], while letters indicate proposed new experiments (see legend on the right).

A unified probabilistic framework for PL and SL

Given the diminishing difference between PL and SL, a parsimonious approach to sensory learning is to define a framework that can seamlessly integrate studies and results in the two domains. A particularly suitable scheme is the probabilistic learning framework that has emerged in the field of machine learning [89], cognitive psychology [90], and neuroscience [91,92] over the last two decades. This framework inherently combines sensory bottom-up and experience-based top-down influences relying on their relative uncertainty to describe information processing in the brain [92–94]. More recent hierarchical extensions of the framework under the name of Hierarchical Bayesian Models (HBM) can potentially capture the full complexity of human learning including high cognitive functions such as abstract concept formation, language acquisition and causal learning [95,96].

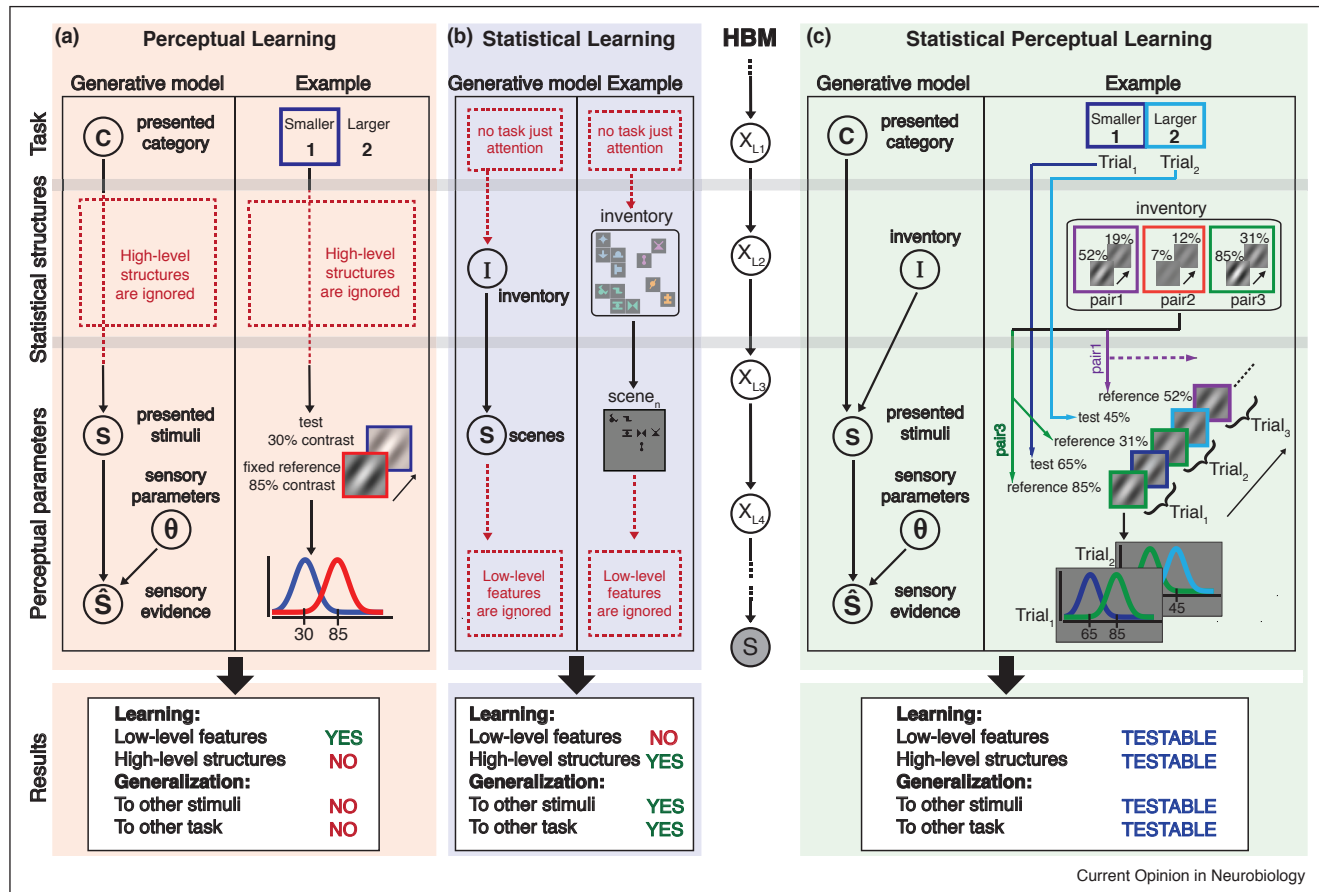
Our main proposal is that that PL and SL should be treated jointly in the framework of HBM, since they are not two separate types of learning, but two extreme testing paradigms of the same complex learning mechanism, in which either more complex structures and context (in case of PL) or the treatment of low level fine sensory features (in case of SL) have been deliberately eliminated (Figure 3).

Although there were earlier studies linking the probabilistic framework to either PL [97,98] or to SL [99,100], no studies have explored the benefit of treating PL and SL jointly under the same HBM framework. This is surprising, as the HBM framework inherently fits the overwhelming majority of natural learning situations, where both details of features and the more global structure and context of the sensory information might be relevant for successfully solving the task at hand.

By explicitly capturing different aspects of the input and the learning task through structured priors, the HBM approach is compatible and includes as special cases the Reweighting Models [56], two-stage models [101], and the Reverse Hierarchy Theory [102] of PL. By using a hierarchy of latent variables, the HBM approach is also compatible and includes as a special case the two-layer probabilistic chunk learning models of SL that are already known to capture human behavior better than the alternative associative learning and counting models [100,103]. Therefore, HBM can accommodate the wide variety of recently established results in the domains of both PL and SL, and facilitates a clearer separation of their causes.

Importantly, the integrated viewpoint of HBM also provides a useful guiding principle to identify the kind of experiments

Figure 3



Current Opinion in Neurobiology

Unifying PL and SL in a probabilistic framework. **HBM:** The scheme of the general Hierarchical Bayesian Model that provides a unified computational framework for classical perceptual (a) and statistical learning paradigms (b), as well as for the combination of the two (c). (a–c): Probabilistic interpretation of the three paradigms, each with the instantiation of the generative HBM within the given paradigm (left) and one example experiment (right) together with levels not controlled by the paradigm (red dashed rectangles). **Bottom row:** Features of each paradigm and questions that they can address. (a) PL example of a two alternative forced choice contrast discrimination task. (b) SL example of visual patterns learning. (c) Joint Statistical Perceptual Learning (SPL) of contrast discrimination with structured reference stimuli. The reference contrast is not selected randomly but it follows the order defined by sequentially chosen reference contrast-pairs from the inventory. While PL with randomly varying reference contrast levels is excessively hard, we expect that providing a statistical structure to the changes across reference levels (imitating natural conditions) enables and enhances PL. In the HBM of SPL, the observer's perception is formalized with a probability distribution over the stimulus (S) given her sensory evidence (\hat{S}):

$$(1) P(\hat{S}|\hat{S}) \propto \int P(\hat{S}|\hat{S}, \theta) P(\hat{S}|\hat{I}) P(\theta, \hat{I}) d\theta d\hat{I}$$

where θ , and \hat{I} denote, the sensory parameters and the structure of the task (c.f. inventory), respectively, and $P(\hat{S}|\hat{S})$ captures the observer's belief of the true stimulus given her sensory representation. Since under natural conditions, the observer does not know the structure (\hat{I}) or the sensory parameters (θ) given the structure, s/he has to learn them jointly:

