Title
Learning (to Learn) from Spatial Attention Cues During Infancy

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 34(34)

ISSN
1069-7977

Authors
Wu, Rachel
Kirkham, Natasha

Publication Date
2012

Peer reviewed

Learning (to Learn) from Spatial Attention Cues During Infancy

Winner of a Robert J. Glushko Dissertation/Thesis Prize in Cognitive Science

Rachel Wu (r.wu@bbk.ac.uk) and Natasha Z. Kirkham (n.kirkham@bbk.ac.uk)

Centre for Brain and Cognitive Development, Department of Psychological Sciences, Birkbeck, University of London
Malet Street, London, WC1E 7HX, UK

Abstract

Human infants develop a variety of attentional mechanisms that allow them to extract relevant information from a cluttered world. We know that both social and non-social cues shift infants' attention, but not how infants use these cues to learn basic events. With over 450 infants, four extensive eye-tracking studies in this thesis established a controlled paradigm for investigating how attention cues shape early learning. The results showed that infants' ability to learn about structures in their environment (i.e., predicting the appearance of audio-visual events and forming expectations about co-occurring features) is dependent on the presence and nature of attention cues. By 8 months, infants learned these events significantly better with social cues (e.g., eye gaze, infant-directed speech, expression of interest) than with non-social cues (e.g., flashing squares) or without any attentional cueing. Importantly, when presented with multiple events to learn and cued by a face to one specific event, infants learned the cued event and ignored the non-cued event. The last study found that familiar communicative social signals (i.e., an engaging face that spoke to the infant) boosted 9-month-olds' learning about cued events. In particular, the engaging face supported learning from non-social cues, providing evidence for a mechanism explaining how infants learn to learn from unfamiliar attention cues such as pointing or arrows. Our results showed that though social cues may temporarily detract attention away from certain learning events in the world, they appear to stimulate infants to display better learning about the cued event than when infants learn with other attention cues or on their own without attention cues.

Keywords: attention cues; pattern learning; infant eye-tracking; cognitive development; social cues.

Introduction

The relationship between attention and learning is reciprocal: learning is enabled and facilitated by attentional selection, and attentional selection builds on previously learned knowledge. For example, studies with adults have shown that appropriate attentional selection can help filter distracters to learn about targets (e.g., Hillyard, Hink, Schwent, & Pincton, 1973; Van Voorhis & Hillyard, 1977) and that learned knowledge about targets and distracters allows more efficient target selection among distracters (e.g., Chun & Jiang, 1998; Kruschke, 2011; Mazza, Turatto, Umilta, & Eimer, 2007). This bi-directional relationship between attention and learning is highlighted in the developmental perspective, when considering such abilities in inexperienced learners (i.e., infants).

Human infants have to learn a great deal of information in a complex world that is filled with ambiguity. Not only are many different features and dimensions present in the environment, but also they are often unrelated to any reinforcement or feedback. Selective attention, focusing on some information but ignoring other aspects, is a crucial component for learning, especially in situations involving uncertainty (e.g., Dayan, Kakade, & Montague, 2000). There are two broad solutions to this complexity and ambiguity inherent in the learning environment: (a) innate constraints on the cues selected for processing, or (b) rapid learning-to-learn mechanisms that assess cue-reliability and thus guide learning. In the bottom-up solution, there are environmental and innate biases that constrain what infants attend to (e.g., Brazelton, Scholl, & Robey, 1966; Johnson, Dziurawiecb, Ellisc & Morton, 1991), and in the top-down solution, infants can sample their environment and quickly learn what sources of information make the most sense (e.g., Saffran, Aslin, & Newport, 1996; Fiser & Aslin, 2002). Mechanisms of learned information selection are generally considered top-down and thus may be tuned to specific task demands. Many experiments with older children and adults show that top-down selection produces efficient learning and allows for pre-activation of relevant neural circuits, which may lead to better processing of the stimulus (e.g., Driver & Frith, 2000). Would infants learn differently with different types of attention cues?

