Hearing in a "Moving" Visual World:

Coordinate Transformations Along the Auditory Pathway

Shawn M. Willett, Jennifer M. Groh, Ross K. Maddox

S.M. Willett (corresponding author)
Department of Neurobiology, and Center for Cognitive Neuroscience,
Duke University, 308 Research Drive, LSRC B203 Box 90999, Durham, NC, 27708
E-mail: shawn.willett@duke.edu phone: (919) 684-6729

J.M. Groh
Departments of Neurobiology, Psychology and Neuroscience, and Center for Cognitive Neuroscience,
Duke University, 308 Research Drive, LSRC B203 Box 90999, Durham, NC, 27708
E-mail: jmgroh@duke.edu phone: (919) 681-6536

R.K. Maddox
Departments of Neuroscience and Biomedical Engineering, and Del Monte Institute for Neuroscience
University of Rochester, 201 Robert B. Georgen Hall, Box 270168, Rochester, NY, 14627
E-mail: ross.maddox@rochester.edu phone: (585) 275-1835

In press, Springer Handbook for Auditory Research

Keywords:

Reference frame, multisensory integration, multimodal, eye-centered, hybrid, head-centered, superior colliculus, inferior colliculus, intraparietal cortex, frontal eye field, sound localization

Abstract

This chapter reviews the literature on how auditory signals are transformed into a coordinate system that facilitates interactions with the visual system. Sound location is deduced from cues that depend on the position of the sound with respect to the head, but visual location is deduced from the pattern of light illuminating the retina, yielding an eye-centered code. Connecting sights and sounds originating from the same position in the physical world requires the brain to incorporate information about the position of the eyes with respect to the head. Eye position has been found to interact with auditory signals at all levels of the auditory pathway that have been tested, but usually yields a code that is in a hybrid reference frame – neither head- nor eye-centered. Computing a coordinate transformation, in principal, may be easy, which could suggest that the looseness of the computational constraints may permit hybrid coding. A review of the behavioral literature addressing the effects of eye gaze on auditory spatial perception and a discussion of its consistency with physiological observations concludes the chapter.

1 Introduction

No sensory system is an island. The auditory system works in concert with other sensory systems to help organisms understand the events occurring in their environments. The process of integrating sensory information from different senses usually proceeds so seamlessly that animals are not aware of it, and it only becomes obvious in cases where the brain is swayed by one sense to overlook the evidence in another sense. Two classic audio-visual examples involve ventriloquism, in which sounds are erroneously perceived as coming from the mouths of puppets, and the McGurk effect (McGurk and Macdonald 1976), in which the sound /bɑ/ is dubbed to a video of a person mouthing /ga/, leading to a non veridical perception of /dɑ/ (see Lee, Maddox, Bizley, Chap. 4 for in-depth discussion of these multisensory illusions).

Illusions such as these reflect a deep intertwining of sensory pathways, with communication occurring between the pathways at multiple levels and taking multiple forms. In the case of interactions between hearing and vision specifically, eye movements play a critical role. In humans and monkeys, the eyes move about three times per second, and cover about an 80 degree range of space. Every time the eyes move, the visual input stream is disrupted and shifted to a new location on the retina. In contrast, the auditory system's input depends on the locations of sounds with respect to the head and ears. Eye movements in relation to the head, then, prevent a simple static connection between the visual and auditory domains. Rather, one or both sensory systems must adjust its processing based on these eye movements in order to be able to communicate with the other system.

This chapter will review what is known about where and how this happens in the brain (Sect. 3-5) and its consequences for auditory perception (Sect. 6) and attention (Sect. 7).

2 The Why and How of Linking Visual and Auditory Signals in Space

Combining visual and auditory information can be useful to help resolve ambiguities in sensory input. In the McGurk effect for example, some phonemes are acoustically similar, such as /bɑ/ vs /gɑ/, or /fɑ/ vs. /sɑ/, but the lip movements associated with generating those sounds look very different. Thus, watching someone's lips move while listening to their speech can greatly facilitate comprehension. However, it is critical that the visually observed lip movements used to resolve auditory ambiguities belong to the person who is actually speaking. At a cocktail party with many talkers, determining which person's lips to associate with which person's voice is necessary to derive any benefit from lipreading.

In principle, this can be accomplished by matching stimuli by a common spatial origin, but the visual and auditory systems use very different methods of determining spatial location. The optics of the eye creates an image of the visual scene on the retina. For sound, the brain must deduce location by comparing sound loudness and timing differences across the two ears as well as from direction dependent spectral variations. These different methods mean that the original information available to the brain specifies locations in different reference frames. The retina provides the brain with information about the eye-centered location of visual stimuli. The cues on which sound localization are based provide information about the location of sounds with respect to the head and ears.

However, it is critical to note that although the *cues* are head-centered, *it does not follow that the brain representations are*. In fact, as shall be described further below (Sect. 3 and 4), there is no known auditory brain representation that appears to encode sound location in a strictly head-centered frame of reference. Rather, eye movements and the resulting changes in eye position with respect to the head and ears modulate auditory processing at multiple stages throughout the auditory pathway and in multiple ways.

3 Auditory Reference Frames in the Superior Colliculus

Interactions between eye movements and auditory processing were first found in the intermediate and deep layers of the superior colliculus (SC) of monkeys (Jay and Sparks 1984; Lee and Groh 2012) and cats (Populin and Yin 1998; Populin et al. 2004). Attention focused on the SC because it was known to play a role in guiding saccadic eye movements (Robinson

1972; Sparks 1975), which can be made to visual, auditory (Zahn et al. 1979; Zambarbieri et al. 1982), and tactile (Groh and Sparks 1996) targets. It was also known that the SC exhibited responses to auditory stimuli in anesthetized animals such as hamsters (Chalupa and Rhoades 1977), mice (Drager and Hubel 1975), and cats (Meredith and Stein 1986a; Meredith and Stein 1986b). Furthermore, stimulation studies (Robinson 1972) and recording studies involving visual stimuli (Mays and Sparks 1980) suggested that the SC likely used an eye-centered reference frame specifying the direction and amplitude of the eye movement necessary to look at the saccade goal. Jay and Sparks (1987b, 1987a) therefore postulated that the SC must convert auditory information, originally determined from head-centered cues, to an eye-centered reference frame to accurately move the eyes to auditory targets.