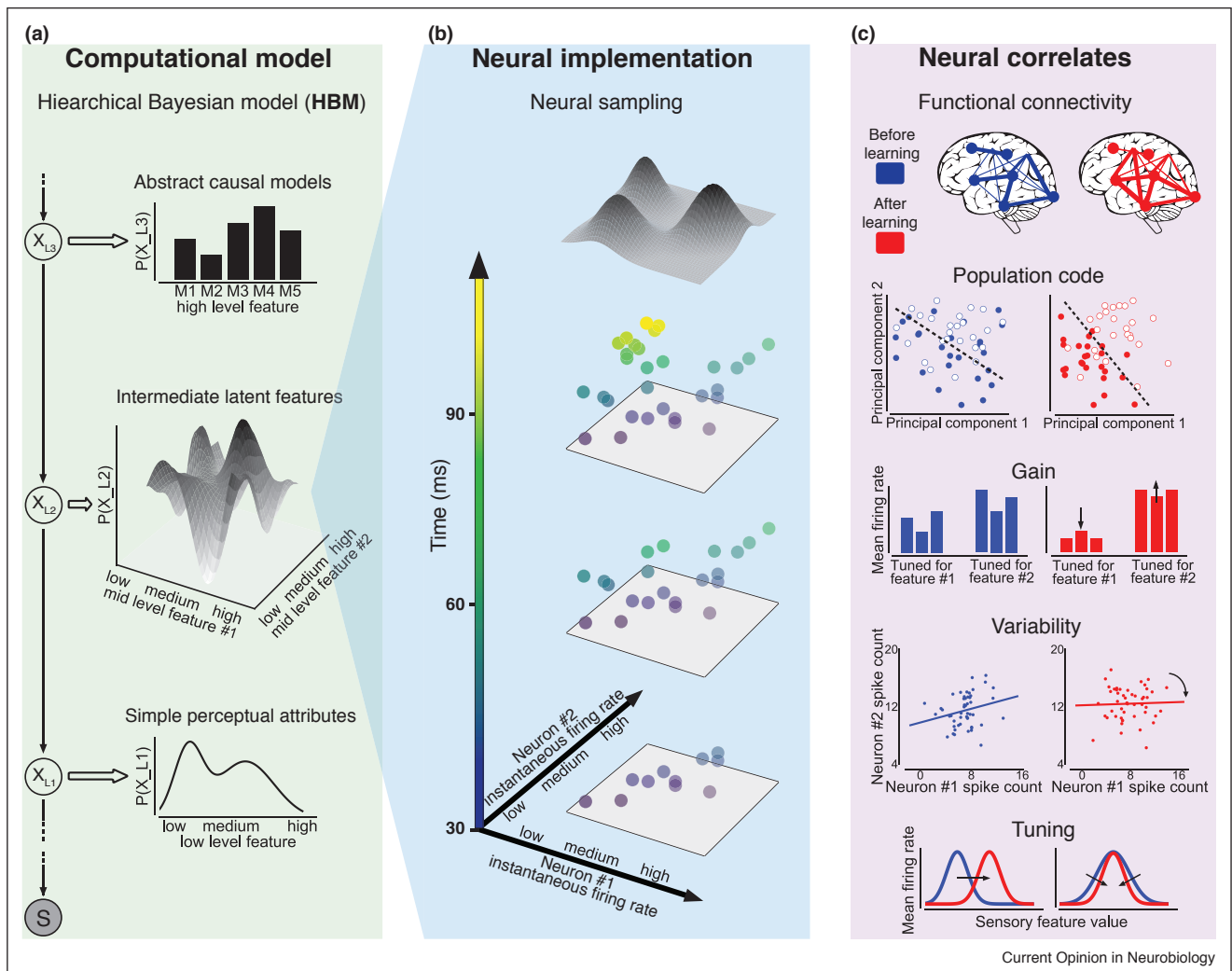
$$(2) P(\theta, \hat{I}|\hat{S}_{1:t}, F_{1:t}) \propto \int P(\hat{S}_t|\hat{S}_t, \theta) P(\hat{S}_t|\hat{I}_t, \hat{I}) d\hat{S}_t P(\theta, \hat{I}|\hat{S}_{1:t-1}, F_{1:t-1})$$

where F denotes the feedback (not shown in the graphical model) and t is the trial number. The three terms on the right side of Eq. (2) can be derived from the generative model (shown in c) and represent the low-level sensory model ($P(\hat{S}_t|\hat{S}_t, \theta)$), the high-level representation of the stimulus based on the task structure ($P(\hat{S}_t|\hat{I}_t, \hat{I})$), and the prior distribution which is the posterior at the previous time step ($P(\theta, \hat{I}|\hat{S}_{1:t-1}, F_{1:t-1})$). In this framework, classic PL (a) is framed as parameter learning [98], and classic SL (b) as structure learning [100]. PL without SL emerges when there is no uncertainty in the task structure or the feedback shows the true stimulus, thus the term $P(\hat{S}_t|\hat{I}_t, \hat{I})$ becomes a Dirac-delta. SL without PL is captured when there is no uncertainty in the sensory process thus the term $P(\hat{S}_t|\hat{S}_t, \theta)$ becomes a Dirac-delta. When PL and SL occur jointly, the interaction between the two types of learning can be investigated by using a PSL paradigm (c) and modelled by Eq. (2).

that could advance a fuller understanding of the nature of human and animal learning. The first type of experiments (Figure 2, Groups A,B) could use multi-element stimuli and ir/relevant cover stories with a PL task to explore how the effect of such sensory and cognitive context could be systematically captured as a consequence of priors acquired earlier by SL

(Figure 3c). These experiments could handle in a coherent manner rowing [70], generalization results of double-training [67,68,104], imagination-based learning [66], interaction between orientation detection and categorization [105] and perceptual biases due to SL [79,80,81,82,84,85,87]. The second type of experiments (Figure 2, Groups C,D) could

Figure 4



Linking the proposed HBM framework for PL and SL to different neural correlates through a probabilistic sampling-based neural implementation. (a) In the HBM (left), the stimulus (S) is jointly described by observed and latent features of the environment, which are represented by momentary posterior distributions, $P(X_{L_i})$, over possible values of latent variables, X_{L_i} , at different levels of abstraction. (b) According to the neural sampling hypothesis, covarying neural activities within different cortical areas directly represent the probability distributions over the latent variables of the HBM as samples from that distribution. For each probability distribution (depicted here for latent variables at a middle level of abstraction shown in (a)), the individual samples of the joint instantaneous firing rates of neurons at a given time frame (dots) accumulate through time (y axis, also color code of dots), and they jointly approximate the probability distribution of the latent variable (grey 2D distribution on top) with an increasing precision. (c) Various previously reported neural correlates of sensory learning that can be potentially derived from the sampling-based probabilistic representation of latent variables. These include shifts and sharpening of tuning curves, decorrelation of neural responses, and changes in gain, population codes [108*,112**], and, functional connectivity of neural clusters.

extend the first one by using natural scene inputs instead of artificial stimuli and could be applied to explain the high generalization of bird (and other) experts [22,23], task-structure learning [78**] and increased PL performance after video-game playing [77,106].