Both bottom-up and top-down (learned) attention cues shift infants' attention: social stimuli (e.g., eye gaze, infant-directed speech, faces), which infants learn to follow by four months of age, and non-social stimuli (e.g., bright lights/colors, motion), which capture infants' attention from birth. In following these stimuli, infants can develop attentional mechanisms for extracting relevant information from a cluttered multimodal world, because these stimuli can cue infants to the location of relevant events. In turn, relying on these attention cues can shape learning because what controls attention can gate processing (Moran & Desimone, 1985). Though it is well known that both social and non-social cues shift infants' attention, it is unclear how these attention cues differentially affect learning of basic events. We know that infants are capable of learning about the structure of their environment in the simplified laboratory setting. For example, infants recognize that certain sights and sounds belong together (e.g., toys dropping, cars driving, people talking, Bahrick, 2004; Lewkowicz, 2000) and track the co-occurrence of visual features (e.g., Fiser & Aslin, 2002; Kirkham, Slemmer, & Johnson, 2002). In the typical cluttered environment, however, infants are often presented with multiple events to learn. How do infants know which events are important to learn? Would different attention cues similarly help infants select information to learn?

Summary of Thesis Studies

The tight coupling of attention and learning, especially for the young learner, determines how information is selected, retained, and eventually applied. This thesis capitalized on recent advances in methodologies (i.e., eye-tracking) that allow for paradigms that include more ecologically valid environments with noise and distraction testing young age groups.

Learning from cues (specifically attention cues) is an ideal context for studying this interaction between selective attention and learning efficacy. The thesis investigated whether infants learn differently from different attention cues (or no cues), and how infants learn to learn from attention cues. In a distraction-filled environment, visual spatial attention cues (e.g., other's eye gaze or flashing lights) can highlight events in a particular location and facilitate processing of those events (see Posner, 1980). This is a critical aspect for the young learner, who may not know what to learn at a given moment. The components of a spatial cueing experimental paradigm include a cue, target, and distracter(s). Commonly, the cues shift attention to a particular location to prepare the viewer for the target event in that location. Cues can either attract attention (bright flashing lights, big moving objects) or shift attention from themselves to another location (learned attention cues: eye gaze, arrows).

My 3-year PhD investigated how different attention cues (or no cues) affect learning during infancy. My studies showed for the first time in one cohesive paradigm that the presence and nature of these attention cues mediate what infants learn about the structures in the environment (i.e., predicting the appearance of audio-visual events or forming expectations about co-occurring features). A sample of over 450 infants across four extensive eye-tracking studies (Wu & Kirkham, 2010; Wu, Gopnik, et al., 2011; Wu, Kirkham, et al., under revision) showed that by 8 months, infants learned the structures in their environment significantly better with subtle social cues (e.g., eye gaze, infant-directed speech) than with salient non-social cues (e.g., flashing squares) or without any attentional cueing. Importantly, when presented with multiple events to learn and cued by a face (rather than a flashing square) to one specific event, infants learned the cued event rather than the non-cued event. These results show that when naïve learners do not know what to learn, social objects (faces) shape the likelihood of learning cued targets, and that attending to social cues provides infants with an optimal strategy for learning appropriate events despite the presence of distractions.