Answering this question required evaluating responses to sounds as a function of both their position with respect to the head and their position with respect to the eyes, i.e. with the eyes in several different positions with respect to the head (Fig. 1A). The shift in initial eye position is key because it forces the eye-centered and head-centered reference frames out of alignment. If both the eyes and head are oriented forward (or aligned in any direction), then the eye-centered and head-centered reference frames are in register, meaning no coordinate transformation is needed to accurately orient the eyes to a head-centered cue.

PLACE FIGURE 1 NEAR HERE

Jay and Sparks (1984, 1987a) were the first to implement this clever experimental manipulation of shifting initial eye position. They recorded the activity of single neurons while head restrained monkeys made saccades to visual (LED) or auditory (bandpass filtered noise) targets from different fixation positions (-24, 0, or 24 degrees in horizontal azimuth). They then mapped the receptive field of each neuron as a function of initial fixation location. If a neuron encoded auditory stimuli in a head-centered reference frame, then its responses should be governed by sound location with respect to the head regardless of eye position. A schematic of a perfectly head-centered cell is shown in Fig. 1B. A head-centered response pattern would have superimposed receptive fields if responses are plotted in head-centered space, but receptive fields would be shifted by the amount of the initial fixation if responses are plotted in eye-centered space. In contrast, in a perfectly eye-centered response pattern, receptive fields would be shifted by initial fixation if responses are plotted in head-centered space but superimposed if plotted in eye-centered space. A schematic of a perfectly eye-centered cell is shown in Fig. 1C.

Jay and Sparks actually found something between these two canonical cases. Specifically, they found that initial eye position affected the majority of auditory responses in the SC, but did not appear to produce perfectly eye-centered response patterns. The response of an example cell modulated by eye position is shown in Fig. 2A. Each column displays the activity of the same neuron in three different trial conditions. While the target remained at 20 degrees with respect to the head across trials, the monkey fixated at three different locations (-24°, 0° or 24°), meaning the target was at the same place in reference to the head but in three different locations in reference to the eyes. When the receptive fields for this cell are plotted in head-centered coordinates the responses are shifted. In contrast, when the receptive fields for this cell are plotted in eye-centered coordinates the responses are more closely superimposed (Fig. 2B). These results indicate this particular cell’s response depended not only on the position of the auditory target with respect to the head but also on the position of the eyes in the orbit. Since the fixation locations were 24° apart, one would predict that if auditory receptive fields of SC neurons are encoded in an eye-centered reference frame the receptive fields would shift 24°, which this particular example neuron appears to do. However, the population results reveal auditory responses were only shifted on average 12.9°; while the visual receptive fields were shifted on average by 21.7°, closer to the expected amount for an eye-centered reference frame (Fig. 2C). The auditory receptive field shift is only about one half of what is expected and indicates that auditory sensory responses in the SC are neither head- nor eye-centered, but rather are encoded in a hybrid reference frame. How, then, are primates able to accurately move their eyes towards auditory stimuli (Metzger et al. 2004)?

PLACE FIGURE 2 NEAR HERE

It took a study nearly three decades later to start unraveling this mystery. Lee and Groh (2012) advanced understanding of the coordinate transform by teasing apart the time course of activity in the SC (Fig. 3A). It had long been known that SC neurons exhibit both “sensory” and “motor” activity, defined as activity time-locked to the stimulus and the movement, respectively (Fig. 3B) (Wurtz and Goldberg 1972; Sparks 1978). Indeed, Jay and Sparks had demonstrated that the motor burst occurs for auditory saccades (Jay and Sparks 1987a), but they had not quantitatively analyzed the coordinate frame of these signals. Using essentially the same procedures as Jay and Sparks but analyzing the motor period as well as the sensory period, Lee and Groh (2012) found that the auditory reference frame evolved over time. In the sensory period the auditory reference frame was encoded in a hybrid reference frame, as previously reported by Jay and Sparks. However, in the motor period, auditory-evoked signals appear to correspond to target location in an eye-centered reference frame. The population results are shown in Fig. 4, which plots index values reflecting how well head-centered vs. eye-centered reference frames describe the activity for each neuron in the population. Neurons exhibiting predominantly eye-centered signals are plotted in orange and cluster below the unity line, whereas neurons exhibiting predominantly head-centered signals are plotted in blue and cluster above it. For visual signals, both sensory- and motor- periods are dominated by eye-centered signals. In stark contrast, for auditory signals, the sensory period is predominantly hybrid but the motor period is dominated by eye-centered response patterns. This shift from hybrid encoding in the sensory period to more eye-centered encoding in the motor period of auditory stimuli likely allows for accurate saccades to auditory targets regardless of initial eye position (Metzger et al. 2004).

PLACE FIGURES 3 AND 4 NEAR HERE

The intermediate and deep layers of the SC comprise a comparatively “late” sensory structure, situated well on the oculomotor side of the brain’s sensorimotor continuum. Since the auditory reference frame already appears to be hybrid in the SC, where does the process of adjusting the auditory reference frame begin? The SC receives inputs from four structures with auditory activity: parietal cortex, frontal eye fields, auditory cortex, and the inferior colliculus (Sparks and Hartwich-Young 1989). The following sections outline what is known about the auditory reference frame in these structures.

4 Reference Frames Throughout the Brain

4.1 Reference Frames in Parietal and Frontal Cortex

Parietal cortex is known to exhibit activity related to both auditory (Stricanne et al. 1996; Linden et al. 1999) and visual cues, as well as eye and limb movements (Andersen and Buneo 2002), and is thought to play a role in spatial processing (Mullette-Gillman et al. 2005, 2009). Early studies from Andersen and colleagues indicated that changes in eye position affected visual signals in parietal cortex (Andersen and Mountcastle 1983; Andersen et al. 1985). These studies originally characterized the representation as eye-centered with eye position contributing to the gain of the response; however, the study design involved confounds that rendered the reference frame ambiguous (Mullette-Gillman et al. 2009). A more recent quantitative analysis indicated that at least in the banks of the intraparietal sulcus, this visual representation was essentially a hybrid between eye- and head-centered coordinates (Mullette-

Gillman et al. 2005, 2009). This finding was exciting from the standpoint of visual-auditory integration, because it suggested some “movement” of the visual reference frame to meet auditory signals in a common middle ground. Indeed, the auditory signals, while weaker and less prevalent than the visual signals, also showed eye position effects and the net result was a hybrid reference frame similar to the visual reference frame. Unlike in the SC, this reference frame was stable across time and did not become eye-centered at the time of the saccade (Mullette-Gillman et al. 2009).

