

## News from the field

Published online: 27 April 2011  
© Psychonomic Society, Inc. 2011

### DURATION REPRODUCTION

#### There are known unknowns

Jazayeri & Shadlen (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, **13**, 1020–1026.

The precision of our discriminations varies with stimulus length (Weber, 1851), such that we can notice both when the school clock is 5 min late with its hourly chime as well as when a musical note is played half a beat late (~0.3 s, in a mid-tempo song). Exactly how we appropriately scale the precision of our discriminations remains something of a mystery, but perhaps even more mysterious is how we cope with our variable precision. A recent paper by Jazayeri and Shadlen suggests a one-word solution to this mystery: optimally. They had observers attempt to reproduce the duration of a visual stimulus with a key press. Trials were blocked by the category of stimulus durations. Long durations were between 0.847 and 1.200 s, medium durations were between 0.671 and 1.023 s, and short durations were between 0.494 and 0.847 s. Given that there must be some imprecision in our estimates of duration, it's only natural that our we hedge our bets and err towards the average duration in a block of trials. What is surprising is that Jazayeri and Shadlen's observers knew just how much to hedge their bets. That is, their biases were the same as those of an ideal (i.e. Bayesian) observer whose precision is limited by Weber's Law: big biases for long durations and small ones for short durations.

Several labs including my own have presented evidence for Bayesian-esque decisions when uncertainty was manipulated by changing signal-to-noise ratios in the stimulus. Jazayeri and Shadlen's exploitation of Weber's Law is cleverer. Their observers' optimal decisions constitute the best evidence to date that we know the extents of our own uncertainties.—J.A.S.

### SPEECH

#### Intersensory synchrony

Vroomen J., & Stekelenburg, J. J. (2011) Perception of intersensory synchrony in audiovisual speech: Not that special. *Cognition*, **118**, 75–83

Language is a fascinating tool in human communication, and speech perception is often considered to be *that* special; a special mode of perception may be induced by speech sounds. It is not easy to demonstrate this proposition, however. The reason is mainly because it is too hard to equalize low-level (physical) factors of speech and non-speech sounds. Vroomen and Stekelenburg (2011) investigated the perception of temporal order of audiovisual speech. They tried to determine why the timing difference between speech sounds and lip movements is more difficult to detect than that between non-speech sounds and lip movements. The performance levels in speech and non-speech conditions have been explained by assuming the specialness of speech sounds for the perceptual system: speech sounds are strongly combined with lip movements in the perceptual system.

Vroomen and Stekelenburg tested this explanation by employing *sine-wave speech*. Sine-wave speech is typically synthesized with a few components varying in frequency and amplitude, which simulate the formant transitions of speech sounds. It is perceived as non-speech if listeners are without any bias, but it is perceived as speech if listeners are informed that it is speech. Thus, the sine-wave speech that they employed could become non-speech sounds for non-biased listeners and become speech sounds for biased listeners, while keeping the low-level factors equal. Two experiments of temporal-order judgments and simultaneity judgments revealed that the detection performance of the timing difference between the sine-wave speech and lip

movements was unchanged, no matter whether the sine-wave speech was perceived as non-speech or as speech. In brief, in this study, speech was not *that* special.—S.G.

## SPATIO-TEMPORAL GROUPING

### Can attention to the Ternus display modulate the perceived motion?

Aydin, M., Herzog, M.H. & Oğmen H. (2011). Attention modulates spatio-temporal grouping. *Vision Research*, 51 (4), 435–446.

The Ternus display consists of two frames separated by a temporal interval of varying duration. Each frame contains a row of three elements with partial spatial inter-frame overlap in the elements' spatial location. The perceived motion in the Ternus display depends on the temporal interval. When this interval is relatively short, 'element motion' prevails—the two elements with spatial overlap appear stationary, while the outermost element appears to jump from one end of the row to the other. However, when this interval is relatively long, 'grouping motion' prevails—all three elements appear to move as a group. Throughout the years several different hypotheses were offered to explain the bistable nature of the Ternus display, but regardless of the specific mechanism offered, many of the accounts are consistent with the view that the perceived motion in the Ternus display is the outcome of spatio-temporal grouping processes.