A sampling-based probabilistic implementation for HBMs exploring PL and SL

One of the main obstacles hindering progress in PL and SL research is due to correlating widely different aspects

of neural activity with learning (Figure 11, x axis). Although our proposal of introducing HBMs for the computational treatment of learning seems to further complicate this problem, in fact, the probabilistic view offers a unification and clarification on earlier results. As the probabilistic computational framework inherently requires a new type of conversion and approximation from abstract computational descriptions by probability distribution to neural signals [91,107], the new representations can provide a principled way to establish a rigorous

link between the different types of neural correlates of learning. In particular, sampling-based approximations have been argued to fit well the available neural evidence for perception and learning in the brain [92,108*] (Figure 4). Various other implementational frameworks can also capture top-down influences of neural signals such as effects of decision making and attention based on recurrency. These include recursive neural network models [109], Predictive Coding [110] or Probabilistic Population Codes for Bayesian inference making [111]. However, sampling-based methods offer a potentially more precise link between computations and various manifestations neural correlations including neural tuning curves, response means and variability, correlations and population sparseness [112**] that can likely be recursively extended to higher levels of the hierarchy.

Conclusions

We proposed that PL and SL should be treated uniformly and jointly under the HBM framework because this would enable addressing more natural and complex learning problems than before, and because combined with the probabilistic sampling approximation, such a treatment could link more successfully abstract computations of learning with various cortical and subcortical processes. Following this approach, a number of new experimental paradigms can be developed that combine the characteristics of PL and SL paradigms for a more in-depth investigation of human and animal learning and its neural correlates.

Conflict of interest

Nothing declared.

Acknowledgements

This research was supported by the Wellcome Trust, the European Research Council, NIH and by ONR.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Sagi D: **Perceptual learning: learning to see.** *Curr. Opin. Neurobiol.* 1994, **4**(2):195-199 [http://dx.doi.org/10.1016/0959-4388\(94\)90072-8](http://dx.doi.org/10.1016/0959-4388(94)90072-8).
2. Fahle M, Poggio T, Poggio TA: *Perceptual Learning*. MIT Press; 2002 <http://dx.doi.org/10.7551/mitpress/5295.001.0001>.
3. Adini Y, Wilkowsky A, Haspel R, Tsodyks M, Sagi D: **Perceptual learning in contrast discrimination: the effect of contrast uncertainty.** *J. Vis.* 2004, **4**(12):993-1005 <http://dx.doi.org/10.1167/4.12.2>.
4. Yu C, Klein SA, Levi DM: **Perceptual learning in contrast discrimination and the (minimal) role of context.** *J. Vis.* 2004, **4**(3):169-182 <http://dx.doi.org/10.1167/4.3.4>.
5. Ball K, Sekuler R: **Direction-specific improvement in motion discrimination.** *Vision Res.* 1987, **27**(6):953-965. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/3660656> [https://linkinghub.elsevier.com/retrieve/pii/0042-6989\(87\)90011-3](https://linkinghub.elsevier.com/retrieve/pii/0042-6989(87)90011-3) <https://www.ncbi.nlm.nih.gov/pubmed/3660656> <https://www.ncbi.nlm.nih.gov/pubmed/3660656>.
6. Fiorentini A, Berardi N: **Perceptual learning specific for orientation and spatial frequency.** *Nature* 1980, **287**(5777):43-44 <http://dx.doi.org/10.1038/287043a0>.
7. Ahissar M, Hochstein S: **Task difficulty and the specificity of perceptual learning.** *Nature* 1997, **387**(6631):401-406 <http://dx.doi.org/10.1038/387401a0>.
8. Karni A, Sagi D: **Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity.** *Proc. Natl. Acad. Sci. U. S. A.* 1991, **88**(11):4966-4970. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/2052578> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC51788> <http://www.pnas.org/cgi/pmidlookup?view=long&pmid=2052578> <https://www.ncbi.nlm.nih.gov/pubmed/2052578> In: <https://www.ncbi.nlm.nih.gov/pubmed/2052578>.
9. Spang K, Grimsen C, Herzog MH, Fahle M: **Orientation specificity of learning vernier discriminations.** *Vision Res.* 2010, **50**(4):479-485 <http://dx.doi.org/10.1016/j.visres.2009.12.008>.
10. O'Toole AJ, Kersten DJ: **Learning to see random-dot stereograms.** *Perception* 1992, **21**(2):227-243 <http://dx.doi.org/10.1068/p210227>.
11. Jeter PE, Doshier BA, Liu S-H, Lu Z-L: **Specificity of perceptual learning increases with increased training.** *Vision Res.* 2010, **50**(19):1928-1940 <http://dx.doi.org/10.1016/j.visres.2010.06.016>.
12. Miyamoto D, Hirai D, Fung CCA, Inutsuka A, Odagawa M, Suzuki T, Boehringer R, Adaikkan C, Matsubara C, Matsuki N et al.: **Top-down cortical input during NREM sleep consolidates perceptual memory.** *Science* 2016, **352**(6291):1315-1318 <http://dx.doi.org/10.1126/science.aaf0902>.
13. Karni A, Tanne D, Rubenstein B, Askenasy J, Sagi D: **Dependence on REM sleep of overnight improvement of a perceptual skill.** *Science* 1994, **265**(5172):679-682 <http://dx.doi.org/10.1126/science.8036518>.
14. Karni A, Sagi D: **The time course of learning a visual skill.** *Nature* 1993, **365**(6443):250-252 <http://dx.doi.org/10.1038/365250a0>.
15. Watanabe T, Sasaki Y: **Perceptual Learning: Toward a Comprehensive Theory.** *Annu. Rev. Psychol.* 2015, **66**(1):197-221 <http://dx.doi.org/10.1146/annurev-psych-010814-015214>.
16. Aberg KC, Herzog MH: **Different types of feedback change decision criterion and sensitivity differently in perceptual learning.** *J Vis* 2012, **12**(3) <http://dx.doi.org/10.1167/12.3.3>.
17. Shiu LP, Pashler H: **Improvement in line orientation discrimination is retinally local but dependent on cognitive set.** *Percept Psychophys* 1992, **52**(5):582-588 Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/1437491> <https://www.ncbi.nlm.nih.gov/pubmed/1437491>.
18. Schoups AA, Vogels R, Orban GA: **Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularly.** *J Physiol* 1995, **483**(Pt 3):797-810 Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/7776259> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1157819> <https://onlinelibrary.wiley.com/resolve/openurl?genre=article&sid=nlm:pubmed&issn=0022-3751&date=1995&volume=483&issue=&page=797> <https://onlinelibrary.wiley.com/resolve/openurl?genre=article&sid=nlm:pubmed&issn=0022-3751&date=1995&volume=483&issue=&page=797>.
19. Fahle M, Morgan M: **No transfer of perceptual learning between similar stimuli in the same retinal position.** *Curr Biol* 1996, **6**(3):292-297 <https://www.ncbi.nlm.nih.gov/pubmed/8805246> [https://linkinghub.elsevier.com/retrieve/pii/S0960-9822\(02\)00479-7](https://linkinghub.elsevier.com/retrieve/pii/S0960-9822(02)00479-7) <https://www.ncbi.nlm.nih.gov/pubmed/8805246>.
20. Crist RE, Kapadia MK, Westheimer G, Gilbert CD: **Perceptual learning of spatial localization: specificity for orientation, position, and context.** *J Neurophysiol* 1997, **78**(6):2889-2894 <http://dx.doi.org/10.1152/jn.1997.78.6.2889>.
21. Schoups A, Vogels R, Qian N, Orban G: **Practising orientation identification improves orientation coding in V1 neurons.** *Nature* 2001, **412**(6846):549-553 <http://dx.doi.org/10.1038/35087601>.
22. Tanaka JW, Curran T, Sheinberg DL: **The training and transfer of real-world perceptual expertise.** *Psychol Sci* 2005, **16**(2):145-151 <http://dx.doi.org/10.1111/j.0956-7976.2005.00795.x>.

