

How does learning and memory shape perceptual development in infancy?

Lauren L. Emberson*

Department of Psychology, Princeton University, Princeton, NJ, United States

*Corresponding author: e-mail address: lauren.emberson@princeton.edu

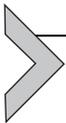
Contents

1. Introduction	130
1.1 Incorporating learning and memory into perceptual development	132
2. Feed-forward/bottom-up models of perceptual development	134
2.1 Models of visual development are predominantly and historically bottom-up	136
2.2 Commitments of bottom-up models	137
3. Challenges to an exclusively bottom-up model	140
4. A feedback/top-down model of perceptual development	145
4.1 Commitments of a top-down model of perceptual development	147
4.2 Evidence for top-down perceptual development in infancy	148
5. Comparison to related models	150
5.1 Scott, Markant and colleagues: Top-down attention and face perception	150
5.2 Johnson, Westermann, and Mareschal: Interactive specialization and connectionist models of object categorization	151
5.3 Frost and colleagues: Statistical learning in adults	152
6. Moving forward	153
6.1 A demonstration of fast, task/context-specific changes in infant perception	153
6.2 Interaction with the development of long-range neural connectivity	153
6.3 Determining the long-term impact of rapid changes in perception	154
6.4 Differentiate the sources of top-down modulation	155
6.5 Why do developmental changes happen when they do?	155
7. Conclusion	156
References	156

Abstract

There is a rapid and substantial development in perceptual abilities, in general, and vision, in particular, in the first year of life after birth (0–12 months). At the same time that these changes in perception are occurring, it is now known that young infants are also engaging in sophisticated types of learning and memory suggesting early

competency in non-perceptual abilities and availability of neural systems located higher in the cortical hierarchy than perception. Thus, there is a developmental co-occurrence between early visual development and an infant's sophisticated learning and memory abilities. Along with extensive research in adults indicating a role of memory in vision, this developmental co-occurrence suggests that an infant's learning and memory abilities could play a role in supporting and shaping perceptual development. However, the dominant model of perceptual development is exclusively feed-forward/bottom-up and does not include the feedback neuro-anatomical connections that would allow sophisticated learning and memory systems to directly shape perception in a top-down fashion. Thus, the first step toward understanding whether and how learning and memory is involved in perceptual development is to either accept this exclusively feed-forward/bottom-up model or to adopt a top-down model that incorporates feedback from learning and memory systems. Focusing on visual development, the differential commitments of these two types of models of perceptual development are outlined. The current evidence for top-down models is presented and a series of future directions are articulated.



1. Introduction

Humans are born with highly immature visual systems. Even though a newborn infant can open her eyes and look around, her brain is unable to effectively process this barrage of new sensory input. Visual development has been an important topic in developmental psychology for many decades (Aslin & Smith, 1988; Gibson, 1969) with a focus on changes in visual capacities starting from birth, when an infant first receives rich visual input, and continuing through their first year (i.e., 0–12 months). Indeed, rapid changes in vision have been documented during this developmental period with changes occurring in both low-level vision (e.g., spatial acuity) and high-level vision (e.g., face perception, see [Arteberry & Kellman, 2016](#) for recent comprehensive review of developmental changes that have been observed in vision and other perceptual modalities as well). While the changes in perception themselves are important to document, in this “what happens when?” approach less is known about the developmental mechanisms that support these changes. Indeed, it is these mechanisms of development that are the focus of this article. To borrow from [Aslin and Smith \(1988\)](#), the goal of the current chapter is to “address a recurrent theme in perceptual development: What is the mechanism that propels development?” (p. 441).

This chapter considers the question “how does learning and memory shape perceptual development?” and brings the latest in our understanding

of early development to bear. This is a mechanistic question considering whether and how learning and memory is part of the mechanism by which perception changes early in life. The phrasing of this question is significant for two reasons. First, despite the fact that perceptual development is characterized by a great deal of change in capacities and many of these changes are believed to be precipitated by sensory experience, the role of learning and memory has been rarely considered in relation to perceptual development. There is an important historical dissociation between the study of *learning and memory* and *perceptual learning and development*. Indeed, the concept of *learning and memory* refers to more sophisticated types of changes in brain and behavior than are believed capable by perceptual systems alone. Examples of phenomena that are considered part of *learning and memory* are the storage and subsequent ability to recall a particular complex event like meeting a person at a particular time and place, the higher-order patterns in a sequence, and different sorts of conditioning (e.g., learning that a tone predicts a shock). Importantly, all of these learning and memory phenomena involve cortices and systems located beyond the perceptual system of the stimuli used to elicit learning (e.g., visual sequence learning does not just depend on the visual system: [Curran, 1997](#); conditioning with a tone does not only depend on the auditory system: [Fanselow & Ledoux, 1999](#)). Indeed, the dissociation in the study of changes in perceptual learning and conditioning (i.e., the belief that changes in perception and conditioning behaviors were not identical processes) is historically important in the field of psychology at the end of the behaviorist era and helped to form the basis of the field of perceptual learning and development ([Gibson, 1969](#)).

The second reason the phrasing of the above question is significant is because it is only relatively recently that developmentalists have recognized that infants are capable of sophisticated types of learning and memory. This belief that many sophisticated types of learning and memory were not available to young infants was based, in part, on the existence of the phenomenon of infantile amnesia (i.e., where we do not have access to memories from our infancy as adults) and the substantial changes in explicit memory that are seen throughout childhood ([Bauer, 2004](#)). However, with increasingly sensitive methods of investigating infant behavior as well as the introduction of infant-appropriate neuroimaging techniques (e.g., functional near-infrared spectroscopy, fNIRS: a light-based method for recording hemodynamic changes in the surface of the cortex: [Aslin, Shukla, & Emberson, 2015](#)), it has become increasingly clear that infants rapidly learn from their experiences and can even retain memories over substantial

periods of time and use cortices beyond their perceptual systems to do so. Three examples to illustrate: Saffran, Aslin, and Newport (1996) demonstrated that infants are sensitive to higher-order statistical properties of their auditory input (e.g., learning to segment multisyllabic words from a continuous stream of speech so that bi-ta-ku is a word but ku-pu-do is not) and sparked the field of statistical learning (see also Aslin, Saffran, & Newport, 1998 for a key follow-up; early visual statistical learning: temporal, Kirkham, Slemmer, & Johnson, 2002; spatial, Fiser & Aslin, 2002, spatio-temporal: Tummeltshammer & Kirkham, 2013). The work of the late Carolyn Rovee-Collier helped to demonstrate that infants are able to form conjunctive memories (e.g., between a context and an action); these memories are sensitive to reactivation and can be retained for weeks and months (Rovee-Collier, 1997). While it is controversial whether this type of memory is explicit and/or mediated by the medial temporal lobe as claimed by Rovee-Collier, there are convergent findings that relational memory, which is found to be impaired in MTL-amnesiacs (Hannula, Ryan, Tranel, & Cohen, 2007), is available in infancy (Richmond & Nelson, 2009). Finally, developmental neuroimaging techniques have revealed that infants are employing their prefrontal cortices to support learning starting early in infancy (Emberson, Cannon, Palmeri, Richards, & Aslin, 2017; Nakano, Watanabe, Homae, & Taga, 2009; Werchan, Collins, Frank, & Amso, 2016). Importantly, all these demonstrations of sophisticated types of learning and memory are occurring at the same time (i.e., before 12 months) as the dramatic changes in visual perception that are occurring in the first year of life. Thus, infants are capable of sophisticated types of learning and memory starting early in life, and these types of learning and memory developmentally co-occur with key phases of visual development.

1.1 Incorporating learning and memory into perceptual development

Since infants are engaging in sophisticated types of learning and memory during the same developmental period when they are experiencing rapid changes in their visual development (i.e., before 12 months of age), it is essential to determine whether learning and memory is influencing experience-based changes in vision, and perception more broadly, early in life. With this in mind, the goal of this article is to present two opposing types of theories or models of perceptual development. One allows the impact of these learning and memory abilities directly on perception starting early in

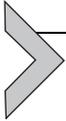
life and the other type does not. The key difference between these two types of models is the availability of neural connections from cortices supporting learning and memory (frontal cortex, MTL, striatum) and perceptual systems. These neural connections are feedback connections: For example, the striatum cannot influence the feed-forward input into the visual system; instead, it must influence the visual system through feedback connections. For this reason, we consider these two contrasting models as *exclusively feed-forward/bottom-up* and *feedback/top-down*.