The goal of Aydin et al.'s study was to test whether attention can affect the perceived motion in Ternus display. Previous studies have already demonstrated that attention can modulate spatial grouping. The dynamic nature of the Ternus display allowed Aydin et al. to also explore attentional effects on temporal grouping. To that end a dual task paradigm was employed. On the 'Ternus-only' condition the observers saw a classical Ternus display and indicated whether they perceive element or group motion. On the 'dual-task' condition the observers were required to perform two tasks. The primary task involved a stream of 10 items presented at the center of the display. Some of these items were squares and the others were disks, and the observers had to indicate whether the number of squares was odd or even. In addition, the Ternus display was presented at the periphery and the observers had to indicate their perceived motion. As is typically the case with dual tasks, it was assumed that due to the need to divide attention between two tasks, less attentional resources are allocated to the Ternus display in the dual-task condition. The results show considerable attentional modulation of perceived motion: the reports of group motion were less

frequent in the dual-task condition than the Ternus-only condition. This finding suggests that the perception of group motion relies more heavily on attentional resources than the perception of element motion. One possible interpretation of these results is that the perception of group motion suffers more heavily from the depletion in attentional resources because it requires more complex grouping, involving matching elements across different retinotopic locations at different points of time. Thus, this study demonstrates, using a single paradigm, attentional modulation of grouping processes over both space and time.—Y.Y.

## AUDITORY SCENE ANALYSIS

### Feature integration gets a hearing

Shamma, S. A., Elhilali, M., & Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. *Trends in Neurosciences*, 34(3), 114–123.

One of the enduring problems in visual perception is the 'binding problem'. Visual objects have multiple features of color, shape, motion, and so forth. These features are analyzed in different portions of the brain. How do they all get bound into an coherent object representation in which this color gets bound to this piece of the object while that color gets bound to that piece, and so on? Following Anne Treisman's pioneering work, many of the proposed solutions involve a role for attention. Without attention, we seem to lack explicit knowledge of what goes with what, even if the features come from the same spot in the visual field.

Now, consider the problem in audition. Auditory 'objects' have features like pitch, loudness, timbre, and location. However, when there are multiple sound sources in the environment all of those features from all of those objects form a single, complex waveform that enters the ear. How is the listener supposed to sort this out into an auditory world with distinct sound sources that are perceived with their true combinations of features? This is the problem of "auditory scene analysis", discussed by Shamma et al. (2011). In vision, the problem is made somewhat easier by the spatial layout of the stimuli. One object is at one location in the image and another, ignoring issues of occlusion, is at another location. Not so in audition where spatial location is a feature that must be derived from the decidedly non-spatial waveforms arriving at each ear. Moreover, compared to auditory stimuli, visual stimuli are delightfully stable over time. The relationship of the color and orientation of an object is likely to persist or, at worst, change relatively slowly as the object moves. Auditory stimuli are temporal stimuli. The mix of features that was present 100 msec ago is irretrievably gone

and has been replaced by a new mix. When there are multiple sources making multiple streams of sound, the binding problem becomes a problem of forming coherent streams out of ever-changing stimuli.

One good idea in auditory scene analysis is to, in effect, substitute tonotopy for spatiotopy. On the cochlea, different frequencies activate different spatially separated receptors, laid out in a tonotopic array and this tonotopy is preserved in parts of the auditory cortex as retinotopy is preserved in the visual cortex. Perhaps streams can be separated if they stimulate spatially separated populations of neurons. As Shamma et al. review, this appealing idea does work, but it fails to account for a number of phenomena. Most importantly for Shamma et al., it doesn't account for the role of temporal coherence. Visual features co-occur in space. Auditory features happen together in time. One could argue that this co-occurrence does all the work and that both visual and auditory binding can be solved "pre-attentively". This is not true in vision and, as Shamma et al. argue, it doesn't work in audition either. In audition, the act of attending can change the formation of auditory streams. They use the example of an orchestra that can be heard as a single thing or, with some attention, pulled apart into strings, woodwinds, and so forth. On a mechanistic level, they present evidence that attention can change the neuronal and behavioral responses to temporal coherence and, thus, the binding of auditory features. The senses use similar tricks to solve similar problems. Apart from its virtues as an advance in auditory research, the Shamma et al. article is a valuable example of how ideas from one modality can inform the study of others.—J.M.W.

## OLFACTION AND VISION

### What the nose sees?

Seigneuric, A., Durand, K., Jiang, T., Baudouin, J., & Schaal, B. (2010). The nose tells it to the eyes: Crossmodal associations between olfaction and vision. *Perception*, 39 (11), 1541–1554.

Different sensory modalities tend to be studied by different people. Researchers specialize in studying the visual system, the auditory system, and so forth. However, the brain integrates sensory information at various levels. While Proust wrote of "the visual memory which, being linked to that taste [the famous madeleine], has tried to follow it into my conscious mind", most of the modern work on cross-modal interaction has favored visual-auditory interactions. This is probably partly due to sheer demographics: vision and audition probably boast more researchers than the remaining senses combined. But mappings between these two senses are probably easier to

study as well, since they are both distal, spatial modalities. High-pitched tones can be easily matched to high spatial locations or high spatial frequencies (Evans & Treisman, 2010). Whatever the cause, we do not have a great deal of information about cross-modal interactions involving the other senses.