23. Devillez H, Mollison MV, Hagen S, Tanaka JW, Scott LS, Curran T: **Color and spatial frequency differentially impact early stages of perceptual expertise training.** *Neuropsychologia* 2018 **11**:011. <http://dx.doi.org/10.1016/j.neuropsychologia.2018.11.011>.
 24. Aslin RN: **Statistical learning: a powerful mechanism that operates by mere exposure.** *Wiley Interdiscip Rev Cogn Sci* 2017, **8**:1-2 <http://dx.doi.org/10.1002/wcs.1373>.
 25. Saffran JR, Kirkham NZ: **Infant Statistical Learning.** *Annu Rev Psychol* 2018, **69**(1):181-203 <http://dx.doi.org/10.1146/annurev-psych-122216-011805>.
 26. Saffran JR, Aslin RN, Newport EL: **Statistical Learning by 8-Month-Old Infants.** *Science* 1996, **274**(5294):1926-1928 <http://dx.doi.org/10.1126/science.274.5294.1926>.
 27. Fiser J, Aslin RN: **Unsupervised statistical learning of higher-order spatial structures from visual scenes.** *Psychol Sci* 2001, **12**(6):499-504 <http://dx.doi.org/10.1111/1467-9280.00392>.
 28. Kirkham NZ, Slemmer JA, Johnson SP: **Visual statistical learning in infancy: evidence for a domain general learning mechanism.** *Cognition* 2002, **83**(2):B35-B42 <https://www.ncbi.nlm.nih.gov/pubmed/11869728> <https://linkinghub.elsevier.com/retrieve/pii/S0010027702000045> <https://www.ncbi.nlm.nih.gov/pubmed/11869728>.
 29. Newport EL: **Statistical language learning: computational, maturational, and linguistic constraints.** *Lang Cogn* 2016, **8**(03):447-461 <http://dx.doi.org/10.1017/langcog.2016.20>.
 30. Erickson LC, Thiessen ED: **Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition.** *Dev Rev* 2015, **37**:66-108 <http://dx.doi.org/10.1016/j.dr.2015.05.002>.
 31. Bulf H, Johnson SP, Valenza E: **Visual statistical learning in the newborn infant.** *Cognition* 2011, **121**(1):127-132 <http://dx.doi.org/10.1016/j.cognition.2011.06.010>.
 32. Fiser J, Aslin RN: **Statistical learning of new visual feature combinations by infants.** *Proc Natl Acad Sci U S A* 2002, **99**(24):15822-15826 <http://dx.doi.org/10.1073/pnas.232472899>.
 33. Fiser J, Aslin RN: **Encoding multielement scenes: statistical learning of visual feature hierarchies.** *J Exp Psychol Gen* 2005, **134**(4):521-537 <http://dx.doi.org/10.1037/0096-3445.134.4.521>.
 34. Slone LK, Johnson SP: **When learning goes beyond statistics: Infants represent visual sequences in terms of chunks.** *Cognition* 2018, **178**:92-102 <http://dx.doi.org/10.1016/j.cognition.2018.05.016>.
 35. Conway CM, Christiansen MH: **Modality-constrained statistical learning of tactile, visual, and auditory sequences.** *J Exp Psychol Learn Mem Cogn* 2005, **31**(1):24-39 <http://dx.doi.org/10.1037/0278-7393.31.1.24>.
 36. Glicksohn A, Cohen A: **The role of cross-modal associations in statistical learning.** *Psychon Bull Rev* 2013, **20**(6):1161-1169 <http://dx.doi.org/10.3758/s13423-013-0458-4>.
 37. Ongchoco J, Uddenberg S, Chun M: **Statistical learning of movement.** *J Vis* 2016, **16**(12):1079 <http://dx.doi.org/10.1167/16.12.1079>.
 38. Lengyel G, alalytė G, Pantelides A, Ingram JN, Fiser J, Lengyel M, Wolpert DM: **Unimodal statistical learning produces multimodal object-like representations.** *Elife* 2019;8 <http://dx.doi.org/10.7554/eLife.43942>.
 39. Brady TF, Oliva A: **Statistical learning using real-world scenes: extracting categorical regularities without conscious intent.** *Psychol Sci* 2008, **19**(7):678-685 <http://dx.doi.org/10.1111/j.1467-9280.2008.02142.x>.
 40. Turk-Browne NB, Isola PJ, Scholl BJ, Treat TA: **Multidimensional visual statistical learning.** *J Exp Psychol Learn Mem Cogn* 2008, **34**(2):399-407 <http://dx.doi.org/10.1037/0278-7393.34.2.39>.
 41. Toro JM, Trobalón JB: **Statistical computations over a speech stream in a rodent.** *Percept Psychophys*. 2005, **67**(5):867-875 <http://dx.doi.org/10.3758/bf03193539>.
 42. Rosa-Salva O, Fiser J, Versace E, Dolci C, Chehaimi S, Santolin C, Vallortigara G: **Spontaneous Learning of Visual Structures in Domestic Chicks.** *Animals (Basel)* 2018, **8**(8) <http://dx.doi.org/10.3390/ani8080135>.