Feed-forward/bottom-up models of perceptual development have been dominant historically and are consistent with the view that perceptual learning and development is separate from the sophisticated learning and memory abilities that infants possess. For example, [Aslin and Smith \(1988\)](#) argue that early vision is constrained by limitations in the sensitivity of the perception system (e.g., reductions in visual spatial acuity) and, when these low-level constraints are relieved later in the first year of life, this allows developmental changes in the representation of perceptual input (e.g., objects or forms) that emerge later in the first postnatal year (e.g., 6–12 months). Then, heading into childhood, one sees the emergence of more cognitive abilities (e.g., task-based attention) to support yet further developments in perception. Thus, low-level perception develops, then higher-level perception develops, then more cognitive factors impact perception after these initial stages are over. They do not consider the role of memory or familiarity on visual development. Since low-level perception, high-level perception, and cognition are considered along a progression from earlier to later in the cognitive or cortical hierarchy, this is an exclusively feed-forward/bottom-up model of perceptual development. Despite their historical importance and early usefulness, in recent years, the field has been questioning these exclusively feed-forward/bottom-up models of perceptual development.

By contrast, feedback/top-down models of perceptual development will allow more influence of not only learning and memory, but other cognitive factors such as language, on perceptual development. If these abilities are found to influence perception, perceptual developmental cannot be exclusively feed-forward and bottom-up but must incorporate feedback neural connections to allow for the top-down influence of perception starting early in life.

There is not sufficient evidence to argue for one or the other of these models at this time. Thus, the goal of this article is to consider these types of models as clearly as possible in order to facilitate future work. To this end,

I will outline the different commitments of these models, highlight some of the benefits and limitations of each and present some of the evidence. The ultimate motivation for this article is not to argue for or against a particular view but to move the field forward in the investigation of the mechanisms supporting perceptual development.



2. Feed-forward/bottom-up models of perceptual development

The historically-dominant model of visual development is one that is exclusively feed-forward or bottom-up, as outlined above. As the goal of this chapter is to be explicit about these types of models and their commitments or predictions, we start by defining why these models are feed-forward or bottom-up.

For ease of explanation, I will divide the hierarchy of visual processing into three parts (Fig. 1): lower-level perception (e.g., sensitivity to sensory features like color, luminance, frequency), higher-level perception (e.g., representations for objects), and cognitive systems (systems located neuro-anatomically after or beyond the perceptual cortex that are believed to mediate more cognitive and less perceptual functions, e.g., sophisticated types of learning and memory, endogenous attention, reward, executive function systems).

The connections ascending this hierarchy (lower-level perception, higher-level perception, cognitive) are feed-forward (neuro-anatomically speaking), and this hierarchy also represents an increase in functional complexity from less to more cognitive. The connections descending this hierarchy are feedback connections, neuro-anatomically (e.g., LOC can only influence V1 through feedback neural connections, similarly for the medial temporal lobe, MTL) and also represent a transfer of information from more to less cognitively-complex regions of the brain (e.g., knowledge of objects which is representational can influence sensitivity to lines or contours which is thought to be more sensory). A bottom-up model only allows connections between adjacent levels, which limits which systems can influence perception to those lower in the hierarchy (or sensory input). However, top-down models allow the influence of more systems on perception starting early in life. Thus, top-down models allow more regions of the brain to be involved in perceptual development. Influence between perceptual modalities is also considered top-down as the cortical visual system is only influenced by auditory input through feedback neural connections.

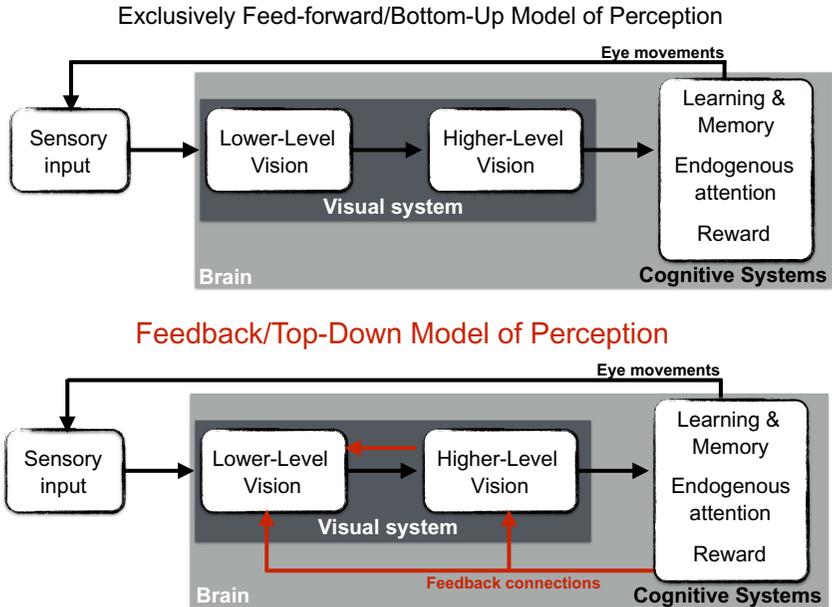


Fig. 1 A simplified boxology of a cortical hierarchy starting from sensory input to lower-level vision (e.g., sensitivity to sensory features such as contrast, color), higher-level vision (e.g., object and face perception) and cognitive systems (e.g., learning and memory as mediated by cortices outside perceptual systems like the medial temporal lobe, MTL, basal ganglia or striatum and frontal cortex; attention mediated by the frontoparietal network; reward from amygdala and ventral striatum). Red arrows denote what are defined as feedback neural connections. Black arrows denote either what are defined as feedforward neural connections or ways that sensory input and perception can change that are not dependent on feedback connections (e.g., changes in sensory input; modulation of eye movements directly from cognitive or attentional systems).

It goes without saying that this division of the hierarchy of visual processing is highly simplified, but the focus on these models is the neural connections and flow of information, and this division allows us to consider relatively few and cognitively significant connections. By contrast, some conceptions of feedback in perception consider connections between particularly layers of the cortex (Friston, 2005; Rao & Ballard, 1999) and others consider top-down to be conscious and volitional changes in perception (Firestone & Scholl, 2016). This division of the hierarchy proposed here, however, charts an intermediate course between these two extremes and allows the focus of just a few types of feedback connections in relation to a broadly defined functional and anatomical hierarchy and could potentially be investigated early in development using behavioral and neuroimaging techniques.

2.1 Models of visual development are predominantly and historically bottom-up

While the field of perceptual development has not historically created explicit models, the models classically employed have been bottom-up. Eleanor Gibson (e.g., 1969) excluded any cognitive influences on perceptual development and, even more extremely, also rejected constructivist approaches such that *representations*, arguably part of higher-level vision, were not included. Instead, the infant or learner discovered invariants in their sensory input that occur over time and space that allow them to see the world directly. Gibsonian perceptual development has a largely flat hierarchy where perception development occurs in isolation from other cognitive abilities.

Aslin and Smith's (1988) comprehensive review of perceptual development explicitly lays out a very similar hierarchy to what is presented here with sensory primitives (corresponding to low-level perception), perceptual representations, and cognitive and language influences (see also summary of this model above). When considering visual development, the authors explicitly state that the hierarchy is bottom-up where the first phase of development is constrained to sensory primitives and, when sensory primitives are sufficiently developed, that permits the subsequent development of perceptual representations. Once there has been adequate development of perceptual representations, the authors argue that in early childhood there can be other (cognitive) influences on vision.

Turning now to more contemporary models, a major shift in the field of perceptual development in the 21st century was the discovery of perceptual narrowing as a visual and, arguably, domain-general, perceptual developmental phenomenon. Very young infants have the capacity to equally (but poorly) perceive many types of faces (e.g., human and monkey faces, faces of all races) and many types of speech sounds (i.e., those of many different languages). Infants differentially develop perceptual abilities for the faces and speech sounds that are present in their environment compared to ones that are not present in their environment (faces: Kelly et al., 2007; Pascalis et al., 2005; speech: Werker & Tees, 1984). This phenomenon of perceptual narrowing presents a compelling case that an infant's experience shapes their perceptual development in a way that supports the emergence of more sophisticated perceptual abilities and has been the focus of the field of perceptual development for most of the 21st century.

Theorizing about how perceptual narrowing occurs has focused largely on the debate about whether it arises from sensory experience or is mediated by biologically-determined critical periods. This particular debate is beyond the scope of this review. But, focusing on the models of how experience supports perceptual narrowing, both early and more recent reviews have argued that it is mere sensory experience that drives this phenomenon as opposed to the involvement of learning and memory or other more cognitive systems. [Scott, Pascalis, and Nelson \(2007\)](#) note that the development of specialization “corresponds to improved perceptual discrimination for stimuli predominant in the environment relative to declining perceptual discrimination for stimuli not present in the environment” (p. 200) and point to synaptic pruning and Hebbian learning as potential neural mechanisms supporting perceptual narrowing. Through these mechanisms, “as perceptual experience with human faces increases, the strength of the neural circuit responding to human faces is strengthened” (p. 201). A more recent review suggests a similar mechanism: [Maurer and Werker \(2014, p. 171\)](#) suggest that it is the “simple accrual of experience” that results in perceptual narrowing. Both of these models point to differences in an infant’s sensory experience as the cause of perceptual narrowing (e.g., Caucasian infants predominantly see Caucasian faces). These explanations speak to the view that perceptual development only requires the interaction of sensory input and a developing perceptual system that are necessary to result in perceptual development with development occurring at the level of relevant representation (i.e., face or speech) rather than involving the interaction from cognitive systems such as learning and memory located beyond the perceptual systems that may be activated by this differential experience.