Seigneuric and colleagues (Seigneuric, Durand, Jiang, Baudouin, & Schaal, 2010) aim to ameliorate this lack, studying the connections between vision and olfaction. They presented observers with a visual scene comprising 12 objects on a tabletop. The objects were primarily food items (strawberries, bacon, coffee, fish, bananas, apricots, melons, vanilla beans), plus two flowers (rose and lavender) and a bar of soap. The authors then recorded oculomotor behavior during 30 s of free exploration of the image. Observers were told they would be asked questions about the image later, but not given any more specific instructions.

One clever wrinkle is that the odorants were hidden inside a cardboard box used as a chin-rest. The eye tracking equipment did not actually require a chin-rest; the box was simply an excuse to get the odorant placed near the observers' nostrils without their awareness. The odorants were diluted to subthreshold intensity. Few observers reported detecting an odor, and of those who did, none could recognize the odor. Each observer received a different odor, and contributed a single trial to the experiment.

The authors hypothesized that the subthreshold odor would increase the visual salience of the corresponding objects, manifesting in a reduced latency to first fixating the object. Additionally, the odor should prime recognition of the corresponding object, reducing fixation time (reductions in both cases relative to observers who received a different odor). The results supported their hypotheses: observers exposed to the (subthreshold) scent of, for example, bacon fixated the plate of bacon and eggs roughly 740 ms earlier than observers exposed to, say, vanilla. Once there, they spent 76 ms less fixating on the bacon than they would have otherwise.

The results suggest that crossmodal links between olfaction and vision can operate automatically, since observers were not aware of the odors and had no strategic reason to use them. The reduced fixation time is an interesting result. One might have expected observers to linger longer on the bacon while inhaling its smoky scent. However, the authors argue that the implicit task in free exploration is to recognize the objects, and that the odor facilitated more rapid recognition of its corresponding object.

The finding that imperceptible odors modulate oculomotor behavior is an intriguing beginning to the study of crossmodal links between olfaction and vision.—T.S.H.

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features *Journal of Vision*, 10(1), 6.1–12. doi:10.1167/10.1.6

## PERCEPTION AND ACTION

### What you see is not what you do

Spering, M., Pomplun, M. & Carrasco, M. (2011). Tracking without perceiving: A dissociation between eye movements and motion perception. *Psychological Science*, 22(2), 216–225.

Some of the most interesting scientific discoveries have been stumbled upon while investigators had designs on other goals. Such was the case for Spering, Pomplun and Carrasco who wisely put aside their primary investigation when initial results showed an unexpected divergence between eye movements and reports of motion perception. The team was studying the effects of adaptation to one component of a motion plaid presented dichoptically. Participants adapted to either a vertical, moving grating presented to one eye, or, a horizontal, moving grating presented to the other eye for 1.5 s, before a 500-ms, test presentation of both stimuli. Participants reported most often that they perceived only one component motion in the test, either the horizontal or vertical, replicating previous work on binocular rivalry (Wolfe, JM, 1984, *Vis Research*, 24, 471–478). What Spering and colleagues did differently was they simultaneously tracked eye movements. The surprising result was that on almost all the trials participants' eyes did not follow the perceived direction of motion but instead moved along the diagonal, consistent with the pattern motion that was the sum of the two components. Dissociation of eye movements and motion perception suggested that perception and action use different motion information. There are numerous studies demonstrating dissociations between perception and action in a variety of domains, but motion perception has usually been tightly linked to action, as both seem to stem from responses in the brain's motion center, MT/V5. What makes this result especially interesting is that the dissociation is not one merely from magnitude, such as a difference in speed, which might be expected from a difference in response gain of perceptual and action systems. Rather, the dissociation is from motion-direction, indicating that the two systems may differ in how motion information is integrated. After recognizing the importance of their discovery, Spering and colleagues conducted several experiments to generalize results across stimulus conditions, and to rule out alternative accounts including report bias and intentional eye movements. The results provide a convincing argument that motion information can be used differently by eye movement and perceptual systems, and open the doorway for more discoveries describing the differences between what we see and do.—A. E.S.

## OLFACTION

### Rethinking the nature of olfactory receptors

Franco, M.I., Turin, L, Mershin, A., & Skoulakis, E.M.C. (2011). Molecular vibration-sensing component in *Drosophila melanogaster* olfaction. *Proceedings of the National Academy of Sciences*, 108, 3797–3802.