 43. Santolin C, Saffran JR: **Constraints on statistical learning across species.** *Trends Cogn Sci* 2018, **22**(1):52-63 <http://dx.doi.org/10.1016/j.tics.2017.10.003>.
 44. Santolin C, Rosa-Salva O, Vallortigara G, Regolin L: **Unsupervised statistical learning in newly hatched chicks.** *Curr Biol* 2016, **26**(23):R1218-R1220 <http://dx.doi.org/10.1016/j.cub.2016.10.011>.
 45. Castro L, Wasserman EA, Lauffer M: **Unsupervised learning of complex associations in an animal model.** *Cognition* 2018, **173**:28-33 <http://dx.doi.org/10.1016/j.cognition.2017.12.015>.
 46. Fiser J: **The other kind of perceptual learning.** *Learn Percept* 2009, **1**(1):69-87 <http://dx.doi.org/10.1556/lp.1.2009.1.6>.
 47. Frost R, Armstrong BC, Siegelman N, Christiansen MH: **Domain generality versus modality specificity: the paradox of statistical learning.** *Trends Cogn Sci* 2015, **19**(3):117-125 <http://dx.doi.org/10.1016/j.tics.2014.12.01>.
 48. Kim R, Seitz A, Feenstra H, Shams L: **Testing assumptions of statistical learning: is it long-term and implicit?** *Neurosci Lett* 2009, **461**(2):145-149 <http://dx.doi.org/10.1016/j.neulet.2009.06.030>.
 49. Nemeth D, Janacek K, Londe Z, Ullman MT, Howard DV, Howard JH: **Sleep has no critical role in implicit motor sequence learning in young and old adults.** *Exp Brain Res* 2009, **201**(2):351-358 <http://dx.doi.org/10.1007/s00221-009-2024-x>.
 50. Simor P, Zavecz Z, Horvth K, Itet N, Trk C, Pesthy O, Gombos F, Janacek K, Nemeth D: **Deconstructing procedural memory: different learning trajectories and consolidation of sequence and statistical learning.** *Front Psychol* 2018, **9**:2708 <http://dx.doi.org/10.3389/fpsyg.2018.02708>.
 51. Turk-Browne NB, Jungé J, Scholl BJ: **The automaticity of visual statistical learning.** *J Exp Psychol Gen* 2005, **134**(4):552-564 <http://dx.doi.org/10.1037/0096-3445.134.4.552>.
 52. Musz E, Weber MJ, Thompson-Schill SL: **Visual statistical learning is not reliably modulated by selective attention to isolated events.** *Atten Percept Psychophys* 2014, **77**(1):78-96 <http://dx.doi.org/10.3758/s13414-014-0757-5>.
 53. Altmann GTM: **Abstraction and generalization in statistical learning: implications for the relationship between semantic types and episodic tokens.** *Philos Trans R Soc Lond B Biol Sci* 2017, **372** <http://dx.doi.org/10.1098/rstb.2016.0060>.
 54. Marcus GF, Vijayan S, Bandi Rao S, Vishton PM: **Rule learning by seven-month-old infants.** *Science* 1999, **283**(5398):77-80 Retrieved from: <https://www.ncbi.nlm.nih.gov/pubmed/9872745> <http://www.sciencemag.org/cgi/pmidlookup?view=long&pmid=9872745>.
 55. Saffran JR, Pollak SD, Seibel RL, Shkolnik A: **Dog is a dog is a dog: Infant rule learning is not specific to language.** *Cognition* 2007, **105**(3):669-680 <http://dx.doi.org/10.1016/j.cognition.2006.11.004>.
 56. Doshier B, Lu Z-L: **Visual Perceptual Learning and Models.** *Annual Review of Vision Science* 2017, **3**(1):343-363 <http://dx.doi.org/10.1146/annurev-vision-102016-061249>.
 57. LeMessurier AM, Feldman DE: **Plasticity of population coding in primary sensory cortex.** *Curr Opin Neurobiol* 2018, **53**:50-56 <http://dx.doi.org/10.1016/j.conb.2018.04.029>.
- This paper tabulates and characterizes the diverging neural manifestations of sensory learning in local cortical circuits, bottom-up and top-down projections and even in distant motor pathways, and reviews their location in the cortex. The wide spectrum of changes in responses and variability in neural populations due to learning emphasizes the need for an integrated approach to understand population codes of learning.
58. Maniglia M, Seitz AR: **Towards a whole brain model of Perceptual Learning.** *Curr Opin Behav Sci* 2018, **20**:47-55 <http://dx.doi.org/10.1016/j.cobeha.2017.10.004>.
 59. Diaz JA, Queirazza F, Philastides MG: **Perceptual learning alters post-sensory processing in human decision-making.** *Nature Human Behaviour* 2017, **1**(2):0035 <http://dx.doi.org/10.1038/s41562-016-0035>.