2.2 Commitments of bottom-up models

The explicit architecture of the bottom-up models results in particular commitments for how perceptual development should proceed. These commitments can be operationalized as testable hypotheses and will be considered in relation to the recent empirical findings. [Table 1](#) summarizes these commitments of exclusively feed-forward/bottom-up models.

A key commitment of an exclusively feed-forward/bottom-up model is that perceptual development is entirely stimulus-driven. In other words, what sensory input an infant receives drives what develops. This constraint

Table 1 Differential commitments of feed-forward/bottom-up and feedback/top-down models of perceptual development.

Feed-forward/bottom-up	Feedback/top-down
<i>Unidirectional</i> : feed-forward flow of information only from the lowest levels of perception to cognitive systems	<i>Bidirectional</i> : feedback creates a bidirectional flow of information across the levels of perception and cognition
The same sensory experience will result in the same changes in perception	The same sensory experience can yield different changes in perception if and only if there is a change in structural/cognitive/task context
<i>Passive</i> : no engagement of higher-level perception or cognitive systems is necessary	<i>Active</i> : certain types of experience engage higher-level perceptual or cognitive systems which then feedback to change perception
Development proceeds <i>sequentially</i> from lower to higher-level of perception (and then cognition)	Development can be <i>simultaneous</i> across levels of the hierarchy
<i>Slow</i> changes in perception (i.e., requiring a lot of sensory experience) arise from mechanisms like Hebbian learning and synaptic pruning. These are <i>structural</i> changes in the brain	Potential for <i>rapid</i> changes (i.e., from a brief exposure) in perception through top-down modulation of perceptual systems. These changes are <i>not structural</i>
<i>Robust but inflexible</i> changes in perception that are <i>constant across contexts</i>	<i>Flexible</i> changes in perception that can <i>change across contexts</i>

results in the strong prediction that if infants receive the same sensory input that they will have the same development outcomes; For example, if two infants see the same distribution of faces, they will have the same face discrimination abilities.

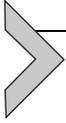
Feed-forward/bottom-up models would result in development that is, to some extent, *sequential* from lower- to higher-levels of perception. Drawing from [Aslin and Smith \(1988\)](#), sensory primitives must be adequately developed before development at the higher-representational level is possible and which then facilitate cognitive abilities later on development. Indeed, sequential development is a common feature in experience-based models of language development. Since there is no strong evidence of sequential development (e.g., speech perception develops before lexical representations), the requirement of sequential development has been recognized as a key weakness in these models (see speech sounds and word learning or lexical development: [Saffran & Thiessen, 2007](#); [Swingley, 2008](#)).

Bottom-up models predict that learning and development will be *local* and arise *slowly*. Bottom-up models invoke mechanisms like Hebbian learning and pruning (e.g., [Scott et al., 2007](#)) which can strengthen representations of stimuli that an infant experiences. While it is not worked out how these mechanisms would give rise to perceptual learning, these mechanisms result in structural changes in the cortex through modifying the strength and/or distribution of synapses in the cortex. Since they are working at the level of individual synapses, they will arise from highly local learning that is specific to the level of the relevant representations (i.e., Hebbian learning of faces will occur in regions of the cortex that are selective for faces). Moreover, since Hebbian learning and pruning are mechanisms that alter the structure of the cortex, these types of changes will arise over long periods of time and a great deal of experience. For example, there is an important developmental period of pruning that occurs over many years (and even decades) starting early in life ([Webb, Monk, & Nelson, 2001](#)). Recent advances in imaging technologies (e.g., two photon microscopy) have allowed researchers the ability to see how synapses change in response to experience in animal models. While this work has revealed that synapses are very dynamic and change their structure as a result of experiences over days (as opposed to years), these synaptic changes are produced through extreme variations in sensory input and usual complete deprivation of input (e.g., from a limb: [Zuo, Yang, Kwon, & Gan, 2005](#); from one or both eyes; see [Zhou, Lai, & Gan, 2017](#)). Given that the differences in sensory experience arising from seeing only Caucasian faces as opposed to seeing a wider variety of faces are very different from complete visual deprivation, it is unlikely that synaptic changes arising from face experience would modify synapses in order to affect experience even on the scale of days. Thus, the mechanisms proposed by bottom-up models of perceptual development will result in highly local and slow changes in perception.

The information flow of feed-forward/bottom-up models is unidirectional from the lowest levels of the visual hierarchy moving upward toward cognitive systems without any feedback or top-down connections. Without these feedback connections, perceptual development is passive: It occurs without any cognitive engagement from the infant, and without the interaction of higher-level to lower-level perception. In an extreme expression of this idea, perceptual experience could support perceptual development without any region in the brain beyond the perceptual cortices.

Finally, bottom-up models will result in changes in perception that are quite robust but also inflexible. Since structural changes in the cortex are invoked (e.g., Hebbian learning and pruning), sensory experience will result

in perceptual changes that have resulted from structural changes in the cortex and, therefore, will affect perception regardless of task context. For example, if perceptual narrowing results in a structural strengthening of representations for the faces that an infant has seen, the perceptual benefit of this experience should be present regardless of the context in which that face occurs.



3. Challenges to an exclusively bottom-up model

There are several crucial behavioral phenomena that cannot be explained with an exclusively feed-forward or bottom-up model of perceptual development. As outlined above, a bottom-up model is committed to the fact that if sensory input is the same, then the perceptual outcomes will also be the same. That is, if two infants see the same faces, then they will develop the same perceptual abilities. There are findings with both face and object perception in infancy that challenge this assumption. Another commitment of the exclusively feed-forward/bottom-up model is that changes in perception occur slowly over a lot of experience. However, there have been empirical demonstrations that experience with novel objects over short periods (e.g., less than a minute) can support changes in perception. These rapid changes cannot be readily explained using an exclusively feed-forward/bottom-up model of perception.

One important group of studies that challenge an exclusively feed-forward/bottom-up model conclude that higher-level object representations or perceptual categories can influence lower-level perception (e.g., scene segmentation, Gestalt perceptual grouping) in infants as young as 3–4 months. Considering the hierarchy in Fig. 1, object representations or perceptual categories are part of higher-level vision, and scene segmentation is considered part of low-level vision. For perceptual categories to influence scene segmentation, infants must be employing feedback connections from high- to low-level vision. Moreover, in the majority of these studies, changes in perception arise from a minute or two of experience in a single session in the lab. Thus, these changes are very rapid with changes in perception.

In one line of investigation, Quinn and colleagues examined whether young infants can override their bottom-up perception (lower-level vision) based on prior experience with a particular object category (higher-level vision). Fig. 2 (top panel) presents a schematic of the major findings from this work. First, Quinn and colleagues established that an infant's

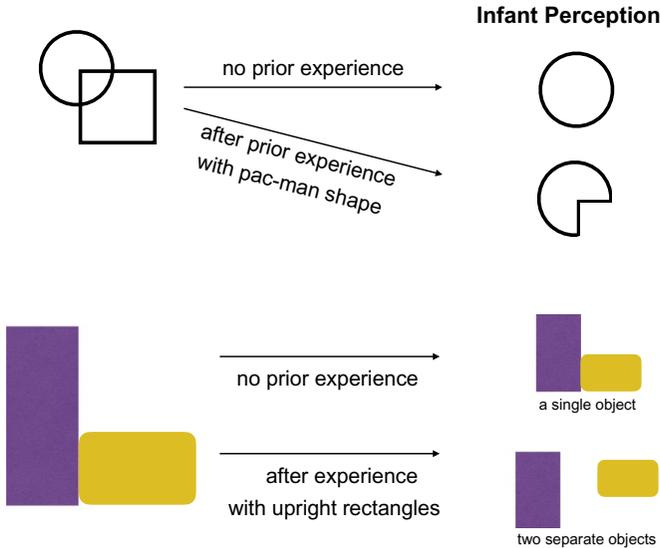


Fig. 2 Schematic of findings from Quinn and colleagues (top panel) and Needham and colleagues (bottom panel) that challenge an exclusively feed-forward/bottom-up model. In both this series of findings, infants make rapid changes in their perception of visual scenes as a result of variable experience. Left side of the figure presents the type of scene that infants were presented with and the right side of the figure presents key features of the infant’s perception of the scene based on behavioral findings (e.g., looking time or looking preference). For key papers in these seminal experiments, see [Needham, Dueker, and Lockhead \(2005\)](#) and [Quinn and Schyns \(2003\)](#).