Study 1 In the first study, Experiment 1 used social cues to direct 8- and 4-month-old infants' attention to multimodal events (i.e., animations of toys accompanied by specific sounds), while identical distracter events were presented in another location. Experiment 2 directed 8-month-olds' attention with colorful flashes to the same events. Experiment 3 measured 8-month-olds' baseline learning without attention cues. The test trials in all experiments played only the sounds previously associated with a particular animation, and looking time was measured to each now blank location that previously contained an object. Results showed that the 8-month-olds exposed to social cues learned about the cued audiovisual event (i.e., they predicted its appearance in the correct rather than incorrect cued location) (Figure 1). The 4-month-olds, however, displayed only general spatial learning from social cues (i.e., they looked to both correct and incorrect cued locations to predict its appearance), suggesting that infants *learn to learn* from a face stimulus between 4 to 8 months. Eight-month-old infants cued with the colorful flashes looked indiscriminately to both correct and incorrect cued locations during test (similar to the 4-month-olds learning from social cues) despite attending for equal duration to the training events as the 8-month-olds with social cues. Results from Experiment 3 (no learning) indicated that the learning effects from Experiments 1 and 2 resulted from exposure to the different cues and multimodal events. In summary, this first series of experiments shows that infants' attention to target events is captured equally well by both social and non-social cues, but learning is deeper and more precise with social cues. This study is published in the *Journal of Experimental Child Psychology* (Wu & Kirkham, 2010).

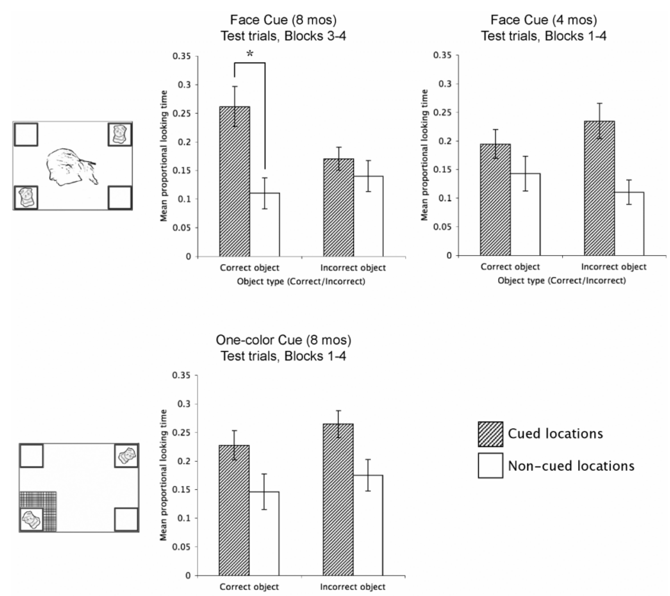

Figure 1: Stimuli and results from Study 1. Mean proportional looking times to locations during test trials from Blocks 3 and 4 in the Face Cue condition (8 months) and Blocks 1 to 4 in the Face Cue condition (4 months) and One-color Cue condition (8 months). The 8-month-olds in the Face Cue condition looked more to the cued correct object location, whereas the 4-month-olds in the same

condition looked longer to cued locations than to non-cued locations regardless of object–sound mappings. The 8-month-olds in the One-color Cue condition (flashing squares) also looked longer only to cued locations than to non-cued locations regardless of multimodal information. *p=.01

Study 2 Study 2 used visual statistical learning as the dependent measure, and compared how infants learn from social cues to that with no cues. In laboratory experiments, infants can learn patterns of co-occurring visual features (e.g., Fiser & Aslin, 2002). Once infants learn the statistical regularities, however, how do they use the knowledge? In addition, which patterns do infants learn when presented with many options (as in the cluttered world outside of the laboratory)? In this study, infants were shown clusters of 3 shapes, where two always co-occurred and a third changed on every trial. Infants were either shown these shape patterns on their own (Experiments 1 and 2) or shown the pattern cued by a face (Experiment 3) or also with a distracter pattern in a non-cued location (Experiment 4). Test trials displayed shapes moving apart: either the co-occurring shapes (looking longer related to a preference for the inconsistent/violation of expectation) or the non-co-occurring shapes (looking longer related to a preference for the consistent). Tracking co-occurring units and inferring that they are larger fused units help identify integral object parts for object individuation, recognition, and categorization. Experiment 1 showed that 9-month-old infants interpret co-occurring features as larger fused units (i.e., infants looked longer when co-occurring features split apart). The other three experiments showed that social cues (compared to no cues) help 9-month-olds choose patterns among distracters during learning and test (Figure 2). These findings suggest that by 9 months, infants can use feature co-occurrence to learn about objects and that social cues shape such foundational learning in distraction-filled environments. In particular, though social cues may temporarily detract attention away from certain learning events in the world, they appear to stimulate infants to display the learning better in complex situations than when infants learn on their own without social cues. Task difficulty also mediates how inferences (made from visual statistical processing) are exhibited during test trials. This study is published in *Developmental Psychology* (Wu, Gopnik, et al., 2011).