Much like the SC, the frontal eye fields (FEF) are integral to generating eye movements (Robinson and Fuchs 1969; Schiller et al. 1979) to visual cues (Mohler et al. 1973; Schiller et al. 1980). However, until the mid-1990s, it remained unknown if the FEF contributed to the generation of saccades to auditory stimuli. Russo and Bruce (1994) investigated the response of FEF neurons to auditory and visually evoked saccades from multiple initial fixation positions. Although Russo and Bruce found the responses of FEF neurons vary with changes in eye position for both modalities, they did not quantitatively investigate which frame of reference these neurons used to encode space. More recently, studies have indicated that auditory signals in FEF are largely encoded in a hybrid reference frame in both sensory and motor periods (Caruso, Pages, Sommer, Groh, preliminary results). While this might seem inconsistent with a native visual eye-centered reference frame, the available evidence indicates that in the FEF, the visual code is only about 60% eye-centered (Sajad et al. 2015; Caruso et al. 2017). These data suggest that visual and auditory signals in the FEF are similar to each other but do not employ a completely pure eye- or head-centered coordinate frame. The coding of auditory cues in eye-centered coordinates thus appears to be uniquely reflected in the motor response of SC neurons.

4.2 Reference Frames in the Auditory Pathway: Inferior Colliculus and Auditory Cortex

The studies described so far concern auditory signals in association cortex or the oculomotor system. These areas could encode auditory stimuli in hybrid coordinates due to hybrid input from the auditory system or they could receive a head-centered input and transform it to a more hybrid reference frame. In what reference frame, then, do auditory areas encode auditory stimuli?

It is intuitive to assume neurons in the auditory system would use a head-centered reference frame to encode target location because the cues originally needed to compute auditory target location in space are head-centered, relying on differences in level and timing of the sound between the two ears. However, current evidence, so far, does not bear this theory out. Several studies investigating core auditory cortex (Werner-Reiss et al. 2003; Fu et al. 2004), belt auditory cortex (Maier and Groh 2010), and inferior colliculus (Groh et al. 2001; Zwiers et al. 2004), identified numerous examples of neurons sensitive to the combination of sound and eye position (Porter et al. 2006, 2007). In aggregate, the reference frame of signals in both structures is generally hybrid, similar to the SC (sensory phase), parietal cortex, and FEF. These data are shown in Fig. 5, using a displacement index. Values of 0 correspond to head-centered, values of 1 indicate eye-centered, and values of 0.5 indicate an intermediate or hybrid reference frame in which head- and eye-centered information is equally weighted. Both the auditory cortex (yellow bars) and the inferior colliculus (orange bars) have a mean distribution centered around a displacement index score 0.5, showing both regions encode auditory targets with a hybrid reference frame, similar to those used in parietal cortex (pink bars) and during the sensory phase of the SC response (blue bars).

PLACE FIGURE 5 NEAR HERE

This finding is surprising because it shows that the auditory system, largely thought to be independent of the visual and oculomotor systems, is in fact quite dependent upon eye position, with auditory signals encoded in a hybrid reference frame throughout the brain, except for the SC's eye-centered encoding of auditory evoked saccades (Fig. 5). If there are any regions possessing a solely head-centered reference frame, they would need to be prior to the inferior colliculus in the auditory pathway. The reference frame of these areas, principally the lateral lemniscus, the superior olivary complex and cochlear nucleus have yet to be probed and require further investigation.

5 Why Hybrid? Some Insights from Models of Coordinate Transformations

An enduring mystery is why a hybrid reference frame, the most commonly observed scenario, may be useful to the brain. Insights into this question can be gleaned from considering the computational steps involved in transforming signals from one coordinate system to another.

The first models for coordinate transformations of auditory signals were developed by Groh and Sparks (1992). Their vector subtraction model capitalized on the comparatively straightforward mathematics of computing a coordinate transformation. If the brain possesses a signal representing sound location with respect to the head (for example, suppose there is a sound at 24° to the right with respect to the head) and another signal representing eye position with respect to the head (e.g. the eyes might be at 10° to the right with respect to the head), then subtraction of the eye position signal from the sound-with-respect-to-the-head signal (24° - 10°) will yield a signal of sound-with-respect-to-the-eyes (the sound is 14° to the right with respect to the eyes). This core computation forms the crux of the model and is accomplished through subtracting eye position information via an inhibitory synapse.

At the time this model was created, little was known about how sound location was encoded in the primate brain. As noted above, subsequent work has failed to identify any brain regions that encode sound location with respect to the head in studies that explicitly manipulate eye position, so the potential inputs to this coordinate transformation remain unknown. However, another aspect of auditory coding does support this model: the use of rate coding, in which the activities of auditory neurons are monotonically related to the horizontal component of sound location. This coding format has now been identified in the inferior colliculus of primates (Groh et al. 2003) as well as other mammals (McAlpine and Grothe 2003), the auditory cortex of monkeys (Woods et al. 2006; Werner-Reiss and Groh 2008), and cats (Middlebrooks et al. 1994; Middlebrooks et al. 1998), and the primate SC (Lee and Groh 2014). Given that eye position signals are also rate coded (Fuchs and Luschei 1970; Luschei and Fuchs 1972), this suggests that, indeed, the core computation of subtraction can be easily accomplished using known signal types. Relatedly, other computational modeling studies argued that a hybrid code can serve as a basis from which any coordinate transformation may be computed (Pouget and Sejnowski 1997; Deneve et al. 2001). But since the same can be said of inputs in pure coordinates (Groh and Sparks 1992), these models may better be interpreted as how the nervous system computes reference frames as opposed to why it implements any particular reference frame.

The ease of this computation may provide an explanation for why the hybrid format is used. Specifically, the computation may be so easy that it is underconstrained. Neural populations are not under strong selective pressure to produce a purely eye-centered code until the point at which a specific behavioral response requires it – namely the eye-centered coding of a saccadic motor command in the SC (Lee and Groh 2012).