(3–4 months old) initial perception of two overlapping shapes (an example is depicted on the left of the top panel of [Fig. 2](#)) is that of a complete circle. Across a series of experiments, Quinn and colleagues convincingly show that prior experience with the “pac-man” interpretation of this shape (i.e., not a circle) leads infants to segment this scene differently and to perceive the scene as containing the “pac-man” and not the complete circle. This is evidence that (a) infants retain experience with the “pac-man” and form an object representation for this shape and, crucially, (b) infants use this object representation in their lower-level vision to interpret a novel visual scene. Quinn and colleagues interpret this as evidence of a top-down influence in vision ([Quinn & Schyns, 2003](#); [Quinn, Schyns, & Goldstone, 2006](#)). Specifically, the authors conclude that “perceptual units formed during concept acquisition can be entered into a perceptual system’s working ‘featural’ vocabulary and be used for subsequent object segmentation” ([Quinn et al., 2006](#), pp. 120–121). Why can these results not be explained

by priming or bottom-up explanations? First, the experience infants have with each shape is controlled suggesting that it cannot be priming. Second, infants receive a variety of experience with the pac-man or circle shapes such that the shapes occur with a numerous occluding objects. With the variability of experience, it is less likely that low-level vision (sensitive to retinal position and other low-level features of a scene) could result in the change in perception observed by Quinn and colleagues. Thus, this series of behavioral findings is a challenge to an exclusively bottom-up model in two ways. First, it is evidence that object representations in higher-level vision are able to influence lower-level vision in very young infants. Second, these changes in perception occur rapidly after a minute or two of experience.

A convergent series of findings is presented by Needham and colleagues. In these studies, infants' prior experience with multiple similar exemplars of a shape bias their segmentation of a visual scene. [Fig. 2](#), bottom panel, presents a schematic of the basic findings. Infants are familiarized with an upright rectangle abutting another object. Infants without any prior experience interpret these objects are being connected (one object). However, with a short amount of experience with other exemplars of the upright rectangle (e.g., three exemplars with different colors and patterns), they change their interpretation to consider the upright rectangle to be a separate object ([Needham & Baillargeon, 1998](#); [Needham et al., 2005](#); [Needham & Modi, 1999](#)). There are two key features to these findings, similar to Quinn and colleagues' findings. First, these changes in perception happen as a result of variable experience with the object again suggesting the involvement of higher-level vision and the creation of object representation or perceptual categories (with the increased variability in these series of studies the authors conclude that category knowledge is involved in these perceptual effects). Second, a brief period of exposure to these variable exemplars is sufficient to have infants change their segmentation of a visual scene suggesting that young infants are capable of rapid changes in their perception as a result of experience.

Moving onto face perception and perceptual narrowing, a number of studies have found evidence that receiving the same sensory input does not necessarily result in the same perceptual outcomes. To highlight, two prominent and complementary examples, [Scott and Monesson \(2009\)](#) employed a training study to provide infants, starting at 6 months of age and occurring over 3 months, with the same visual experience with monkey faces (via a storybook read by their parents), while the task context of the faces or environmental structure of the faces was different. Specifically,

infants received variation in the linguistic context: One group of infants simply saw the pictures, another group saw the pictures labeled with the category label “monkey,” and the last group saw the pictures but each monkey had their own name (i.e., the individuation group). According to an exclusively feed-forward/bottom-up model of perceptual development, these infants should all have the same perceptual outcomes as, based on the assumption of the task, they all viewed the monkey faces equally during the same period of time. Yet, only infants in the individuation group, where each monkey had its own name, maintained their perceptual discrimination of monkey faces. In other words, having parents give proper names for each monkey picture affected infants’ visual development beyond their visual experience. While this study has been highly influential, there is an important limitation to note which is that the sensory input was not quantified between the groups, and it is assumed by the researchers that infants were read the book for equal amounts across all the different groups. However, given that there is more linguistic detail and potentially a more interesting story in the individuation book, it is possible that this group also received more sensory experience of the monkey faces.

Another study, that rigorously controls for sensory input, argues that attentional progressing is an integral part of an infant’s face perception development. While an infant may receive the same bottom-up sensory input across stimulus types, they may differentially engage higher-level systems such as endogenous attention depending on prior history, task context, etc. If there are systematic differences in an infant’s engagement of these higher-level systems across stimulus types, it may lead to developmental changes in perceptual processing of these stimuli. The majority of studies allow for just such a differential engagement as they present different stimuli for infants to orient to and engage with. These studies are able to control for the duration of sensory input but do not have a way to quantify or equate the engagement of endogenous attention across stimulus types (e.g., [Kelly et al., 2007](#); [Pascalis et al., 2005](#)). By contrast, [Markant, Oakes, and Amso \(2016\)](#) used an ingenious attentional manipulation, relying on inhibition of return, to control the engagement of attentional systems of infants across face categories. Inhibition of return is a phenomenon where a reflexive or automatic shift of attention to a particular spatial location (e.g., orienting toward a sudden flash on a screen) is followed by an internal inhibition of attentional processing to input from that previously used spatial location and enhanced selective-attention to a new location. After the initial cue was presented followed by a brief delay (600 ms), Markant and colleagues presented faces

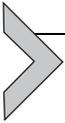
to infants either in the location where infants had previously been cued and was now attentionally inhibited or to the previously uncued location that was now attentionally enhanced. This manipulation allowed the researchers to equate the attentional engagement of infants across face types because an infant's attention is being manipulated by the initial cue (which is identical across face types). By controlling for attentional engagement, Markant and colleagues found that when attention is equated across face types, differences in face perception across own and other-race faces disappear. Of course, not all of face perception development is explained by attention. ERP studies have shown that early, perceptual ERP components (e.g., the N290, the infant analog to the N170 in adults) show differential development for experienced vs. non-experienced faces (e.g., [Scott, Shannon, & Nelson, 2006](#)). However, [Markant et al. \(2016\)](#) provide some evidence that the other-race effect may not be exclusively driven by declines in perceptual sensitivity or perceptual representations driven by bottom-up input but may be driven in part by the differential engagement of high-level systems such as attention.

In addition to these intriguing findings with infants, the understanding of the adult perceptual system has shifted in the last decades to suggest that adult perception is highly permeable to feedback or top-down information. Take, for example, the primary visual cortex (V1): A region that has been viewed as the canonical bottom-up perceptual region as it is the first area in the cortex to receive incoming visual information from the environment. Yet even this region of the brain receives 5–8 times more feedback or top-down information from the rest of the brain than it receives in bottom-up input from the environment (45% vs. 6–9%). Moreover, it is believed that this top-down information enables perceptual systems to be effective by attuning themselves to the environment ([Gilbert & Li, 2013](#)).

In the field of perceptual learning, it has been demonstrated in a number of ways that even highly local and specific changes at the lowest levels of vision arise through the influence of higher levels of the visual system and what are being considered more cognitive factors. Perceptual learning is affected by other learning and memory processes (e.g., selective attention, [Ahissar & Hochstein, 1993](#); reward [Seitz, Nanez, Holloway, Tsushima, & Watanabe, 2006](#)). Moreover, while bottom-up changes would result in slow changes in perception over a great deal with experience, we now know that there are rapid changes in perceptual abilities as a result of relatively small amount of perceptual experience that indicate top-down mechanisms are operating in adults ([Sagi & Tanne, 1994](#)) and that it has been proposed that changes occur from higher- to lower-levels in the visual hierarchy

(see Ahissar & Hochstein, 2004). Baldassarre et al. (2012) found that perceptual learning was predicted in part by individual variation in fronto-occipital connectivity in a resting state scan taken before training suggesting that the frontal lobe is involved in perceptual learning. Given these findings, perceptual learning has at least some dependence on cortices and systems that exist beyond the early regions of the visual system that are modified through perceptual learning.

Of course, even though current views of the mature visual system incorporate feedback connections during perceptual learning, this does not mean that infants employ the same mechanisms during visual development. As reviewed in Advantages of Bottom-Up Models, there are neuro-anatomical constraints present in the infant brain that suggest they may be particularly and selectively limited in interactions across brain regions and the use of feedback connections. However, if perceptual development is driven through an exclusively feed-forward/bottom-up model, it is beholden on developmental scientists to explain how a highly top-down, bidirectional adult perceptual system could arise from one that develops in a purely bottom-up fashion.