60. Li W, Pich V, Gilbert CD: **Perceptual learning and top-down influences in primary visual cortex.** *Nat Neurosci* 2004, **7(6)**:651-657 <http://dx.doi.org/10.1038/nn1255>.
61. Law C-T, Gold JI: **Shared mechanisms of perceptual learning and decision making.** *Top Cogn Sci* 2010, **2(2)**:226-238 <http://dx.doi.org/10.1111/j.1756-8765.2009.01044.x>.
62. Law C-T, Gold JI: **Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area.** *Nat Neurosci* 2008, **11(4)**:505-513 <http://dx.doi.org/10.1038/nn2070>.
63. Kahnt T, Grueschow M, Speck O, Haynes J-D: **Perceptual learning and decision-making in human medial frontal cortex.** *Neuron* 2011, **70(3)**:549-559 <http://dx.doi.org/10.1016/j.neuron.2011.02.054>.
64. Donovan I, Carrasco M: **Endogenous spatial attention during perceptual learning facilitates location transfer.** *J Vis* 2018, **18(11)**:7 <http://dx.doi.org/10.1167/18.11.7>.
65. Donovan I, Szpiro S, Carrasco M: **Exogenous attention facilitates location transfer of perceptual learning.** *J Vis* 2015, **15(10)**:11 <http://dx.doi.org/10.1167/15.10.11>.
66. Tartaglia EM, Bamert L, Mast FW, Herzog MH: **Human Perceptual Learning by Mental Imagery.** *Curr Biol* 2009, **19(24)**:2081-2085 <http://dx.doi.org/10.1016/j.cub.2009.10.060>.
67. Xiao L-Q, Zhang J-Y, Wang R, Klein SA, Levi DM, Yu C: **Complete transfer of perceptual learning across retinal locations enabled by double training.** *Curr Biol* 2008, **18(24)**:1922-1926 <http://dx.doi.org/10.1016/j.cub.2008.10.030>.
68. Wang R, Zhang J-Y, Klein SA, Levi DM, Yu C: **Vernier perceptual learning transfers to completely untrained retinal locations after double training: a "piggybacking" effect.** *J Vis* 2014, **14(13)**:12 <http://dx.doi.org/10.1167/14.13.12>.
69. Wang R, Wang J, Zhang YJ, Xie YX, Yang XY, Luo HS, Yu C, Li W: **Perceptual Learning at a Conceptual Level.** *J Neurosci* 2016, **36(7)**:2238-2246 <http://dx.doi.org/10.1523/jneurosci.2732-15.2016>.
The authors demonstrate generalization of perceptual learning, as indicated by complete transfer of performance, across widely different types of stimuli including luminance gratings and axis of bilaterally symmetric dot patterns in an orientation discrimination task, and between first- and second-order motion signals in a motion direction discrimination task. These results of powerful generalization provide compelling evidence that improving abilities during perceptual learning prominently rely on higher-level cognitive processes.
70. Kuai S-G, Zhang J-Y, Klein SA, Levi DM, Yu C: **The essential role of stimulus temporal patterning in enabling perceptual learning.** *Nat Neurosci* 2005, **8(11)**:1497-1499 <http://dx.doi.org/10.1038/nn1546>.
71. Jeter PE, Doshier BA, Petrov A, Lu Z-L: **Task precision at transfer determines specificity of perceptual learning.** *J Vis* 2009, **9(1)**:1-13 <http://dx.doi.org/10.1167/9.3.1>.
72. Chang DHF, Kourtzi Z, Welchman AE: **Mechanisms for extracting a signal from noise as revealed through the specificity and generality of task training.** *J Neurosci* 2013, **33(27)**:10962-10971 <http://dx.doi.org/10.1523/JNEUROSCI.0101-3.2013>.
73. Lengyel G, Fiser J: **The relationship between initial threshold, learning, and generalization in perceptual learning.** *J Vision* 2019, **19(4)**:28 <http://dx.doi.org/10.1167/19.4.28>.
74. Chang DHF, Mevorach C, Kourtzi Z, Welchman AE: **Training transfers the limits on perception from parietal to ventral cortex.** *Curr Biol* 2014, **24(20)**:2445-2450 <http://dx.doi.org/10.1016/j.cub.2014.08.058>.
75. Hussain Z, Bennett PJ, Sekuler AB: **Versatile perceptual learning of textures after variable exposures.** *Vision Res* 2012, **61**:89-94 <http://dx.doi.org/10.1016/j.visres.2012.01.005>.
76. Green CS, Kattner F, Siegel MH, Kersten D, Schrater PR: **Differences in perceptual learning transfer as a function of training task.** *J Vis* 2015, **15(10)**:5 <http://dx.doi.org/10.1167/15.10.5>.
77. Green CS, Shawn Green C, Li R, Bavelier D: **Perceptual Learning During Action Video Game Playing.** *Top Cogn Sci* 2010, **2(2)**:202-216 <http://dx.doi.org/10.1111/j.1756-8765.2009.01054.x>.
78. Kattner F, Cochrane A, Cox CR, Gorman TE, Green CS: **Perceptual Learning Generalization from Sequential Perceptual Training as a Change in Learning Rate.** *Curr Biol* 2017, **27(6)**:840-846 <http://dx.doi.org/10.1016/j.cub.2017.01.046>.
The authors provided evidence for learning at different computational levels during perceptual learning through separating immediate generalization as the measure of classical learning versus changes in learning speed as the measure of higher-level statistical learning "learning to learn". Depending on the preceding training protocol, observers in a novel dot motion direction categorization task showed different speed of learning even though their initial level of generalization in the task was identical.
79. Zhao J, Ngo N, McKendrick R, Turk-Browne NB: **Mutual interference between statistical summary perception and statistical learning.** *Psychol Sci* 2011, **22(9)**:1212-1219 <http://dx.doi.org/10.1177/0956797611419304>.
80. Zhao J, Yu RQ: **Statistical regularities reduce perceived numerosity.** *Cognition* 2016, **146**:217-222 <http://dx.doi.org/10.1016/j.cognition.2015.09.018>.
81. Zhao J, Al-Aidroos N, Turk-Browne NB: **Attention is spontaneously biased toward regularities.** *Psychol Sci* 2013, **24(5)**:667-677 <http://dx.doi.org/10.1177/0956797612460407>.
82. Otsuka S, Saiki J: **Gift from statistical learning: Visual statistical learning enhances memory for sequence elements and impairs memory for items that disrupt regularities.** *Cognition* 2016, **147**:113-126 <http://dx.doi.org/10.1016/j.cognition.2015.11.00>.
83. Barakat BK, Seitz AR, Shams L: **The effect of statistical learning on internal stimulus representations: Predictable items are enhanced even when not predicted.** *Cognition* 2013, **129(2)**:205-211 <http://dx.doi.org/10.1016/j.cognition.2013.07.003>.
84. Luo Y, Zhao J: **Statistical Learning Creates Novel Object Associations via Transitive Relations.** *Psychol Sci* 2018, **29(8)**:1207-1220 <http://dx.doi.org/10.1177/0956797618762400>.
In several experiments using the temporal statistical learning paradigm, the authors showed that observers made new transitive associations between items that never co-occurred together in the experiment. Specifically, experiencing two base pairs, A-B and B-C, observers inferred the transitive relation of A-C either with colored circles or with country names. These results point toward multi-level hierarchical internal representation and an automatic learning across these levels.
85. Yu RQ, Zhao J: **Implicit updating of object representation via temporal associations.** *Cognition* 2018, **181**:127-134 <http://dx.doi.org/10.1016/j.cognition.2018.08.015>.
This paper shows that after temporal statistical learning of consecutive pairing of two colored circles or two faces, participants automatically and without explicit awareness of the pair structures transferred the physical property change of the first item of the pair (color of circle or expression of the face) to the second one during their perceptual judgement, while they failed to do so in the reverse direction. This finding provides evidence that beyond simple associations, representations of objects in the brain are automatically updated based on the relational structure of those representation.
86. Sotiropoulos G, Seitz AR, Seris P: **Changing expectations about speed alters perceived motion direction.** *Curr Biol* 2011, **21(21)**:R883-R884 <http://dx.doi.org/10.1016/j.cub.2011.09.013>.
87. Piazza EA, Denison RN, Silver MA: **Recent cross-modal statistical learning influences visual perceptual selection.** *J Vis* 2018, **18(3)**:1 <http://dx.doi.org/10.1167/18.3.1>.
88. Kver H, Gill K, Tseng Y-TL, Bao S: **Perceptual and neuronal boundary learned from higher-order stimulus probabilities.** *J Neurosci* 2013, **33(8)**:3699-3705 <http://dx.doi.org/10.1523/JNEUROSCI.3166-12.2013>.
89. Ghahramani Z: **Probabilistic machine learning and artificial intelligence.** *Nature* 2015, **521(7553)**:452-459 <http://dx.doi.org/10.1038/nature14541>.
90. Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND: **How to Grow a Mind: Statistics, Structure, and Abstraction.** *Science* 2011, **331(6022)**:1279-1285 <http://dx.doi.org/10.1126/science.1192788>.
91. Knill DC, Pouget A: **The Bayesian brain: the role of uncertainty in neural coding and computation.** *Trends Neurosci* 2004, **27(12)**:712-719 <http://dx.doi.org/10.1016/j.tins.2004.10.007>.