6 Behavioral Investigations of Gaze's Effect on Auditory Reference Frame

The presence of hybrid signals has also led to considerable interest in whether there are behavioral signatures of this computation. The rationale is as follows: if signals that are in a hybrid code are read out under the erroneous assumption that they are actually either purely head-centered or purely eye-centered, then they should produce errors in sound localization that depend on the position of the sound with respect to the eyes. These errors might vary in magnitude but should be intermediate between the two reference frames. That is, an eccentric fixation position of 20 degrees might be associated with errors in sound localization of 10 degrees. Accurate sound localization would only occur if the eyes were in some privileged position that brought the two reference frames into alignment.

It is readily apparent that this is not the case. Broadly speaking, one's ability to localize sounds is not obviously impaired when the eyes are directed to some position away from straight ahead (if it were, locating a ringing cell phone would prove quite difficult). In monkeys, for which hybrid coding is well observed physiologically, the accuracy of saccades to auditory targets is not adversely affected by starting from different initial fixation positions (Metzger et al. 2004). To be sure, initial fixation position does affect the final eye position for a given target, but this effect is comparable to the undershooting observed for saccades to visual targets – in which saccades typically travel about 90% of the distance to a target. Indeed, many studies with human subjects have reported effects of eye position on various types of sound localization tasks. However, the magnitude of these effects is modest under natural viewing conditions in which the eyes move frequently, and may only become large when subjects maintain fixation eccentrically for minutes at a time, as was done in some of these studies. This section and Sect. 7 review those studies and then discuss whether they are consistent with the reported neurophysiology.

In binaural lateralization studies with short fixations and frequent eye movements, the effect of gaze on auditory localization appears to depend on the specifics of the paradigm, but the majority of studies find small shifts (less than 10% of the gaze magnitude) that are actually in the opposite direction of gaze. Lewald and Ehrenstein (1996) asked subjects to adjust interaural level difference over earphones while they maintained an eccentric fixation, finding that auditory lateralization shifted away from gaze by 1–3 dB. In a series of follow up studies, Lewald found again that localization of sounds shifted away from eye position as long as there was an absolute visual reference to compare against the location of the sound (Lewald 1997, 1998). Complicating matters, Lewald (1997) notes that with eccentric eye position, both the perception of a central sound and a central visual stimulus shift away from gaze. That is, if eyes are fixated in the right hemifield, central stimuli appear to shift into the left hemifield. Importantly, if the shift of the visual reference exceeds the shift of the probe sound (the sound to be localized), it could cause the subject's response to shift towards eye position – accounting for mixed results. Lewald and Getzmann (2006) found that horizontal (as well as vertical) auditory localization, again, shifted in the opposite direction as gaze, and Lewald and Ehrenstein (2001) found that the shift was also away from gaze in rear space. In other words, horizontal localization shifts are reflected about the coronal plane, as opposed to rotated 180° in azimuth. This result makes sense, since horizontal binaural cues undergo the same transformation (reflection rather than rotation). It is thus safe to say that, depending on the specifics of the paradigm, the work of Lewald and colleagues generally finds a modest shift (about 2–4°) in auditory localization in the direction opposite eccentric eye gazes of 45°.

A number of other studies have investigated the effects of gaze on fixation with what seems to be an important experimental difference: lateral fixations were maintained over long periods, multiple seconds to minutes, rather than redirecting gaze on each trial. These studies have more consistent results and typically find a larger effect, around 40% of the gaze magnitude, but in the same direction as gaze (in contrast to the studies discussed above). Weerts and Thurlow (1971) found that when subjects expected the probe sound to come from a visible loudspeaker at ±20° azimuth, localization shifted by 4–8° in the direction of gaze. Through further manipulations they determined that lateral gaze with no expectation of stimulus origin resulted in smaller shifts on the order of 2°, but that expectation on its own resulted in no shift at all, demonstrating that gaze direction and subject expectation yielded the biggest overall localization shift. A number of follow up studies confirmed these results of localization shifts towards eye position (Bohlander 1984). Razavi et al. (2007) showed that those shifts increased with fixation duration, approaching large steady-state shifts after several minutes of around 8° when fixation was 20° to the side. Notably, they tested many sound locations and found the shift to be largely consistent across auditory space. Looking at both horizontal and vertical localization, Cui et al. (2010b) found a shift in the same direction as gaze, with a similar asymptotic time course to other studies from that group (Razavi et al. 2007; Cui et al. 2010a). This is in accord with a previous study that tested only vertical localization and vertical gaze offsets (Getzmann 2002).

In short, the difference in the sequencing of fixations from trial to trial appears to be what is driving the differing results between sets of studies. Studies that employ naturalistic (i.e., short) fixations show only modest gaze-driven localization shifts in the opposite direction of gaze (Lewald 1997; Lewald and Getzmann 2006). This is consistent with daily human experience: there is little trouble registering the locations of what is heard with what is seen. Studies that employ non-naturalistic fixations (i.e. long, constant fixations, often for several minutes at a time) show larger localization shifts in the same direction as gaze (e.g. Razavi et al. 2007). These larger shifts were around 40% of the magnitude of eccentric gaze—consistent with the partial shift of a hybrid reference frame.

The mechanism that explains these results is not known, but could be the result of decay in the accuracy of the relevant signals across time. In particular, sensing of eye position relies at least in part on corollary discharge, or the copy of the motor command that was issued to bring the eyes to that location in space (Guthrie et al. 1983; Sommer and Wurtz 2008). Memory for such corollary discharge signals may decay across seconds to minutes, producing shifts in the perceived location of sounds with respect to the (mis-sensed) eyes. This idea is similar to proprioceptive drift of occluded limb position, in which 15–120 seconds after an occluded limb movement the limb drifts back towards the body (Wann and Ibrahim 1992). Such a model—an initial hybrid shift time locked to an eye movement that decays with continued fixation—allows the disparate behavioral localization results to be reconciled with physiological observations.

7 Effects of Gaze Direction on Auditory Spatial Acuity

Most of the behavioral studies of auditory-oculomotor reference frames have been based on looking for gaze-driven biases of absolute sound localization. As seen above, studies with natural fixation lengths find small biases, a fact that is in line with the general ease of registering visual and auditory space in daily life. A related but distinct test of spatial perception is the ability to discriminate subtle differences in two sounds' locations (or small changes to binaural lateralization cues). Human auditory spatial acuity is best in front of the head and worsens with increasing distance from the median plane (Hafter and Maio 1975; Middlebrooks and Onsan 2012). This is partly a physical acoustical effect (Mills 1958); however, auditory spatial discrimination is poorer for lateral sounds even when controlling for the acoustics, suggesting that the neural resolution of horizontal space coding is better for central locations than for lateral ones (Maddox et al. 2014).