4. A feedback/top-down model of perceptual development

Following these challenges to an exclusively feed-forward/bottom-up model of perceptual development, here we consider alternative models that incorporate feedback connections. According to these models, starting early in life, experience can shape perceptual development in a top-down fashion using feedback connections. Experience with statistics and structures in the environment allows the engagement of higher-level cognitive systems such as the learning and memory systems located beyond the perceptual cortex. These systems engage in sophisticated learning to uncover the complex and higher-order patterns present in the environment accumulated over variable experiences. These higher-level systems uncover and encode types of patterns that are not learnable by perceptual systems alone. Once these patterns are learned, learning and memory systems can then use feedback connections to shape an infant's perception in a top-down fashion and guide perceptual development.

The term *top-down* has been used in a number of different ways from volitional control of perception (Firestone & Scholl, 2016) to feedback between adjacent cortical regions of the visual system (Clark, 2013). Here,

I chart an intermediate course between these two extremes. Building from the simplified three-part hierarchy presented in Fig. 1, information can be considered to be top-down if it connects two parts of this hierarchy using feedback connections (Fig. 1). For example, feedback from demonstrably cognitive processes such as sophisticated learning and memory or reward systems or from other perceptual modalities to vision at either level would be considered top-down (e.g., Emberson & Amso, 2012). Similarly, if higher-level representations within a perceptual system feed back to lower-level representations, this would also be considered top-down (e.g., Lupyan & Spivey, 2008).

Note that the definition of top-down is distinct from the definition of top-down raised by Firestone and Scholl (2016) and their related criticisms of the top-down literature. For example, the findings from (Emberson & Amso, 2012) are considered top-down in the current model as they required the influence of the hippocampus to shape object perception. Yet, they were accepted by Firestone and Scholl in their response to rebuttals (Emberson, 2016; Firestone & Scholl, 2016). Thus, while the provocative editorial by Firestone and Scholl (2016) has raised important criticisms regarding the validity of certain top-down effects in adult perception, this proposal is largely tangential to the current proposal. In particular, the current proposal is distinct as it focuses on learning, neural evidence and has definition of top-down that does not involve volitional or conscious control of perception.

The current model is also distinct from the numerous challenges to the bottom-up and top-down dichotomy in the domain of attention. In attentional research, bottom-up is considered as salience-driven, and top-down is considered relevant to the current goal. Awh, Belopolsky, and Theeuwes (2012) and Hutchinson and Turk-Browne (2012) both argued that this dichotomy is flawed since reward and memory, respectively, manipulate attention in a way that is neither bottom-up nor top-down. The definitions of top-down and bottom-up as defined by the model presented here are not the same as what is used by the broader field of attention but refer to feed-forward and feedback neural connections (also in Gilbert & Li, 2013; Johnson, 2011). Moreover, it could be that the feedback-based integration of learning and memory systems with perceptual systems partly drives the break down in this attentional dichotomy between bottom-up (salience drive) and top-down (task driven) as perceptual representations are shifted through feedback connections.

4.1 Commitments of a top-down model of perceptual development

Top-down models of perceptual development make a series of different commitments from an exclusively bottom-up model (Table 1). First, the flow of information and systems available to influence perceptual development are much less constrained. While bottom-up changes and feed-forward information are still available in a top-down model, there is an additional feedback from higher- to lower-levels of the visual system and from the myriad of cognitive systems (learning and memory, reward, attention) or systems that would require feedback connections to influence cortical vision systems (e.g., auditory). Thus, a top-down model allows for many more interactions between segments of this hierarchy through the introduction of feedback connections (see Gilbert & Li, 2013 for another paper that equates feedback connections with top-down modulation).

Since a top-down model is much more interactive, the changes and systems that give rise to changes in perception can be more distributed. In other words, while some aspects of perceptual development can arise from highly local changes (e.g., changes in face perception arising from changes locally to the fusiform), they may also depend on regions or systems located higher in the hierarchy (e.g., changes in face perception may arise from the engagement of attention systems likely located in the frontoparietal regions of the brain, see Markant et al., 2016; Markant & Scott, 2018). The dependence of perceptual changes on these other regions is not possible in an exclusively feed-forward/bottom-up model of perceptual development but is possible in a feedback/top-down model. A number of neuroimaging studies in adults have now pointed to the involvement of non-visual cortices in experience-based changes in perception (e.g., involvement of the medial temporal lobe in visual changes: Emberson & Amso, 2012; Stokes, Atherton, Patai, & Nobre, 2012; Turk-Browne, Scholl, Johnson, & Chun, 2010). These studies suggest that while perceptual changes occur as a result of experience, these changes occur in concert with regions located beyond the visual cortex.

Because of these interactions and the influence of feedback neural connections, the same sensory input can result in different developmental changes in perception. While an exclusively feed-forward/bottom-up model is strongly constrained by the sensory input an infant receives, a feedback/top-down model allows that the same sensory input can result in differences in perceptual outcomes *if and only if* there are differences in

the structural/cognitive/task context in which the sensory input is received. Thus, two infants who receive the same sensory experience with faces could have differences in their perceptual development if the context or task of this exposure were different. As outlined above in [Section 2](#), this impact of task has already been demonstrated by [Scott and Monesson \(2009\)](#) where infants received the same experience with monkey faces but differed in the context in which this experience was given. This difference in task (i.e., individuation vs. non-individuation or no labels) corresponded to differences in infant perception.

In addition to the slow and gradual tuning that feed-forward/bottom-up models allow, a top-down model allows that experience can shape perception through rapid and flexible changes. Top-down changes in perception do not have to arise from structural changes in the brain (e.g., pruning and Hebbian learning) that alter the representations or sensitivity to particular representations. Instead, these rapid and flexible changes arise from the modulation of perceptual cortices through feedback connections from other systems. To refer to a canonical example, task-driven or endogenous attention has been found in adults to modulate sensory responses in the visual system rapidly (i.e., it does not require months of training or exposure) and flexibly depending on task (e.g., [Luck, Fan, & Hillyard, 1993](#); [McMains & Kastner, 2011](#)). While task-driven or endogenous attention is a canonical example, flexible changes in vision are in principle able to be gained from many different modulatory systems including learning and memory (e.g., [Peterson, De Gelder, Rapcsak, & Gerhardstein, 2000](#)), reward (e.g., [Serences, 2008](#)), and prediction or expectation (e.g., [Kok, Rahnev, Jehee, Lau, & De Lange, 2012](#); [Summerfield & de Lange, 2014](#)). All of these changes in the visual system can be seen after a very short period of exposure (i.e., are rapid) and are, in some cases, dependent on context.

Even though top-down processes can give rise to rapid changes, mediated through top-down processes, they can still result in lasting changes in perception. [Folstein, Palmeri, Van Gulick, and Gauthier \(2015\)](#) argue that perceptual category learning employs selective attention to expand the perception of certain feature dimensions and these changes both arise rapidly and persist over weeks after only a short period of training.

4.2 Evidence for top-down perceptual development in infancy

In [Section 2](#), several behavioral studies are reviewed that indicate that perceptual development is not only driven by the sensory input that an infant

receives. Beyond these challenges to an exclusively feed-forward/bottom-up model, recent behavioral and neuroimaging studies have established that young infants already have the capacity to employ top-down modulation to change activity in their visual system and that disruptions of this ability are associated with poor developmental outcomes. Moreover, there is some initial evidence from clinical studies that the striatum is involved in visual development.

Emberson, Richards, and Aslin (2015) examined responses in the visual systems of 6-month olds when they predicted a visual stimulus. This study used functional near-infrared spectroscopy (fNIRS), a light-based method of recording the hemodynamic response, the same physiological substrate as fMRI, in the surface of the infant cortex (Aslin et al., 2015). Consistent with the view that top-down mechanisms are available early in development, infants exhibited robust visual cortex responses when a visual stimulus was unexpectedly omitted. In other words, visual expectations can provide feedback to drive perceptual cortex responses in infancy as young as 6 months of age.

Importantly, deficits in top-down modulation are associated with poor developmental outcomes. Premature birth has been found to put infants at-risk for language delays, motor deficits and learning disabilities. Emberson, Boldin, Riccio, Guillet, and Aslin (2017) found that, at 6 months, infant born prematurely (<33 weeks gestation) exhibited deficits in neural signals associated with the prediction of visual stimuli. Finding these deficits in top-down processing very early in development suggests that deficits in prediction give rise to developmental difficulties.

Moreover, an isolated study has suggested that the striatum is more involved in visual system development. Mercuri et al. (1997) investigated 37 infants with perinatal brain lesions to determine the locus of lesion that results in visual impairment as measured using behavioral and electrophysiological measurements. Consistent with the view that the young brain is highly plastic, even infants with bilateral occipital lesions had normal visual function. However, infants with basal ganglia lesions (either associated with occipital lesions or isolated) had abnormal visual function even when assessing vision using non-motor tasks (e.g., VEP, p. F111).