92. Fiser J, Berkes P, Orbán G, Lengyel M: **Statistically optimal perception and learning: from behavior to neural representations.** *Trends Cogn Sci* 2010, **14**(3):119-130 <http://dx.doi.org/10.1016/j.tics.2010.01.003>.
93. Rao RPN, Olshausen BA, Lewicki MS (Eds): *Neural information processing series. Probabilistic models of the brain: Perception and neural function.* The MIT Press: Cambridge, MA, US; 2002 <http://dx.doi.org/10.7551/mitpress/5583.001.0001>.
94. Kersten D, Mamassian P, Yuille A: **Object Perception as Bayesian Inference.** *Annu Rev Psychol* 2004, **55**(1):271-304 <http://dx.doi.org/10.1146/annurev.psych.55.090902.142005>.
95. Tenenbaum JB, Griffiths TL, Kemp C: **Theory-based Bayesian models of inductive learning and reasoning.** *Trends Cogn Sci* 2006, **10**(7):309-318 <http://dx.doi.org/10.1016/j.tics.2006.05.009>.
96. Lake BM, Salakhutdinov R, Tenenbaum JB: **Human-level concept learning through probabilistic program induction.** *Science* 2015, **350**(6266):1332-1338 <http://dx.doi.org/10.1126/science.aab3050>.
97. Bejjanki VR, Beck JM, Lu Z-L, Pouget A: **Perceptual learning as improved probabilistic inference in early sensory areas.** *Nat Neurosci* 2011, **14**(5):642-648 <http://dx.doi.org/10.1038/nn.2796>.
98. Michel MM, Jacobs RA: **Parameter learning but not structure learning: a Bayesian network model of constraints on early perceptual learning.** *J Vis* 2007, **7**(1):4 <http://dx.doi.org/10.1167/7.1.4>.
99. Goldwater S, Griffiths TL, Johnson M: **A Bayesian framework for word segmentation: exploring the effects of context.** *Cognition* 2009, **112**(1):21-54 <http://dx.doi.org/10.1016/j.cognition.2009.03.008>.
100. Orbán G, Fiser J, Aslin RN, Lengyel M: **Bayesian learning of visual chunks by human observers.** *Proc Natl Acad Sci U S A* 2008, **105**(7):2745-2750 <http://dx.doi.org/10.1073/pnas.0708424105>.
101. Shibata K, Sagi D, Watanabe T: **Two-stage model in perceptual learning: toward a unified theory.** *Ann N Y Acad Sci* 2014, **1316**:18-28 <http://dx.doi.org/10.1111/nyas.12419>.
102. Ahissar M, Hochstein S: **The reverse hierarchy theory of visual perceptual learning.** *Trends Cogn Sci* 2004, **8**(10):457-464 <http://dx.doi.org/10.1016/j.tics.2004.08.011>.
103. Austerweil JL, Griffiths TL: **A rational model of the effects of distributional information on feature learning.** *Cogn Psychol* 2011, **63**(4):173-209 <http://dx.doi.org/10.1016/j.cogpsych.2011.08.002>.
104. Xiong Y-Z, Zhang J-Y, Yu C: **Bottom-up and top-down influences at untrained conditions determine perceptual learning specificity and transfer.** *Elife* 2016:5 <http://dx.doi.org/10.7554/eLife.14614>.
105. Tan Q, Wang Z, Sasaki Y, Watanabe T: **Category-Induced Transfer of Visual Perceptual Learning.** *Curr Biol* 2019, **29**(8) <http://dx.doi.org/10.1016/j.cub.2019.03.003> 1374-1378.e1373.
106. Green CS, Shawn Green C, Pouget A, Bavelier D: **Improved Probabilistic Inference as a General Learning Mechanism with Action Video Games.** *Curr Biol* 2010, **20**(17):1573-1579 <http://dx.doi.org/10.1016/j.cub.2010.07.040>.
107. Pouget A, Dayan P, Zemel RS: **Inference and computation with population codes.** *Annu Rev Neurosci* 2003, **26**:381-410 <http://dx.doi.org/10.1146/annurev.neuro.26.041002.131112>.
108. Haefner RM, Berkes P, Fiser J: **Perceptual Decision-Making as Probabilistic Inference by Neural Sampling.** *Neuron* 2016, **90**(3):649-660 <http://dx.doi.org/10.1016/j.neuron.2016.03.020>.
This paper demonstrates that a hierarchical probabilistic inference model of perception implemented with sampling-based approximation can naturally capture earlier findings such as task-dependence of neural noise correlations, and the modulation of the choice probabilities and psychophysical kernels across time that were attributed before to some cognitive top-down effects.
109. Pich V, Li W, Reeke GN, Gilbert CD: **Network model of top-down influences on local gain and contextual interactions in visual cortex.** *Proc Natl Acad Sci U S A* 2013, **110**(43):E4108-E4117 <http://dx.doi.org/10.1073/pnas.1317019110>.
110. Aitchison L, Lengyel M: **With or without you: predictive coding and Bayesian inference in the brain.** *Curr Opin Neurobiol* 2017, **46**:219-227 <http://dx.doi.org/10.1016/j.conb.2017.08.010>.
111. Pouget A, Beck JM, Ma WJ, Latham PE: **Probabilistic brains: knowns and unknowns.** *Nat Neurosci* 2013, **16**(9):1170-1178 <http://dx.doi.org/10.1038/nn.3495>.
112. Orbán G, Berkes P, Fiser J, Lengyel M: **Neural Variability and Sampling-Based Probabilistic Representations in the Visual Cortex.** *Neuron* 2016, **92**(2):530-543 <http://dx.doi.org/10.1016/j.neuron.2016.09.038>.
The authors present a normative theory and model of probabilistic neural coding in the brain and provide a quantitative link between the theory and various manifestations of neural activity in the cortex through a sampling-based approximation. The model captures and synthesized a number of previously unexplained empirical observations related to variability, spontaneous activity, and signal, and noise correlations across cell pairs in V1. Importantly, it provides a method to derive theory-based hypotheses that are directly testable by standard neurophysiological measures..
113. Teich AF, Qian N: **Learning and Adaptation in a Recurrent Model of V1 Orientation Selectivity.** *J Neurophysiol* 2003, **89**(4):2086-2100 <http://dx.doi.org/10.1152/jn.00970.2002>.
114. Schwabe L: **Adaptivity of Tuning Functions in a Generic Recurrent Network Model of a Cortical Hypercolumn.** *J Neurosci* 2005, **25**(13):3323-3332 <http://dx.doi.org/10.1523/jneurosci.4493-04.2005>.
115. Doshier BA, Jeter P, Liu J, Lu Z-L: **An integrated reweighting theory of perceptual learning.** *Proc Natl Acad Sci U S A* 2013, **110**(33):13678-13683 <http://dx.doi.org/10.1073/pnas.1312552110>.
116. Law C-T, Gold JI: **Reinforcement learning can account for associative and perceptual learning on a visual-decision task.** *Nat Neurosci* 2009, **12**(5):655-663 <http://dx.doi.org/10.1038/nn.2304>.
117. Petrov AA, Doshier BA, Lu Z-L: **The dynamics of perceptual learning: an incremental reweighting model.** *Psychol Rev* 2005, **112**(4):715-743 <http://dx.doi.org/10.1037/0033-295X.112.4.715>.
118. Mareschal D, French RM: **TRACX2: a connectionist autoencoder using graded chunks to model infant visual statistical learning.** *Philos Trans R Soc Lond B Biol Sci* 2017, **372**(1711) <http://dx.doi.org/10.1098/rstb.2016.0057>.
119. Perruchet P: **What Mechanisms Underlie Implicit Statistical Learning? Transitional Probabilities Versus Chunks in Language Learning.** *Top Cogn Sci* 2019, **11**(3):520-535 <http://dx.doi.org/10.1111/tops.12403>.
120. Schapiro AC, Turk- Browne NB, Botvinick MM, Norman KA: **Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning.** *Philos Trans R Soc Lond B Biol Sci* 2017, **372**(1711) <http://dx.doi.org/10.1098/rstb.2016.0049>.
121. Wang R, Shen Y, Tino P, Welchman AE, Kourtzi Z: **Learning Predictive Statistics: Strategies and Brain Mechanisms.** *J Neurosci* 2017, **37**(35):8412-8427 <http://dx.doi.org/10.1523/JNEUROSCI.0144-17.2017>.
122. Karlaftis VM, Wang R, Shen Y, Tino P, Williams G, Welchman AE, Kourtzi Z: **White-Matter Pathways for Statistical Learning of Temporal Structures.** *eNeuro* 2018, **5**(3) <http://dx.doi.org/10.1523/ENEURO.0382-17.2018>.
123. Alamia A, Solopchuk O, D'Ausilio A, Van Bever V, Fadiga L, Olivier E, Znon A: **Disruption of Broca's Area Alters Higher-order Chunking Processing during Perceptual Sequence Learning.** *J Cogn Neurosci* 2016, **28**(3):402-417 http://dx.doi.org/10.1162/jocn_a_00911.
124. Karuza EA, Newport EL, Aslin RN, Starling SJ, Tivarus ME, Bavelier D: **The neural correlates of statistical learning in a word segmentation task: An fMRI study.** *Brain Language* 2013, **127**(1):46-54 <http://dx.doi.org/10.1016/j.bandl.2012.11.007>.
125. Schapiro AC, Rogers TT, Cordova NI, Turk-Browne NB, Botvinick MM: **Neural representations of events arise from temporal community structure.** *Nat Neurosci* 2013, **16**(4):486-492 <http://dx.doi.org/10.1038/nn.3331>.