Discrimination paradigms test relative rather than absolute localization, so why would they be useful for studying reference frames? If the neural representation of acoustical space is modified by eye gaze, then it is reasonable to expect that such a modification resulting from directed gaze may improve auditory spatial acuity. This improvement could occur by moving the experimental probe sounds into a part of the nonlinear receptive field with a steeper or shallower slope (e.g. Fig. 2B), improving or impairing discrimination performance, respectively.

Maddox et al. (2014) tested that notion by directing eye fixation with an eccentric dot and engaging subjects in a binaural cue discrimination task. They found that in blocks where gaze was held centered, discrimination performance was as expected for both interaural level difference (ILD) and interaural timing difference (ITD) cues: much better for central sounds than lateral ones (Fig. 6). However, when subjects directed their gaze towards the probe sounds, there was an improvement in discrimination of the lateral binaural cues of about 10% (there was also a smaller improvement for centered auditory probes). Such a result, in isolation, could be attributed to auditory attention—visual cues indicating the location of a masked speech stream do improve intelligibility (Best et al. 2007). Surprisingly, though, they found no such benefit in blocks where auditory spatial attention was directed with an acoustic primer, suggesting that eccentric fixation was a crucial component of the observation, and that neither spatial attention nor expectation was the main driver. The authors hypothesize that an enhancement to spatial acuity in the direction of gaze could lead to enhanced spatial release from masking (Marrone et al. 2008) when attempting to selectively attend to one sound that is physically close to another distracting sound, in a sense because the two sound sources become better separated in perceptual space. This would represent an advantage that was distinct from correct alignment of auditory and visual reference frames, one which would be in line with the notion that the major benefit afforded by binaural spatial hearing in many species is separation of competing sounds, rather than precise localization (Grothe and Pecka 2014).

PLACE FIGURE 6 NEAR HERE

However, as with absolute localization as described in the previous section (Sect. 6), there is disagreement between studies for spatial discrimination. Again, the duration of fixation is a possible factor, albeit with opposite results. Wood and Bizley (2015) tested discrimination over a broad range of reference azimuths between ±75°. Their subjects maintained fixation over the course of an entire block of trials (lasting minutes at a time) at -30, 0, or 30°. They confirmed that performance was best near the median plane, but in this case they found no effect of gaze, in stark contrast to previous results (Maddox et al. 2014).

In short, these relative discrimination experiments serve as a convenient complement to the absolute localization experiments of the previous section (Sect. 6): here there is an effect for short fixations that seems to disappear for longer ones. This suggests that relative judgments and absolute judgments are accomplished using at least partly different mechanisms, and are differentially affected by the duration of fixation.

8 Summary and Future Directions

The auditory and visual systems work together to help animals understand the events happening in the environment. In species with mobile eyes, such as humans and monkeys, such movements must be factored in when comparing the locations of sounds to the locations of images of potential visual sources. The neurophysiological processes involved in this computation appear to span a wide range of brain regions. Although auditory location *cues* depend on the sound's position with respect to the head, no purely head-centered brain representation has yet been identified.

The perceptual implications of the neurophysiological findings remain somewhat unclear. On the whole, humans and monkeys are clearly able to localize sounds accurately despite movements of the eyes. However, perceptual errors that depend on eye position do occur, and can vary in direction and size depending on whether the fixation duration is short (<seconds) or long (>minutes) as well as whether the task involves absolute versus relative judgments.

A number of questions remain to be answered. Is there a purely head-centered reference frame for auditory stimuli anywhere in the brain? Where in the neural hierarchy does this occur? Where and how do eye position signals first reach the auditory pathway? Is the commonly observed hybrid reference frame a “bug” or a not-yet-understood feature? How does the process of integrating visual and auditory space accommodate the many simultaneous visual and auditory events that occur in natural scenes? Do the same neural mechanisms underlie eye movements, visual attention, and auditory spatial attention?

These questions can be addressed through continued integration of physiology, behavior, and modeling in this computationally rich system. That the brain somehow manages to weave together information derived from two distinct physical phenomena using completely different sensors in dynamically misaligned physical reference frames is a truly remarkable feat that goes unnoticed in daily life.

Compliance with Ethics Requirements

Shawn M. Willett declares no conflicts of interest.

Jennifer M. Groh declares no conflicts of interest.

Ross K. Maddox declares no conflicts of interest.

References

Andersen, R. A., & Mountcastle, V. B. (1983). The influence of the angle of gaze upon the excitability of the light- sensitive neurons of the posterior parietal cortex. *The Journal of Neuroscience, 3*(3), 532-548.

Andersen, R. A., & Buneo, C. A. (2002). Intentional maps in posterior parietal cortex. *Annual Review of Neuroscience, 25*(1), 189-220.

Andersen, R. A., Essick, G. K., & Siegel, R. M. (1985). Encoding of spatial location by posterior parietal neurons. *Science, 230*(4724), 456-458.

Best, V., Ozmeral, E. J., & Shinn-Cunningham, B. G. (2007). Visually-guided Attention Enhances Target Identification in a Complex Auditory Scene. *Journal for the Association for Research in Otolaryngology, 8*(2), 294-304.

Bohlander, R. W. (1984). Eye Position and Visual Attention Influence Perceived Auditory Direction. *Perceptual and Motor Skills, 59*(2), 483-510.

Caruso, V., Pages, D. S., Sommer, M., & Groh, J. M. (2017). Beyond the labeled line: variation in visual reference frames from intraparietal cortex to frontal eye fields and the superior colliculus. *bioRxiv*, 124628.

Chalupa, L. M., & Rhoades, R. W. (1977). Responses of visual, somatosensory, and auditory neurones in the golden hamster's superior colliculus. *The Journal of Physiology, 270*(3), 595-626.

Cui, Q. N., O’Neill, W. E., & Paige, G. D. (2010a). Advancing age alters the influence of eye position on sound localization. *Experimental Brain Research, 206*(4), 371-379.

Cui, Q. N., Razavi, B., O'Neill, W. E., & Paige, G. D. (2010b). Perception of Auditory, Visual, and Egocentric Spatial Alignment Adapts Differently to Changes in Eye Position. *Journal of Neurophysiology, 103*(2), 1020-1035.