While this study bears follow-up, the finding that the striatum, a region not typically associated with visual function, should be considered in relation to the broader concept of *developmental diaschisis*. Coined in Wang, Kloth, and Badura (2014), developmental diaschisis is a concept where disruptions in a particular brain area during development, such as the basal ganglia in this

case, can affect the development of disparate regions of the brain but not to function in the mature brain. Using this concept, [Wang et al. \(2014\)](#) propose that the cerebellum, a region typically associated with motor function in the mature brain, is involved in the development of non-motor cortical circuits. As part of evidence to support this, they review compelling evidence that abnormal cerebellar development increases an individual's risk for developing autism spectrum disorder (ASD) despite cerebellar dysfunction in adulthood having no such association. Returning to the striatum, [Fuccillo \(2016\)](#) also proposes the striatum to be involved in ASD through developmental diaschisis. [Mercuri et al. \(1997\)](#) suggest that while the basal ganglia is not important for visual function later in life, it is important for early visual development. This type of observation or the existence of developmental diaschisis could not be accommodated in an exclusively bottom-up model of perceptual development.

Together, these findings provide some initial evidence that feedback connections are available to modulate the visual system starting early in infancy ([Emberson et al., 2015](#)), that impairment in this top-down modulation is associated with infants at-risk for developmental delays ([Emberson, Boldin, et al., 2017](#)) and that regions located outside the visual system may be more essential for an infant's visual function than the occipital cortex ([Mercuri et al., 1997](#)).



5. Comparison to related models

The shift toward considering top-down or feedback connections in the context of perceptual development is broader than the current proposal. Here, we consider the relationship between the current proposal and other related proposals.

5.1 Scott, Markant and colleagues: Top-down attention and face perception

The most related proposal to the current one has been put forward by Scott, Markant and colleagues in two papers: [Hadley, Rost, Fava, and Scott \(2014\)](#) and [Markant and Scott \(2018\)](#). These proposals are the first to consider whether top-down attention can play a role in supporting perceptual narrowing of faces in infancy. In particular, [Markant and Scott \(2018\)](#) present an exciting, ground-breaking and concrete model of how early bottom-up attentional biases can interact with later developing top-down attentional biases to give rise to the kind of perceptual selectivity that is seen

at the end of the first postnatal year (e.g., 10–12 months). While there are many points of convergence between the current proposal and the model outlined by Markant and Scott, we highlight two major differences between these proposals.

Scott, Markant and colleagues' model is highly specific to face perception and attentional development. It would not be trivial to simply broaden the [Markant and Scott's \(2018\)](#) proposal to other aspects of vision for two reasons: (1) Their model depends on particular bottom-up attentional biases to faces early in life; (2) They propose that the changes in face perception after 6 months are specifically tied to developmental changes in endogenous attention specific to face perception.

Second, the proposals have different definitions of top-down and the mechanisms of action for top-down attention may be distinct between the proposals. The Scott, Markant and colleagues' proposals rely heavily, though not exclusively, on findings where infants' eye-movements to faces are altered over development as evidence that top-down attention is shaping face perception. While it is the case that sampling different aspects of the visual input results in differences in sensory input, it is not clear whether these subtle, if systematic, changes in eye movements meaningfully change perception (see below) and whether these changes in eye movements arise through use of feedback connections from the anterior attentional system to the visual system. Thus, while Scott, Markant and colleagues' proposals use the term top-down, selective or endogenous attention, it may be that none of these changes rely on feedback neural connections. Indeed, as outlined above, the current proposal uses a concept of top-down that is different from the concepts of top-down and bottom-up attention. Thus, these proposals employ different definitions of top-down and, as such, may be mechanistically independent. It is also possible that there is a great deal of convergence here but additional evidence and clarification about the role of feedback connections would need to be more carefully evaluated in their proposal.

5.2 Johnson, Westermann, and Mareschal: Interactive specialization and connectionist models of object categorization

This proposal owes large debt to the theoretical view of Interactive Specialization (e.g., [Johnson, 2011](#)). However, this important theoretical view has not yet considered feedback or top-down influences: [Johnson \(2011\)](#) highlights the question of “how the specialization of early sensory areas is shaped by top-down feedback” (p. 19) as an important Future Direction

and notes several relevant adult findings but no developmental evidence for top-down feedback.

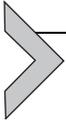
While the theoretical work in this literature has not yet been extended in this direction, one relevant model has been published which suggests that the hippocampus could affect perceptual representations of object categories in development through feedback connections (Westermann & Mareschal, 2012, 2014). In brief, this model has a module characterized by a slow learning rate (cortex) and a fast learning rate (hippocampus) with mutual connections between them. When the fast learning rate module is present, there is more separation between the representations in the slow learning module (cortex) than if it is absent. This model is convergent with the current proposal as connections from the hippocampus to the visual cortex would be feedback (i.e., descending the cortical hierarchy). Moreover, the authors explicitly argue that these changes occur “beyond attending to stimuli differently” (p. 7) and integrate changes in perception.

The differences between that proposal and the current one are threefold. First, there is a clear difference in scope: Westermann and Mareschal (2012, 2014) focus exclusively on the hippocampus and object categorization. Second, it is not clear that the changes found in the cortical module are perceptual. They are certainly representational which is exciting but there are no non-perceptual representations in this module (e.g., semantic) as it only has one layer. Moreover, while the differences between fast and slow learning rates are inspired by neurobiology and have been highly popular in the connectionist modeling literature, it is readily accepted that these models cannot be taken as direct neuroanatomical evidence and convergent research is needed. Indeed, the authors acknowledge that these models “are not intended as models of neuronal processing. Instead, they aim to explain cognitive processing in terms of statistical learning, using only a rough approximation of the actual learning mechanisms of the brain” (p. 370).

5.3 Frost and colleagues: Statistical learning in adults

Another highly interactionist model that proposes the modulation of perceptual systems during learning and experience is the multisensory model of statistical learning proposed by Frost, Armstrong, Siegelman, and Christiansen (2015). This model proposes that regions like the frontal cortex, the hippocampus and the striatum modulate relevant representations in perceptual regions during statistical learning. However, it is unclear whether this model is operational early in development as it is built exclusively from

adult cognitive neuroscientific findings (as virtually no developmental studies exist even in childhood: though see recent work from [Finn, Kharitonova, Holtby, & Sheridan, 2019](#) linking the anatomical structure of the prefrontal cortex and the hippocampus to statistical learning abilities in childhood: 5–8.5 years).



6. Moving forward

The goal of this article is to explicitly lay out differences between an exclusively feed-forward/bottom-up of perceptual development and a model that allows feedback-dependent/top-down connections in order to facilitate future work. Here, we highlight several crucial areas for future investigation to mediate between these two types of models.

6.1 A demonstration of fast, task/context-specific changes in infant perception

A major weakness in the current evidence for a top-down model of perceptual development is that we do not know if infants are able to rapidly modify their perception in a flexible and context-driven fashion. This type of rapid, flexible perceptual change is predicted to be available to young infants in a feedback/top-down model only and could provide a means by which feedback connections can drive broader changes in perceptual development. The current empirical findings are not sufficient: (1) They either show changes in perception as evidenced behaviorally but without a demonstration of flexibility or task-sensitivity (e.g., findings from [Needham & Modi, 1999](#)), or (2) They demonstrating rapid and specific changes but only neurally, no evidence from behavior ([Emberson et al., 2015](#)). While neural findings in particular are suggestive, they must be viewed with reservation because, given the neuroanatomical limitations of infants, it is possible that neural changes based on feedback are not sufficiently robust or rapid to modulate perception as evidenced in behavior.

6.2 Interaction with the development of long-range neural connectivity

Long-range neural connectivity exhibits a protracted developmental trajectory particularly in the white matter connections between disparate cortical regions with changes into adolescence ([Barnea-Goraly et al., 2005](#)). Thus, infants have limited neural connectivity which would presumably limit any potential long-range inter-regional interactions. However, the picture

of connectivity in infancy has been changing with the increased focus on this topic in developmental cognitive neuroscience. Recent findings have demonstrated that while white matter does develop in a protracted development, there are rapid changes in white matter development in the first year or two of life (Dean et al., 2015). Moreover, while it was the dominant view from resting-state measures of functional connectivity, that infants had very poor long-range connectivity (Gao, Lin, Grewen, & Gilmore, 2017; Homae et al., 2010; Smyser, Snyder, & Neil, 2011), some recent studies have suggested that after a period of rapid development there is a core of adult-like connectivity by 6 months of age (Gilmore, Knickmeyer, & Gao, 2018; Sours et al., 2017). In general, it is important to determine how the use of feedback connections to modify perception relates to the development of long-range connectivity (determined both by measures of white matter and functional connectivity). It may be that there is a mutual relationship between these two developmental time-courses. Immaturity in long-range connectivity should constrain the use of feedback connections and the interactions between disparate regions of the brain but it may be that the interactions of disparate regions of the brain can help guide future refinement and development of long-range connectivity as suggested by Johnson (2011). Indeed, in adults, engagement of feedback connections has been shown to increase functional coupling across disparate regions of the brain (Al-Aidroos, Said, & Turk-Browne, 2012; Tompary, Duncan, & Davachi, 2015).