126. Turk-Browne NB, Scholl BJ, Chun MM, Johnson MK: **Neural evidence of statistical learning: efficient detection of visual regularities without awareness.** *J Cogn Neurosci* 2009, **21**(10):1934-1945 <http://dx.doi.org/10.1162/jocn.2009.21131>.
127. Schapiro AC, Kustner LV, Turk-Browne NB: **Shaping of object representations in the human medial temporal lobe based on temporal regularities.** *Curr Biol* 2012, **22**(17):1622-1627 <http://dx.doi.org/10.1016/j.cub.2012.06.056>.
128. Aly M, Chen J, Turk-Browne NB, Hasson U: **Learning Naturalistic Temporal Structure in the Posterior Medial Network.** *J Cogn Neurosci* 2018, **30**(9):1345-1365 http://dx.doi.org/10.1162/jocn_a_01308.
129. Bi T, Chen J, Zhou T, He Y, Fang F: **Function and structure of human left fusiform cortex are closely associated with perceptual learning of faces.** *Curr Biol* 2014, **24**(2):222-227 <http://dx.doi.org/10.1016/j.cub.2013.12.028>.
130. Ramachandran S, Meyer T, Olson CR: **Prediction suppression in monkey inferotemporal cortex depends on the conditional probability between images.** *J Neurophysiol* 2016, **115**(1):355-362 <http://dx.doi.org/10.1152/jn.00091.2015>.
 Single cell neural recordings in macaques' inferotemporal cortex in response to sequentially presented and concatenated stimulus pairs revealed that prediction suppression in neural responses is modulated by the conditional and not just the joint probability between elements of the pairs.
131. Meyer T, Olson CR: **Statistical learning of visual transitions in monkey inferotemporal cortex.** *Proc Natl Acad Sci U S A* 2011, **108**(48):19401-19406 <http://dx.doi.org/10.1073/pnas.1112895108>.
132. Kaposvari P, Kumar S, Vogels R: **Statistical Learning Signals in Macaque Inferior Temporal Cortex.** *Cereb Cortex* 2018, **28**(1):250-266 <http://dx.doi.org/10.1093/cercor/bhw374>.
133. Adab HZ, Popivanov ID, Vanduffel W, Vogels R: **Perceptual learning of simple stimuli modifies stimulus representations in posterior inferior temporal cortex.** *J Cogn Neurosci* 2014, **26**(10):2187-2200 http://dx.doi.org/10.1162/jocn_a_00641.
134. Karuza EA, Emberson LL, Roser ME, Cole D, Aslin RN, Fiser J: **Neural Signatures of Spatial Statistical Learning: Characterizing the Extraction of Structure from Complex Visual Scenes.** *J Cogn Neurosci* 2017, **29**(12):1963-1976 http://dx.doi.org/10.1162/jocn_a_01182.
135. Kuai S-G, Levi D, Kourtzi Z: **Learning optimizes decision templates in the human visual cortex.** *Curr Biol* 2013, **23**(18):1799-1804 <http://dx.doi.org/10.1016/j.cub.2013.07.052>.
136. Liu LD, Pack CC: **The Contribution of Area MT to Visual Motion Perception Depends on Training.** *Neuron* 2017, **95**(2) <http://dx.doi.org/10.1016/j.neuron.2017.06.024> 436-446.e3..
 The authors used pharmacological inactivation of the middle temporal area (MT) in primates before and after a long training on a perceptual task with drifting gratings and found that the strongest causal influence on motion perception shifted from MT to lower-level areas after practice.
137. Gu Y, Liu S, Fetsch CR, Yang Y, Fok S, Sunkara A, DeAngelis GC, Angelaki DE: **Perceptual learning reduces interneuronal correlations in macaque visual cortex.** *Neuron* 2011, **71**(4):750-761 <http://dx.doi.org/10.1016/j.neuron.2011.06.015>.
138. Sanayei M, Chen X, Chicharro D, Distler C, Panzeri S, Thiele A: **Perceptual learning of fine contrast discrimination changes neuronal tuning and population coding in macaque V4.** *Nat Commun* 2018, **9**(1):4238 <http://dx.doi.org/10.1038/s41467-018-06698-w>.
139. Ni AM, Ruff DA, Alberts JJ, Symmonds J, Cohen MR: **Learning and attention reveal a general relationship between population activity and behavior.** *Science* 2018, **359**(6374):463-465 <http://dx.doi.org/10.1126/science.aao0284>.
140. Yang T, Maunsell JHR: **The effect of perceptual learning on neuronal responses in monkey visual area V4.** *J Neurosci* 2004, **24**(7):1617-1626 <http://dx.doi.org/10.1523/JNEUROSCI.4442-03.2004>.
141. Adab HZ, Vogels R: **Practicing coarse orientation discrimination improves orientation signals in macaque cortical area v4.** *Curr Biol* 2011, **21**(19):1661-1666 <http://dx.doi.org/10.1016/j.cub.2011.08.037>.
142. Chen N, Cai P, Zhou T, Thompson B, Fang F: **Perceptual learning modifies the functional specializations of visual cortical areas.** *Proc Natl Acad Sci U S A* 2016, **113**(20):5724-5729 <http://dx.doi.org/10.1073/pnas.1524160113>.
 Using TMS and fMRI imaging techniques, the authors showed that both V3A and MT+ have causal contributions to motion processing. Importantly, after training, the role of V3A in processing noisy motion signals strengthened while the role of MT+ vanished indicating that perceptual learning alters representations across cortical areas.
143. Hua T, Bao P, Huang C-B, Wang Z, Xu J, Zhou Y, Lu Z-L: **Perceptual learning improves contrast sensitivity of V1 neurons in cats.** *Curr Biol* 2010, **20**(10):887-894 <http://dx.doi.org/10.1016/j.cub.2010.03.066>.
144. Yan Y, Rasch MJ, Chen M, Xiang X, Huang M, Wu S, Li W: **Perceptual training continuously refines neuronal population codes in primary visual cortex.** *Nat Neurosci* 2014, **17**(10):1380-1387 <http://dx.doi.org/10.1038/nn.3805>.
145. Goltstein PM, Coffey EBJ, Roelfsema PR, Pennartz CMA: **In vivo two-photon Ca2+ imaging reveals selective reward effects on stimulus-specific assemblies in mouse visual cortex.** *J Neurosci* 2013, **33**(28):11540-11555 <http://dx.doi.org/10.1523/JNEUROSCI.1341-12.2013>.
146. Yu Q, Zhang P, Qiu J, Fang F: **Perceptual learning of contrast detection in the human lateral geniculate nucleus.** *Curr Biol* 2016, **26**(23):3176-3182 <http://dx.doi.org/10.1016/j.cub.2016.09.034>.
147. Rescorla RA: **Pavlovian conditioning and its proper control procedures.** *Psychol Rev* 1967, **74**(1):71-80 Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/5341445>.
148. Amano K, Shibata K, Kawato M, Sasaki Y, Watanabe T: **Learning to associate orientation with color in early visual areas by associative decoded fMRI neurofeedback.** *Curr Biol* 2016, **26**(14):1861-1866 <http://dx.doi.org/10.1016/j.cub.2016.05.014>.
149. Rosenthal CR, Kennard C, Soto D: **Visuospatial sequence learning without seeing.** *PLoS ONE* 2010, **5**(7):e11906 <http://dx.doi.org/10.1371/journal.pone.0011906>.