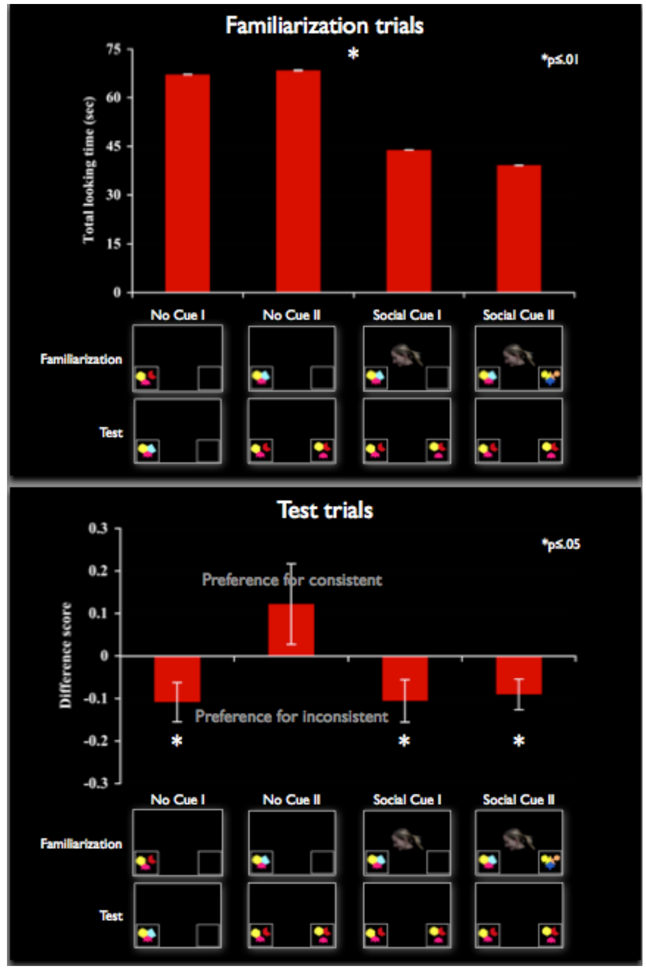

Figure 2: Stimuli and results from Study 2. The top half of the figure shows the total looking times to the target pattern for the familiarization trials across the four conditions (No Cue I, No Cue II, Social Cue I, and Social Cue II). Infants looked longer to the target pattern in the two No Cue conditions, and less in the two Social Cue conditions because they split their attention between the face and target pattern. *$p \leq .01$ The bottom half of the figure displays the difference scores across four conditions (mean difference between proportional looking times for consistent minus inconsistent events during test). A negative value reflected a preference for the inconsistent splits (i.e., looking longer at events showing the separation of features that co-occurred rather than the separation of features that did not co-occur). Infants showed this preference in the easier test condition without a cue (No Cue 1), and with a social cue regardless of task difficulty. *$p \leq .05$

Study 3 Following up from these findings, that infants learn better from social than non-social cues, Study 3 investigated whether infants can *learn to learn* from typical non-social objects that are interactive like social cues (e.g., objects that move when infants look at it, e.g., Johnson, Slaughter, & Carey, 1998). The main procedure consisted of three phases: training, familiarization, and test. During the training phase, 9-month-olds interacted with a centrally presented teapot (similar to Deligianni, Senju, Gergeley, & Csibra, 2011, who found that infants follow such interactive objects). Infants' fixations on the teapot caused it to jump or lift its lid. During the familiarization and test phases, infants were presented with the pattern events from Study 2, with the original face now replaced by the teapot. We found that if infants followed the teapot during familiarization, they seemed to have trouble learning about the cued event, only perhaps learning the cued location (general spatial learning similar to 4-month-olds cued by a face in Study 1). Infants who did not follow the teapot (and perhaps ignored this cue) seemed to learn about the non-cued event. While only preliminary observations can be made at this stage, this pilot study begins to address *how* infants learn to use cues, rather than describing cues they do or do not learn from.