6.3 Determining the long-term impact of rapid changes in perception

While a feedback/top-down model of perceptual development predicts that infants would be capable of rapid changes in perception, perceptual development at a population-level emerges over months. If rapid top-down changes in perception can be demonstrated in infants, it will be necessary to reconcile perceptual changes at these two time-scales. In adults, Folstein et al. (2015) present evidence that rapid changes in visual perception, precipitated by top-down processes including learning and memory, persist for 11 days.

Similarly, it could be that very short periods of top-down modulation result in lasting changes in the cortex and need to be tested in infants. Relatedly, it may be possible that infants have many opportunities to engage in top-down modulation of their perceptual cortices based on their experience. These numerous, rapid, top-down changes in perception could build on

each other and, ultimately, give rise to a broader developmental trajectory. Indeed, if changes in perception can occur arising from a few seconds to a few minutes of experience, infants are likely provided with many such bouts of experience that could support top-down engagement. An exciting area for future investigation is to determine the relationship between the structure of real-world infant experience (Fausey, Jayaraman, & Smith, 2016; Sugden & Moulson, 2018) and the learning abilities that infants demonstrate in the lab to determine how readily an infant could engage in top-down modulation of their perceptual systems in their daily lives.

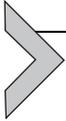
6.4 Differentiate the sources of top-down modulation

A goal of the current conception of a feedback/top-down model is to bring attention to the need to carefully define and empirically evaluate the contribution of top-down modulation and neuroanatomical feedback connections in early perceptual development. However, to this end, the current model lumps many potentially disparate sources of top-down modulation into a single box within an over-simplified boxology (Fig. 1). However, if feedback/top-down models are ultimately favored over bottom-up ones, the differential development and contribution of these many potential sources of top-down modulation must be considered carefully and systematically. In the context of learning and memory, do different types of learning and memory systems differentially modulate visual perception in infancy either based on their developmental emergence, robustness, level of modulation in the visual system, etc.?

6.5 Why do developmental changes happen when they do?

As outlined earlier, exclusively feed-forward/bottom-up models predict a sequential type of development where development proceeds from lower-levels of vision to higher-levels, and then finally more cognitive influences can arise. In this model, it is very clear what develops but less clear how it develops. By contrast, in feedback/top-down models, many aspects of development can occur simultaneously as multiple regions of the brain and levels of processing interact. However, the assertion that every level develops together does not result in a mechanistic explanation of how perceptual development occurs. Indeed, specific testable hypotheses about what should develop when needs to be determined and investigated. For example, if top-down modulation is available early in life and can be precipitated by short amount of experience, why do developmental changes in perception

occur when they do at the population-level? For example, why does the other race effect (i.e., perceptual narrowing) occur from 6 to 10 months and not earlier? Are there constraints on the abilities of higher-level systems before that point? Is there an important change in input that allows the differential engagement of higher-level systems?



7. Conclusion

Perception is where cognition and reality meet (Neisser, 1976). In the mature brain, this meeting has been argued to be a meeting between incoming sensory input (reality) and top-down information from the rest of the brain (cognition), including our powerful learning and memory systems. Here, we ask how does this type of system develop? It has long been proposed that visual development is exclusively bottom-up that employs only feed-forward neural connections and the use of feedback connections and top-down modulation emerges later in development (e.g., in early childhood). Here, I outline an alternative approach where learning and memory abilities, among other cognitive capacities, send top-down signals sent through feedback connections are available starting early in life. Not only is there already evidence that young infants can engage in top-down modulation of their perceptual systems, but there are a number of key behavioral findings that do not fit with an exclusively feed-forward/bottom-up model of perceptual development. The bottom-up or top-down nature of perceptual development is a key issue for perceptual development research but also is essential to determining the fundamental nature of perception.

References

- Ahissar, M., & Hochstein, S. (1993). Attentional control of early perceptual learning. *Proceedings of National Academy of Sciences*, *90*(12), 5718–5722. <https://doi.org/10.1073/pnas.90.12.5718>.
- Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 257–264.
- Al-Aidroos, N., Said, C. P., & Turk-Browne, N. B. (2012). Top-down attention switches coupling between low-level and high-level areas of human visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 14675–14680.
- Arteberry, M. E., & Kellman, P. J. (2016). *Development of perception in infancy: The cradle of knowledge revisited*. New York: Oxford University Press.
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*, 321–324.
- Aslin, R. N., Shukla, M., & Emberson, L. L. (2015). Hemodynamic correlates of cognition in human infants. *Annual Review of Psychology*, *66*, 349–379.
- Aslin, R. N., & Smith, L. B. (1988). Perceptual development. *Annual Review of Psychology*, *39*, 435–473.

- Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*, 437–443.
- Baldassarre, A., Lewis, C. M., Committeri, G., Snyder, A. Z., Romani, G. L., & Corbetta, M. (2012). Individual variability in functional connectivity predicts performance of a perceptual task. *Proceedings of the National Academy of Sciences*, *109*(9), 3516–3521. <https://doi.org/10.1073/pnas.1113148109>.
- Barnea-Goraly, N., Menon, V., Eckert, M., Tamm, L., Bammmer, R., Karchemskiy, A., et al. (2005). White matter development during childhood and adolescence: A cross-sectional diffusion tensor imaging study. *Cerebral Cortex*, *15*, 1848–1854.
- Bauer, P. J. (2004). Getting explicit memory off the ground: Steps toward construction of a neuro-developmental account of changes in the first two years of life. *Developmental Review*, *24*, 347–373.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences*, *36*, 181–204.
- Curran, T. (1997). Higher-order associative learning in amnesia: Evidence from the serial reaction time task. *Journal of Cognitive Neuroscience*, *9*, 522–533.
- Dean, D. C., O’Muircheartaigh, J., Dirks, H., Waskiewicz, N., Walker, L., Doernberg, E., et al. (2015). Characterizing longitudinal white matter development during early childhood. *Brain Structure and Function*, *220*, 1921–1933.
- Emberson, L. L. (2016). Gaining knowledge mediates changes in perception (without differences in attention): A case for perceptual learning. Commentary on: Cognition does not affect perception: Evaluating the evidence for “top-down” effects. *Brain and Behavioral Sciences*, *39*, 28–30.
- Emberson, L. L., & Amso, D. (2012). Learning to sample: Eye tracking and fMRI indices of changes in object perception. *Journal of Cognitive Neuroscience*, *24*, 2030–2042.
- Emberson, L. L., Boldin, A. M., Riccio, J. E., Guillet, R., & Aslin, R. N. (2017). Deficits in top-down sensory prediction in infants at risk due to premature birth report deficits in top-down sensory prediction in infants at risk due to premature birth. *Current Biology*, *27*, 1–6.
- Emberson, L. L., Cannon, G., Palmeri, H., Richards, J. E., & Aslin, R. N. (2017). Using fNIRS to examine occipital and temporal responses to stimulus repetition in young infants: Evidence of selective frontal cortex involvement. *Developmental Cognitive Neuroscience*, *23*, 26–38.
- Emberson, L. L., Richards, J. E., & Aslin, R. N. (2015). Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 9585–9590. <https://doi.org/10.1073/pnas.1510343112>.
- Fanselow, M. S., & Ledoux, J. E. (1999). Why we think plasticity underlying Pavlovian fear conditioning occurs in the basolateral amygdala. *Neuron*, *23*, 229–232.
- Fausey, C. M., Jayaraman, S., & Smith, L. B. (2016). From faces to hands: Changing visual input in the first two years. *Cognition*, *152*, 101–107.
- Finn, A. S., Kharitonaova, M., Holtby, N., & Sheridan, M. A. (2019). Prefrontal and hippocampal structure predict statistical learning ability in early childhood. *Journal of Cognitive Neuroscience*, *31*, 126–137.
- Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for ‘top-down’ effects. *Behavioral and Brain Sciences*, *39*, e229.
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 15822–15826.
- Folstein, J. R., Palmeri, T. J., Van Gulick, A. E., & Gauthier, I. (2015). Category learning stretches neural representations in visual cortex. *Current Directions in Psychological Science*, *24*, 17–23.

- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 360, 815–836.
- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, 19, 117–125.
- Fuccillo, M. V. (2016). Striatal circuits as a common node for autism pathophysiology. *Frontiers in Neuroscience*, 10, 27.
- Gao, W., Lin, W., Grewen, K., & Gilmore, J. H. (2017). Functional connectivity of the infant human brain. *The Neuroscientist*, 23, 169–184.
- Gibson, E. J. (1969). *Principals of perceptual learning and development*. New York: Appleton-Century-Crofts.
- Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, 14, 350–363.
- Gilmore, J. H., Knickmeyer, R. C., & Gao, W. (2018). Imaging structural and functional brain development in early childhood. *Nature Reviews Neuroscience*, 19, 123–137.
- Hadley, H., Rost, G. C., Fava, E., & Scott, L. S. (2014). A mechanistic approach to cross-domain perceptual narrowing in first year of life. *Brain Sciences*, 4, 613–634.
- Hannula, D. E., Ryan, J. D., Tranel, D., & Cohen, N. J. (2007). Rapid onset relational memory effects are evident in eye movement behavior, but not in hippocampal amnesia. *Journal of Cognitive Neuroscience*, 19, 1690–1705.
- Homae, F., Watanabe, H., Otobe, T., Nakano, T., Go, T., Konishi, Y., et al. (2010). Development of global cortical networks in early infancy. *The Journal of Neuroscience*, 30, 4877–4882.
- Hutchinson, J. B., & Turk-Browne, N. B. (2012). Memory-guided attention: Control from multiple memory systems. *Trends in Cognitive Sciences*, 16, 576–579.
- Johnson, M. H. (2011). Interactive specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, 1, 7–21.
- Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science*, 18, 1084–1089.
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 83, B35–B42.
- Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & De Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, 22, 2197–2206.
- Luck, S. J., Fan, S., & Hillyard, S. A. (1993). Attention-related modulation of sensory-evoked brain activity in a visual search task. *Journal of Cognitive Neuroscience*, 5, 188–195.
- Lupyan, G., & Spivey, M. J. (2008). Perceptual processing is facilitated by ascribing meaning to novel stimuli. *Current Biology*, 18, 410–412.
- Markant, J., Oakes, L. M., & Amso, D. (2016). Visual selective attention biases contribute to the other-race effect among 9-month-old infants. *Developmental Psychobiology*, 58, 355–365.
- Markant, J., & Scott, L. S. (2018). Attention and perceptual learning interact in the development of the other-race effect. *Current Directions in Psychological Science*, 27, 163–169.
- Maurer, D., & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology*, 56, 154–178.
- McMains, S., & Kastner, S. (2011). Interactions of top-down and bottom-up mechanisms in human visual cortex. *The Journal of Neuroscience*, 31, 587–597.
- Mercuri, E., Atkinson, J., Braddick, O., Anker, S., Cowan, F., Rutherford, M., et al. (1997). Basal ganglia damage and impaired visual function in the newborn infant. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, 77, F111–F114.

- Nakano, T., Watanabe, H., Homae, F., & Taga, G. (2009). Prefrontal cortical involvement in young infants' analysis of novelty. *Cerebral Cortex*, *19*, 455–463.
- Needham, A., & Baillargeon, R. (1998). Effects of prior experience on 4.5-month-old infants' object segregation. *Infant Behavior and Development*, *21*, 1–24.
- Needham, A., Dueker, G., & Lockhead, G. (2005). Infants' formation and use of categories to segregate objects. *Cognition*, *94*, 215–240.
- Needham, A., & Modi, A. (1999). Infants' use of prior experiences with objects in object segregation: Implications for object recognition in infancy. *Advances in Child Development and Behavior*, *27*, 99–133.
- Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. New York: WH Freeman & Co.
- Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., et al. (2005). Plasticity of face processing in infancy. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 5297–5300.
- Peterson, M. A., De Gelder, B., Rapcsak, S. Z., & Gerhardstein, P. C. (2000). Object memory effects on figure assignment: Conscious object recognition is not necessary or sufficient. *Vision Research*, *40*, 1549–1567.
- Quinn, P. C., & Schyns, P. G. (2003). What goes up may come down: Perceptual process and knowledge access in the organization of complex visual patterns by young infants. *Cognitive Science*, *27*, 923–935.
- Quinn, P. C., Schyns, P. G., & Goldstone, R. L. (2006). The interplay between perceptual organization and categorization in the representation of complex visual patterns by young infants. *Journal of Experimental Child Psychology*, *95*, 117–127.
- Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–87.
- Richmond, J., & Nelson, C. A. (2009). Relational memory during infancy: Evidence from eye tracking. *Developmental Science*, *12*, 549–556.
- Rovee-Collier, C. (1997). Dissociations in infant memory: Rethinking the development of implicit and explicit memory. *Psychological Review*, *104*, 467–498.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
- Saffran, J. R., & Thiessen, E. D. (2007). Domain-general learning capacities. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 68–86).
- Sagi, D., & Tanne, D. (1994). Perceptual learning: Learning to see. *Current Opinion in Neurobiology*, *4*, 195–199.
- Scott, L. S., & Monesson, A. (2009). The origin of biases in face perception. *Psychological Science*, *20*, 676–680.
- Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science*, *16*, 197–201.
- Scott, L. S., Shannon, R. W., & Nelson, C. A. (2006). Neural correlates of human and monkey face processing in 9-month-old infants. *Infancy*, *10*, 171–186.
- Seitz, A. R., Nanez Sr., J. E., Holloway, S., Tsushima, Y., & Watanabe, T. (2006). Two cases requiring external reinforcement in perceptual learning. *Journal of Vision*, *6*, 966–973. <https://doi.org/10.1167/6.9.9>
- Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron*, *60*, 1169–1181.
- Smyser, C. D., Snyder, A. Z., & Neil, J. J. (2011). Functional connectivity MRI in infants: Exploration of the functional organization of the developing brain. *NeuroImage*, *56*, 1437–1452.

- Sours, C., Raghavan, P., Foxworthy, W. A., Meredith, M. A., El Metwally, D., Zhuo, J., et al. (2017). Cortical multisensory connectivity is present near birth in humans. *Brain Imaging and Behavior, 11*, 1207–1213.
- Stokes, M., Atherton, K., Patai, E. Z., & Nobre, A. C. (2012). Long-term memory prepares neural activity for perception. *Proceedings of the National Academy of Sciences of the United States of America, 109*, E360–E367.
- Sugden, N. A., & Moulson, M. C. (2018). These are the people in your neighbourhood: Consistency and persistence in infants' exposure to caregivers' relatives' and strangers' faces across contexts. *Vision Research*. in press.
- Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience, 15*, 745–756.
- Swingle, D. (2008). The roots of the early vocabulary in infants' learning from speech. *Current Directions in Psychological Science, 17*, 308–312.
- Tompary, A., Duncan, K., & Davachi, L. (2015). Consolidation of associative and item memory is related to post-encoding functional connectivity between the ventral tegmental area and different medial temporal lobe subregions during an unrelated task. *Journal of Neuroscience, 35*, 7326–7331.
- Tummeltshammer, K. S., & Kirkham, N. Z. (2013). Learning to look: Probabilistic variation and noise guide infants' eye movements. *Developmental Science, 16*, 760–771.
- Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *The Journal of Neuroscience, 30*, 11177–11187.
- Wang, S. S. H., Kloth, A. D., & Badura, A. (2014). The cerebellum, sensitive periods, and autism. *Neuron, 83*, 518–532.
- Webb, S. J., Monk, C. S., & Nelson, C. A. (2001). Mechanisms of postnatal neurobiological development: Implications for human development. *Developmental Neuropsychology, 19*, 147–171.
- Werchan, D. M., Collins, A. G. E., Frank, M. J., & Amso, D. (2016). Role of prefrontal cortex in learning and generalizing hierarchical rules in 8-month-old infants. *Journal of Neuroscience, 36*, 10314–10322.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*, 49–63.
- Westermann, G., & Mareschal, D. (2012). Mechanisms of developmental change in infant categorization. *Cognitive Development, 27*, 367–382.
- Westermann, G., & Mareschal, D. (2014). From perceptual to language-mediated categorization. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 369*, 20120391.
- Zhou, Y., Lai, B., & Gan, W. B. (2017). Monocular deprivation induces dendritic spine elimination in the developing mouse visual cortex. *Scientific Reports, 7*, 1–11.
- Zuo, Y., Yang, G., Kwon, E., & Gan, W.-B. (2005). Long-term sensory deprivation prevents dendritic spine loss in primary somatosensory cortex. *Nature, 436*, 261–265.