Humans are capable of smelling approximately 100,000 different odors; however, the mechanism by which this vast array of odor molecules is decoded by the olfactory receptor neurons remains a mystery. One popular notion is that olfactory receptor neurons respond to the structure or shape of odor molecules. But this shape-detecting mechanism cannot explain why molecules, that have the same shape, smell different, or why molecules, that have different shapes, smell the same. Over the past decade, an alternative theory, which proposes that olfactory neurons respond not to shape but rather to molecular vibrations, has been gaining support. Molecular vibrations occur when the atoms of a molecule move in a periodic fashion. Although earlier versions of this theory were deemed physically implausible, recent work in physics has validated the possibility. More recently, Franco et al. (2011) provided critical empirical support for the vibration theory of olfaction based on the olfactory abilities of fruit flies. A key innovation in these studies involved creating two molecules with identical shape, but with different vibrations. This was accomplished by selectively replacing the hydrogen atoms of one odorant with deuterium atoms. Deuterium is an isotope of hydrogen and is also called heavy hydrogen, to reflect the extra neutron. Because molecular vibrations can be altered by the density of a particular atom, a “deuterated” molecule can have the same shape as the corresponding hydrogen-only molecule, while also vibrating at a different frequency. If odor quality is determined primarily by shape, then the deuterated and hydrogen-only molecules should be indistinguishable. In contrast, if odor quality is determined by vibration, then the two molecules should be distinguishable. The researchers chose to test these predictions using the fruit fly so as to control the prior odor experiences and abilities of the subjects. With this in mind, the researchers provided a compelling array of behavioral evidence that fruit flies can indeed distinguish between deuterated and hydrogen-only odorants. For instance, when the deuterated molecule was associated with shock, the flies subsequently avoided the deuterated molecule, but not the hydrogen-only molecule (and *vice-versa*). To ensure that the flies were in fact using olfaction to distinguish between the two-odorant molecules (as opposed to some other sense), the researchers repeated this aversive conditioning experiment using fruit flies that were genetically mutated so that they could not smell.

Consistent with the vibration theory, these “anosmic” flies could no longer selectively avoid the deuterated molecule. Altogether, these and other important findings were interpreted to be inconsistent with a shape-only model of smell; rather, the findings were interpreted to support the existence of molecular vibration-sensing receptors, at least in fruit flies. The relevance of these findings to human olfactory abilities awaits further research—BSG.

## GENETICS

### Vision, awareness and genes

Miller SM, Hansell NK, Ngo TT, Liu GB, Pettigrew JD, Martin NG, Wright MJ (2010). Genetic contribution to individual variation in binocular rivalry rate. *Proceedings of the National Academy of Sciences*, 107, 2664–2668.

Many visual stimuli are bistable; that is, they can be perceived in either of two ways. A classic example is the Necker cube. When one stares at such a stimulus for a long period of time, one’s percept typically alternates between the two possible percepts. Such stimuli have been of great interest in the area of consciousness studies because the state of consciousness they evoke can vary sharply over time while the physical stimulus stays the same. Several recent studies provide striking support for the claim that the rate of perceptual alternation evoked by bistable stimuli is largely determined by one’s genes. Miller et al., 2010, focused specifically on binocular rivalry. Their participants viewed stimuli in which a vertical grating moving to the left was presented to one eye and a horizontal grating moving down was presented to the other eye. Such a stimulus evokes alternating percepts of the vertical and horizontal gratings. For a given observer, the binocular rivalry alternation rate tends to be fairly regular; however, alternation rates differ dramatically across observers. Bipolar patients are interesting in this connection because they have slower binocular rivalry alternation rates than controls (Miller et al., 2003). This finding in conjunction with the substantial heritability of bipolar disorder led Miller et al. (2010) to investigate the possibility that binocular rivalry alternation rate might also be partially genetically determined. To investigate this question, Miller et al. (2010) measured the binocular rivalry alternation rate for monozygotic (MZ) vs dizygotic (DZ) twin pairs and found that alternation rates were significantly more highly correlated between MZ twins than between DZ twins. Recently, this result has been extended by Shannon et al., 2011, who measured both the rates of alternation for both binocular rivalry as well as for the bistable percepts evoked

by a Necker cube. These alternation rates were measured for samples of both MZ and DZ twin pairs. The results using the Necker cube exactly paralleled those obtained for binocular rivalry. In each case, alternation rates were significantly more highly correlated between MZ twins than between DZ twins. Although these findings show convincingly that perceptual alternation rate is determined in part by one’s genes, the neural process controlling perceptual alternation rate remains unknown.—C.F.C

Miller SM, Gynther BD, Heslop KR, Liu GB, Mitchell PB, Ngo TT, Pettigrew JD, Geffen LB (2003) Slow binocular rivalry in bipolar disorder. *Psychol Med* 33:683–692.