Study 4 In Study 1, the flashing square cue was unsuccessful at producing specific learning in infants (Wu & Kirkham, 2010). Study 4 investigated whether the pairing of familiar social cues with unfamiliar flashing cues could help infants learn from these novel flashing cues. Nine-month-old infants were eye-tracked during a Training phase, followed by a Generalization phase. In the Social Scaffolding condition, the Training phase consisted of an expressive face that spoke to the infant and then froze with a smile, while identical audio-visual animations appeared in diagonally opposite corners. At the same time, a red flashing square cued the infant to a specific target frame containing an object. In the Extended Practice condition, infants only saw the flashing square and multimodal events. The Generalization phase for both conditions displayed new audio-visual events with only the red flashing square cues. The test trials in each phase played one of the sounds previously associated with a particular animation, and looking time to each location was measured. In the Social Scaffolding condition, infants anticipated that the events would appear in the correct cued locations for both phases. In the Extended Practice condition, infants first displayed general spatial learning (replicating Experiment 2 in Study 1) during the Training phase, and then showed no learning during the Generalization phase (Figure 3). These findings suggest that initial exposure to familiar social cues can elicit and maintain specific learning from novel attention-orienting cues. Moreover, this could provide evidence for a mechanism explaining *how* infants and children can learn to learn from distal attention cues such as pointing fingers and arrows. This is an important first step towards elucidating an emerging ability to use familiar attention cues to support, enhance, and mediate learning about unfamiliar cues, going beyond documenting *which* cues guide attention and learning during infancy to proposing a mechanism for *how* this cascading learning effect occurs. Portions of this study are a CogSci proceeding and EuroCogSci proceeding (awarded Best Student Paper Prize), and is currently under revision for a journal submission.