Deneve, S., Latham, P. E., & Pouget, A. (2001). Efficient computation and cue integration with noisy population codes. *Nature neuroscience, 4*(8), 826.

Drager, U. C., & Hubel, D. H. (1975). Responses to visual stimulation and relationship between visual, auditory, and somatosensory inputs in mouse superior colliculus. *Journal of Neurophysiology, 38*(3), 690-713.

Fu, K.-M. G., Shah, A. S., O'Connell, M. N., McGinnis, T., Eckholdt, H., Lakatos, P., Smiley, J., & Schroeder, C. E. (2004). Timing and Laminar Profile of Eye-Position Effects on Auditory Responses in Primate Auditory Cortex. *Journal of Neurophysiology, 92*(6), 3522-3531.

Fuchs, A. F., & Luschei, E. S. (1970). Firing patterns of abducens neurons of alert monkeys in relationship to horizontal eye movement. *Journal of Neurophysiology, 33*(3), 382-392.

Getzmann, S. (2002). The effect of eye position and background noise on vertical sound localization. *Hearing Research, 169*(1–2), 130-139.

Groh, J. M., & Sparks, D. L. (1992). Two models for transforming auditory signals from head-centered to eye-centered coordinates. *Biological Cybernetics, 67*(4), 291-302.

Groh, J. M., & Sparks, D. L. (1996). Saccades to somatosensory targets. I. behavioral characteristics. *Journal of Neurophysiology, 75*(1), 412-427.

Groh, J. M., Kelly, K. A., & Underhill, A. M. (2003). A Monotonic Code for Sound Azimuth in Primate Inferior Colliculus. *Journal of Cognitive Neuroscience, 15*(8), 1217-1231.

Groh, J. M., Trause, A. S., Underhill, A. M., Clark, K. R., & Inati, S. (2001). Eye Position Influences Auditory Responses in Primate Inferior Colliculus. *Neuron, 29*(2), 509-518.

Grothe, B., & Pecka, M. (2014). The natural history of sound localization in mammals – a story of neuronal inhibition. *Frontiers in Neural Circuits, 8*(116), 1-19.

Guthrie, B. L., Porter, J. D., & Sparks, D. L. (1983). Corollary discharge provides accurate eye position information to the oculomotor system. *Science, 221*(4616), 1193-1195.

Hafter, E. R., & Maio, J. D. (1975). Difference thresholds for interaural delay. *The Journal of the Acoustical Society of America, 57*(1), 181-187.

Jay, M. F., & Sparks, D. L. (1984). Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature, 309*(5966), 345-347.

Jay, M. F., & Sparks, D. L. (1987a). Sensorimotor integration in the primate superior colliculus. I. Motor convergence. *Journal of Neurophysiology, 57*(1), 22-34.

Jay, M. F., & Sparks, D. L. (1987b). Sensorimotor integration in the primate superior colliculus. II. Coordinates of auditory signals. *Journal of Neurophysiology, 57*(1), 35-55.

Lee, J., & Groh, J. M. (2012). Auditory signals evolve from hybrid- to eye-centered coordinates in the primate superior colliculus. *Journal of Neurophysiology, 108*(1), 227-242.

Lee, J., & Groh, J. M. (2014). Different stimuli, different spatial codes: a visual map and an auditory rate code for oculomotor space in the primate superior colliculus. *PLoS ONE, 9*(1), e85017.

Lewald, J. (1997). Eye-position effects in directional hearing. *Behavioural Brain Research, 87*(1), 35-48.

Lewald, J. (1998). The effect of gaze eccentricity on perceived sound direction and its relation to visual localization. *Hearing Research, 115*(1–2), 206-216.

Lewald, J., & Ehrenstein, W. H. (1996). The effect of eye position on auditory lateralization. [journal article]. *Experimental Brain Research, 108*(3), 473-485.

Lewald, J., & Ehrenstein, W. H. (2001). Effect of gaze direction on sound localization in rear space. *Neuroscience Research, 39*(2), 253-257.

Lewald, J., & Getzmann, S. (2006). Horizontal and vertical effects of eye-position on sound localization. *Hearing Research, 213*(1–2), 99-106.

Linden, J. F., Grunewald, A., & Andersen, R. A. (1999). Responses to Auditory Stimuli in Macaque Lateral Intraparietal Area II. Behavioral Modulation. *Journal of Neurophysiology, 82*(1), 343-358.

Luschei, E. S., & Fuchs, A. F. (1972). Activity of brain stem neurons during eye movements of alert monkeys. *Journal of Neurophysiology*.

Maddox, R. K., Pospisil, D. A., Stecker, G. C., & Lee, A. K. C. (2014). Directing Eye Gaze Enhances Auditory Spatial Cue Discrimination. *Current Biology, 24*(7), 748-752.

Maier, J. X., & Groh, J. M. (2009). Multisensory guidance of orienting behavior. *Hearing Research, 258*(1–2), 106-112.

Maier, J. X., & Groh, J. M. (2010). Comparison of Gain-Like Properties of Eye Position Signals in Inferior Colliculus Versus Auditory Cortex of Primates. *Frontiers in Integrative Neuroscience, 4*(121).

Marrone, N., Mason, C. R., & Kidd, G. (2008). Tuning in the spatial dimension: Evidence from a masked speech identification task. *The Journal of the Acoustical Society of America, 124*(2), 1146-1158.

Mays, L. E., & Sparks, D. L. (1980). Dissociation of visual and saccade-related responses in superior colliculus neurons. *Journal of Neurophysiology, 43*(1), 207-232.

McAlpine, D., & Grothe, B. (2003). Sound localization and delay lines – do mammals fit the model? *Trends in Neurosciences, 26*(7), 347-350.

McGurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746-748.

Meredith, A. M., & Stein, B. E. (1986a). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain research, 365*(2), 350-354.

Meredith, M. A., & Stein, B. E. (1986b). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology, 56*(3), 640-662.

Metzger, R. R., Mullette-Gillman, O. D. A., Underhill, A. M., Cohen, Y. E., & Groh, J. M. (2004). Auditory Saccades From Different Eye Positions in the Monkey: Implications for Coordinate Transformations. *Journal of Neurophysiology, 92*(4), 2622-2627.