Shannon RW, Patrick CJ, Jiang Y, Bernat E, He S (2011) Genes contribute to the switching dynamics of bistable perception. *Journal of Vision* 11(3): 8; doi:10.1167/11.3.8.

## NEURAL PLASTICITY

### Language in the blind

Bedny, Pascual-Leone, Dodell-Feder, Fedorenko, & Saxe (2011). Language processing in the occipital cortex of congenitally blind adults. *PNAS*, 108(11), 4429.

The neural substrates for language are often thought to have evolved specifically to process its unique properties. The complexity and combinatorial power of language is thought to require specialized brain regions, traditionally identified as including left frontal and temporal regions. These regions presumably have neural architectures that have become perfectly suited for the computational challenges of language. Despite the evidence for specialization, considerable neural plasticity has also been found. In individuals who are congenitally blind, for example, language-related tasks can result in activation in occipital cortex, in areas devoted primarily to visual processing. Bedny et al. found that this neural activation in occipital regions in the congenitally blind is linked specifically to language processing and not to other cognitive or memory processes. Both sighted and congenitally blind participants were asked to engage in a variety of language tasks, ranging from those requiring phonological and lexical-semantic processing to tasks requiring sentence-level combinatorial processing. These tasks were designed to isolate different components of language processing and to contrast linguistic processing with general auditory, perceptual, and cognitive analysis. Neuroimaging data revealed that the patterns of activation found in occipital regions for congenitally blind participants were similar to patterns

found in classic language areas for sighted participants. The results showed increased neural activation associated with each type of linguistic processing in left occipital cortex in the congenitally blind, even when compared to cognitively demanding control tasks. The authors concluded that as a result of early experience, areas typically devoted to visual processing can be recruited for language related processing, suggesting considerable plasticity in functional localization of linguistic processing.—L.C.N

## STATISTICAL SUMMARY PERCEPTION

### Visual averaging is slow

Whiting & Oriet (in press). Rapid averaging? Not so fast! PSYCHON B REV.

Many recent studies have demonstrated that the human visual system can rapidly extract basic summary statistics from simple scenes. For example, when viewing briefly presented scenes containing a large number of circles with different diameters, observers can report the mean diameter of those circles. Such ‘rapid averaging’ findings have been taken as evidence for a statistical summary representation of a scene, and such a representation could be useful in providing something akin to the gist of a complex scene.

Research on statistical summary perception has proposed that these summaries are derived automatically by analyzing display items preattentively and in parallel across the visual field. However, Whiting and Oriet present new results that suggest that rapid averaging might occur less automatically than previously claimed. In their paper, Whiting and Oriet note that previous studies of rapid averaging used brief exposure durations to argue for a rapid, automatic process; the displays in typical studies are usually unmasked, which allows observers to continue

processing the displays after the stimuli disappear. Further, there may be subtle information across trials that could influence estimates of the average on any specific trial: The average of many individual trial averages (the cumulative mean) could be used to discriminate two test circles on any particular trial. For example, when shown two test circles, observers may be more likely to choose the one closer to the cumulative mean.

To address these possibilities, Whiting and Oriet presented observers with displays containing circles of different diameters. The displays were presented for various durations, and the displays were either masked or unmasked. Finally, the tested items—two circles, one that was the average diameter of those in the display and one that differed—could be near or far from the cumulative mean, which was tested by varying the type of distributions (rectangular or normal) across the displays.

The results suggested that when observers could rely on the cumulative mean (i.e., when the trial mean was near the cumulative mean), observers accurately reported the trial mean in both masked and unmasked trials. However, when observers could only rely on trial-level information only because displays were masked and when the cumulative mean was closer to the distractor, observers were at chance in discriminating the trial average from a distractor.

These findings suggest that previous studies have likely overestimated the speed of statistical averaging. One interesting implication of Whiting and Oriet’s findings is that statistical summaries might exist over different spatial and temporal scales. Observers might be unable to accurately extract trial statistics because statistics have been constantly computed for the previous trials. Whiting and Oriet’s procedure might have interesting implications for separating cumulative statistics and trial statistics, perhaps by using different distributions at the trial level and at the cumulative, cross-trial level.—S.P.V