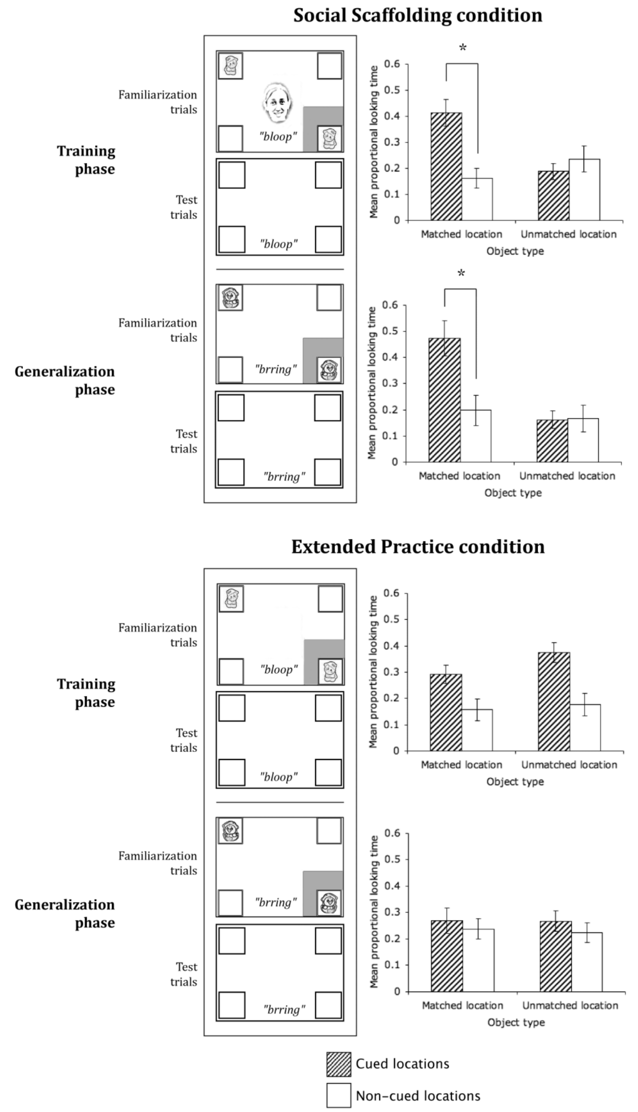

Figure 3: Stimuli and results from familiarization and test trials for both conditions in Study 4. All stimuli were in full color on a black background. When exposed to direct social signals paired with flashing squares, infants predicted objects would appear in matched cued locations (Social Scaffolding condition, Training phase). Infants continued learning specifically from flashing squares even after social signals were no longer presented (Social Scaffolding condition, Generalization phase). With only exposure to flashing squares, infants exhibited general spatial learning and no transfer of learning, suggesting that initial exposure to direct social signals is necessary to elicit and maintain specific learning from novel cues at this age. *p<.03

Conclusions

By studying how infants learn cued events (and do not learn distracter events) with attention cues, this thesis investigated the usefulness of attentional biases for infants' cognitive development. This research is important because it goes beyond describing singular events that infants can learn by exploring how they figure out *what* to learn in the noisy environment. Investigating how infants learn to learn with cues provides a more accurate picture of infants' learning mechanisms in the rich natural environment. My thesis work contributed to the study of the optimal dynamics of selective attention and successful learning in typical development, which in turn would inform populations with learning difficulties. Though the studies in this thesis only investigated learning from visual spatial attention cues, research on the overall question of how one learns to learn among distraction is important for many areas of cognitive development. *Learning to learn* is essentially the interplay of gaining control over one's perceptual input based on previous experience (Aslin, 2008; Frank et al., 2009) and current interactions with the environment (Sarter, Givens, & Bruno, 2001). This is one way of describing how one gains expertise (Gopnik, 2009; Nodine, Kundel, Lauver, & Toto, 1996; Solso, 2001). We are trained to detect and learn from certain stimuli, whether it is learning from people (Bigelow, 1998; Csibra, 2010; Ghazanfar & Santos, 2004) or learning from other cues (e.g., arrows). With some rudimentary initial biases, infants *learn to learn* and become more adapted to function appropriately according to environmental norms.

Testament to the interdisciplinary contribution of this thesis, at least seven current projects on computational modelling, robotics, genetics, and atypical development are based directly on this work and dataset (please see CV). The large dataset (450+ infants) from this thesis is ideal for computational modelling and genetics projects, and the benchmark of typical behavior can be used to compare with data from atypical development.

This work also has led directly to the organization of two attention-learning workshops in London (Jan 2012 – organized by R. Wu, received £1818 to organize workshop) and Tokyo (March 2012 – organized by R. Wu, S. Shimojo, and T. Omori, funded by the Tamagawa-CalTech grant), encouraging discussions among computational modellers, developmental psychologists, neuroscientists, and roboticists to promote the emergence of this sub-field on the interaction of attention and learning.

Acknowledgments

I owe a great deal of gratitude to my supervisors, Natasha Kirkham and Denis Mareschal for their guidance, dedication, and encouragement, as well as to the other collaborators on this project, Teodora Gliga, Alison Gopnik, Daniel Richardson, and Kristen Swan. Thank you to the Glushko-Samuelson Foundation and the Prize committee for the generous award, as well as providing travel grants to present this work at various Cognitive Science meetings. Thank you to Leslie Tucker and Marian Greensmith for help with data collection. This research was supported by a grant to NK from the British Academy, Grant No. SG-47879 and two grants to NK and RW from the University of London Central Research Fund.