Middlebrooks, J. C., & Onsan, Z. A. (2012). Stream segregation with high spatial acuity. *The Journal of the Acoustical Society of America, 132*(6), 3896-3911.

Middlebrooks, J. C., Clock, A. E., Xu, L., & Green, D. M. (1994). A panoramic code for sound location by cortical neurons. *Science, 264*(5160), 842-843.

Middlebrooks, J. C., Xu, L., Eddins, A. C., & Green, D. M. (1998). Codes for Sound-Source Location in Nontonotopic Auditory Cortex. *Journal of Neurophysiology, 80*(2), 863-881.

Mills, A. W. (1958). On the Minimum Audible Angle. *The Journal of the Acoustical Society of America, 30*(4), 237-246.

Mohler, C. W., Goldberg, M. E., & Wurtz, R. H. (1973). Visual receptive fields of frontal eye field neurons. *Brain research, 61*, 385-389.

Mullette-Gillman, O. D. A., Cohen, Y. E., & Groh, J. M. (2005). Eye-Centered, Head-Centered, and Complex Coding of Visual and Auditory Targets in the Intraparietal Sulcus. *Journal of Neurophysiology, 94*(4), 2331-2352.

Mullette-Gillman, O. D. A., Cohen, Y. E., & Groh, J. M. (2009). Motor-Related Signals in the Intraparietal Cortex Encode Locations in a Hybrid, rather than Eye-Centered Reference Frame. *Cerebral Cortex, 19*(8), 1761-1775.

Populin, L. C., & Yin, T. C. T. (1998). Sensitivity of auditory cells in the superior colliculus to eye position in the behaving cat. *Psychophysical and physiological advances in hearing*, 441-448.

Populin, L. C., Tollin, D. J., & Yin, T. C. T. (2004). Effect of Eye Position on Saccades and Neuronal Responses to Acoustic Stimuli in the Superior Colliculus of the Behaving Cat. *Journal of Neurophysiology, 92*(4), 2151-2167.

Porter, K. K., Metzger, R. R., & Groh, J. M. (2006). Representation of Eye Position in Primate Inferior Colliculus. *Journal of Neurophysiology, 95*(3), 1826-1842.

Porter, K. K., Metzger, R. R., & Groh, J. M. (2007). Visual- and saccade-related signals in the primate inferior colliculus. *Proceedings of the National Academy of Sciences, 104*(45), 17855-17860.

Pouget, A., & Sejnowski, T. J. (1997). Spatial Transformations in the Parietal Cortex Using Basis Functions. *Journal of Cognitive Neuroscience, 9*(2), 222-237.

Razavi, B., O'Neill, W. E., & Paige, G. D. (2007). Auditory Spatial Perception Dynamically Realigns with Changing Eye Position. *The Journal of Neuroscience, 27*(38), 10249-10258.

Robinson, D. A. (1972). Eye movements evoked by collicular stimulation in the alert monkey. *Vision Research, 12*(11), 1795-1808.

Robinson, D. A., & Fuchs, A. F. (1969). Eye movements evoked by stimulation of frontal eye fields. *Journal of Neurophysiology, 32*(5), 637-648.

Russo, G. S., & Bruce, C. J. (1994). Frontal eye field activity preceding aurally guided saccades. *Journal of Neurophysiology, 71*(3), 1250-1253.

Sajad, A., Sadeh, M., Keith, G. P., Yan, X., Wang, H., & Crawford, J. D. (2015). Visual–Motor Transformations Within Frontal Eye Fields During Head-Unrestrained Gaze Shifts in the Monkey. *Cerebral Cortex, 25*(10), 3932-3952.

Schiller, P. H., True, S. D., & Conway, J. L. (1979). Effects of frontal eye field and superior colliculus ablations on eye movements. *Science, 206*(4418), 590-592.

Schiller, P. H., True, S. D., & Conway, J. L. (1980). Deficits in eye movements following frontal eye-field and superior colliculus ablations. *Journal of Neurophysiology, 44*(6), 1175-1189.

Sommer, M. A., & Wurtz, R. H. (2008). Brain circuits for the internal monitoring of movements. *Annual Review of Neuroscience, 31*, 317-338.

Sparks, D. L. (1975). Response properties of eye movement-related neurons in the monkey superior colliculus. *Brain research, 90*(1), 147-152.

Sparks, D. L. (1978). Functional properties of neurons in the monkey superior colliculus: Coupling of neuronal activity and saccade onset. *Brain research, 156*(1), 1-16.

Sparks, D. L., & Hartwich-Young, R. (1989). The deep layers of the superior colliculus. *Reviews of Oculomotor Research, 3*, 213-255.

Stricanne, B., Andersen, R. A., & Mazzoni, P. (1996). Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. *Journal of Neurophysiology, 76*(3), 2071-2076.

Wann, J. P., & Ibrahim, S. F. (1992). Does limb proprioception drift? *Experimental Brain Research, 91*(1), 162-166.

Weerts, T. C., & Thurlow, W. R. (1971). The effects of eye position and expectation on sound localization. *Perception & Psychophysics, 9*(1), 35-39.

Werner-Reiss, U., & Groh, J. M. (2008). A Rate Code for Sound Azimuth in Monkey Auditory Cortex: Implications for Human Neuroimaging Studies. *The Journal of Neuroscience, 28*(14), 3747-3758.

Werner-Reiss, U., Kelly, K. A., Trause, A. S., Underhill, A. M., & Groh, J. M. (2003). Eye Position Affects Activity in Primary Auditory Cortex of Primates. *Current Biology, 13*(7), 554-562.

Wood, K. C., & Bizley, J. K. (2015). Relative sound localisation abilities in human listeners. *The Journal of the Acoustical Society of America, 138*(2), 674-686.

Woods, T. M., Lopez, S. E., Long, J. H., Rahman, J. E., & Recanzone, G. H. (2006). Effects of Stimulus Azimuth and Intensity on the Single-Neuron Activity in the Auditory Cortex of the Alert Macaque Monkey. *Journal of Neurophysiology, 96*(6), 3323-3337.

Wurtz, R. H., & Goldberg, M. E. (1972). Activity of superior colliculus in behaving monkey. III. Cells discharging before eye movements. *Journal of Neurophysiology, 35*(4), 575-586.

Zahn, J. R., Abel, L. A., Dell'Osso, L. F., & Daroff, R. B. (1979). The audioocular response: intersensory delay. *Sensory Processes, 3*(1), 60.

Zambarbieri, D., Schmid, R., Magenes, G., & Prablanc, C. (1982). Saccadic responses evoked by presentation of visual and auditory targets. *Experimental Brain Research, 47*(3), 417-427.

Zwiers, M. P., Versnel, H., & Van Opstal, A. J. (2004). Involvement of Monkey Inferior Colliculus in Spatial Hearing. *The Journal of Neuroscience, 24*(17), 4145-4156.

Fig. 1. Schematics of behavioral paradigm and hypothetical neurons with perfect eye or head-centered encoding. A). Monkeys typically fixated one of three locations and then made a saccade to either a visual or auditory target. B & C). Perfectly head-centered cells (B) and perfectly eye-centered cells (C) exhibit different behavior depending on how target location is defined. The left panel plots receptive fields in head-centered space whereas the right panel plots receptive fields in eye-centered space. A head-centered response pattern exhibits well aligned response curves across the different fixation patterns when target location is defined with respect to the head (B left), whereas the responses of an eye-centered response pattern align better when target location is defined with respect to the eyes (C right). (B & C modified from Maier and Groh 2009).

Fig. 2. Auditory responses in the SC as a function of eye position. A). The response of an SC cell to a sound at a fixed head-centered position (20° right with respect to the head; icon of speaker) while the monkey fixated three different positions (-24°, 0°, or 24° as shown in the left, center, and right panels). Sound and fixation positions are schematized in the top row of panels. The middle row indicates the horizontal and vertical eye traces of saccades to this target as well as the instantaneous firing rate of the neuron as a function of time on an individual trial. The bottom row shows a peri-stimulus time histogram for all the trials of that particular condition. This neuron fires much more strongly to this particular head-centered sound when initially fixating at -24° compared to 0° or 24°, consistent with an eye-centered frame of reference. B). A summary of the neuron's auditory response functions as measured at different fixation positions when plotted in head-centered coordinates (top) or eye-centered coordinates (bottom). The response functions are more shifted when plotted in head-centered space but more nearly superimposed in eye-centered space, indicating this neuron encodes sound location in an eye-centered reference frame. C). Population distributions for response function shifts. The average (indicated by arrows) of the auditory shift is 12.9° while the visual shift is 21.7° (A & B modified from Jay and Sparks 1984; C taken from Maier and Groh 2009).

Fig. 3. Analysis of sensory-related vs. motor-related SC activity. A). A schematic of the time course of the experiment. Monkeys fixated on a fixation LED after which a target (visual or auditory) was presented. After a delay, the fixation light was extinguished cuing the monkey to saccade to the target. Neural activity was evaluated prior to stimulus onset (baseline period, 500 ms), after target onset (sensory period, 500 ms) and around the time of the saccade (motor period, 20 ms prior to saccade onset to 20 ms after saccade offset). B). A population peri-stimulus time histogram showing the different phases of the SC populations response. The activity is the normalized to the cell's largest response and then averaged across the population and aligned to the stimulus onset (left 0) and the saccade onset (right 0). The SC population clearly displays a response to sensory and saccade onset which are referred to as the sensory phase and motor phase respectively. (Adapted from Lee and Groh 2012).

Fig. 4. Reference frames of SC cells during the sensory and motor periods to auditory and visual targets. The eye-centered correlation coefficient is a measure of how well response functions align in an eye-centered coordinate frame, and the head-centered correlation coefficient a measure of how well response functions align in a head-centered coordinate frame; -1 indicates perfect anticorrelation, 0 indicates no correlation, and 1 indicates perfect correlation with respect to the designated reference frame. Orange units are classified as eye-centered as the 95% confidence intervals on the head- and eye-centered correlation coefficients (crosshairs) exclude the head-centered reference frame. Blue units are classified as head-centered due to the exclusion of the eye-centered reference frame. Grey units are classified as hybrid as neither reference frame can be excluded. Overall, visual signals in the SC are strongly eye-centered, whereas auditory signals transition from mainly hybrid during the sensory period to mainly eye-centered during the motor period. The eye-centered auditory activity during the SC's motor period is the only place in the brain where a reasonably pure reference frame for auditory signals has been identified. (Modified from Lee and Groh 2012).

Fig. 5. Comparison of the reference frames for auditory and visual stimuli in auditory cortex (AC), the inferior colliculus (IC), lateral/medial intraparietal cortex (LIP), and the superior colliculus (SC). A displacement index values of 0 corresponds to a head-centered reference frame, a value of 0.5 indicates a hybrid reference frame, and a value of 1 indicates an eye-centered reference frame. Again, note the auditory motor period for the SC is the most eye-centered auditory representation measured. HC, head-centered; EC, eye-centered (Modified from Lee and Groh 2012).

Fig. 6. Behavioral paradigm and results showing a gaze-mediated improvement in auditory spatial acuity. (A) The time course of a trial. In visual trials, the dot brightened on fixation and darkened after 800 ms; in auditory trials, the primer was a noise burst. The probe noise bursts lasted 70 ms each, with 30 ms between each. The subject responded by button press any time after the stimulus. Primers provided the same timing information whether visual or auditory, directional or uninformative. (B) Experimental blocks are shown one per quadrant. Each quadrant shows an example of a center trial above a side trial. The positions of the visual or auditory primers, where present, are shown as gray and blue dots, respectively. In auditory trials, subjects were presented with a black screen and not instructed where to direct their eyes. The probe noise bursts are shown as orange bars of different lateralizations centered about the primer. For interaural level difference (ILD, C) and interaural time difference (ITD, D) stimuli, the subject performance is shown for all conditions. Center performance was better than side performance. For ILD, performance was better in visual directional trials than in visual uninformative trials at both the center and side positions. For ITD, directional visual trials showed improved discrimination when the stimulus was located on the side. Auditory primers offered no benefit. Error bars indicate ±1 SEM (across the 15 intrasubject means). Asterisks indicate one-tailed paired t test significance: *$p < 0.00625$, **$p < 0.00125$ (Bonferroni-corrected values of 0.05 and 0.01, respectively). Effect size (within-subjects Cohen’s d) of directional 2 uninformative contrast is bold where >0.5. Arcsine-transformed values were used for t tests and effect sizes; the means and error bars plotted are based on raw percent correct scores. (Adapted from Maddox et al. 2014).