References

Aslin, R. N. (2008). Headed in the right direction: A commentary on Yoshida and Smith. *Infancy, 13*(3), 275-278.

Bahrick, L. E. (2004). Development of intermodal perception. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science* (Vol. 2, pp. 614–617). London: Nature Publishing Group.

Bigelow, A. E. (1998). Infants' sensitivity to familiar imperfect contingencies in social interaction. *Infant Behavior and Development, 21*(1), 149-162.

Brazelton, T. B., Scholl, M. L., & Robey, J. S. (1966). Visual responses in the newborn. *Pediatrics, 37*, 284-290.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*(1), 28-71.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind and Language, 25*(2), 141-168.

Dayan, P., Kakade, S., & Montague, P. R. (2000). Learning and selective attention. *Nature Neuroscience, 3 Supplement*, 1218-1223.

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-months-old infants. *Developmental Psychology*, 47, 1499-1503.

Driver, J., & Frith, C. (2000). Shifting baselines in attention research. *Nature Reviews Neuroscience, 1*, 147-148.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, USA, 99*, 15822-15826.

Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition, 110*(2), 160-170.

Ghazanfar, A. A., & Santos, L. R. (2004). Primate brains in the wild: The sensory bases for social interactions. *Nature Reviews Neuroscience, 5*(8), 603-616.

Gopnik, A. (2009). *The Philosophical Baby: What Children's Minds Tell Us About Truth, Love, and the Meaning of Life*. New York: Farrar, Straus and Giroux.

Hillyard, S. A., Hink, R. F., Schwent, V. L., and Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science, 182*, 177-180.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40* (1-2), 1-19.

Johnson, S. C., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science, 1*(2), 233-238.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*(2), B35-B42.

Kruschke, J. K. (2011). Models of attentional learning. In: E. M. Pothos and A. J. Wills (eds.), *Formal Approaches* *in Categorization*, pp. 120-152. Cambridge University Press.

Lepsien, J., & Nobre, A. C. (2007). Attentional modulation of object representations in working memory. *Cerebral Cortex, 17*(9), 2072-2083.

Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin, 126*(2), 281-308.

Mazza, V., Turatto, M., Umilta, C., & Eimer, M. (2007). Attentional selection and identification of visual objects are reflected by distinct electrophysiological responses. *Experimental Brain Research, 181*, 531- 536.

Mitchell, J. F., Sundberg, K. A., & Reynolds, J. H. (2009). Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron, 63*, 879-888.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229*(4715), 782-784.

Nodine, C. F., Kundel, H. L., Lauver, S. C., & Toto, L. C. (1996). Nature of expertise in searching mammograms for breast masses. *Academic Radiology, 3* (12), 1000-1006.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3-25.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8- month-old infants. *Science, 274*(5294), 1926-1928.

Sarter, M., Givens, B., & Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: Where top-down meets bottom-up. *Brain Research Reviews, 35*(2), 146-160.

Scerif, G., & Wu, R. (in press). Developmental Disorders. In A.C. Nobre & S. Kastner (Eds.) *The Oxford Handbook of Attention*. Oxford: OUP

Schlesinger, M., Amso, D. & Johnson, S. P. (under review). Simulating the role of visual selective attention during development of perceptual completion.

Solso, R. L. (2001). Brain activities in a skilled versus a novice artist: An fMRI study. *Leonardo, 34*(1), 31-34.

Van Voorhis S., & Hillyard S. (1977). Visual evoked potentials and selective attention to points in space. *Perception & Psychophysics, 22*, 54–62.

Wu, R., & Kirkham, N. Z. (2010). No two cues are alike: Depth of learning during infancy is dependent on what orients attention. *Journal of Experimental Child Psychology, 107*, 118-136.

Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (2011). Infants learn about objects from statistics and people. *Developmental Psychology, 47 (5)*, 1220-1229.

Wu, R., Kirkham, N. Z., Swan, K. A., & Gliga, T. (under revision). Direct social signals scaffold learning from novel attention cues during infancy.