2. Experiment 1: Audio-visual coherence

Experiment 1 was designed to answer the question whether humans integrate auditory and visual information to form a motor command for smooth pursuit eye movements. If auditory information was integrated and beneficial for the pursuit command, we would expect pursuit to require a lower visual signal for a given gain. This should have been reflected in lower PSE values if both information move coherently (Congruent condition). If the integration was based on the strength of motion rather than its direction, the same effect should have been observed, if both move in opposite directions (Incongruent condition). We included a condition with a stationary sound in order to check for effects of the presence of any auditory signal.

2.1. Methods

2.1.1. Participants

Participants were 8 undergraduate students from Giessen University (mean age = 24, 7 female). Participants gave informed consent and received course credit (5 out of 8) or 8€/h in return. Everyone had normal or corrected to normal vision.

2.1.2. Setup and stimuli

All stimuli have been displayed using the Psychtoolbox (Brainard, 1997; Pelli, 1997) in MATLAB (The Mathworks, Inc., Natick, MA, USA). Visual stimuli were gabor patches with a spatial frequency of sf= 1 cycles/deg and a stimulus size with a Gaussian SD of 0.7deg. Auditory stimuli were pure tones with a frequency of f=1000Hz. They were created and played with MATLAB using the

C. Wolf, *Audio-Visual Integration in Smooth Pursuit Eye Movements,* BestMasters,
DOI 10.1007/978-3-658-08311-3_2, © Springer Fachmedien Wiesbaden 2015

Psychtoolbox and an HDSPe AIO sound card (RME, Haimhausen, Germany). Sounds were displayed to subjects via Sennheiser HD 280 Pro headphones (Sennheiser electronic GmbH & Co. KG, Wedemark, Germany) and their horizontal location was computed using the interaural time difference (ITD). Auditory motion was simulated by dynamically changing the ITD. Eye movements of the right eye have been recorded with an EyeLink 1000 (SR Research, Osgoode, Ontario, Canada) at a viewing distance of 47cm at 1000Hz.

In session1, contrast values of target Gabor patches varied in five discrete steps from 0.007 to 0.05 Michelson contrast (0.007, 0.01, 0.02, 0.03, and 0.05) and moved with a velocity of 12deg/s. Additionally, there have been distractors moving at 8 or 16deg/s with 0.05 contrast. Gabor patches have been displayed on a neutral gray background with a luminance of 47.6 cd/m². Auditory velocities were adjusted within the experiment by two interleaved staircases for each target contrast and each distractor. Staircase values were updated in the logarithmic domain and had starting values of $10^{0.4}$ and $10^{1.6}$ (approximately 2.5 and 39.8deg/s) and a stepsize of 1.259^2 (about 2.5deg/s). In session2, contrast values and auditory velocities were individually determined from the results of session1 (see section 2.1.5). Contrast values varied logarithmically around the set contrast in seven values from half to twice the contrast.

2.1.3. Data and eye movement analysis

Eye velocity signals were retrieved by differentiating the eye position signal. We used the EyeLink 1000 saccade detection algorithm to determine saccade onset and offset. This algorithm uses a velocity threshold of 22deg/s and an acceleration threshold of 3800deg/s². For pursuit analysis, we removed saccades from velocity traces by linear interpolation. Velocity traces were then filtered by a moving average.

To determine pursuit onset, we fitted a regression line of 80ms to every sample. The best fitting regression line between 10 and 200deg/s² was selected and their interception with the x-axis determined pursuit onset (Schütz, Braun, & Gegenfurtner, 2007). Gain was computed as average velocity from 150ms after the onset (closed-loop) until the end of the trial or until 100ms before the eye reached a horizontal velocity of 0deg/s.

Data from session2 have been converted into dichotomic responses in order to compute oculometric functions. For each subject, we compared the gain of all traces against the median eye trace and converted each trial into a faster or slower response. For oculometric functions, a cumulative Gaussian has been fitted to the data using the psignifit toolbox version 2.5.6 for MATLAB (Wichmann & Hill, 2001a). To compare oculometric functions between the conditions, we applied repeated measures ANOVA whose p-values were corrected according to Greenhouse-Geisser if necessary. All eye movement analysis was executed offline using MATLAB. Inferential statistics were computed in SPSS (Version 21.0).

2.1.4. Procedure

Session1. Each trial consisted of a sequentially displayed visual and auditory stimulus in a randomized order. A fixation cross was displayed at the beginning of each trial to indicate participants that they can initiate a trial by pressing a button on a controller. The visual stimulus was a Gabor patch which appeared on the screen and was stationary for a randomized period of 500 to 1000ms before it started moving horizontally with a constant speed for a randomized duration between 750 and 1250ms (ramp paradigm). The auditory stimulus was a pure tone that was as well first displayed stationary (500-1000ms) and centered (i.e. no ITD) and then started moving for 750-1250ms with a constant velocity. Both stimuli have been separated by the fixation cross that reappeared for 500ms. Subjects were asked to pursue the visual as well as the auditory stimulus and indicate afterwards which of the two moved faster by a response on a keyboard. There have been 350 trials per subject, divided into 4 blocks.

Session2. The second session took place within one week after session1. Session2 was split up into four blocks, one for each condition. Each condition consisted of 180 trials and was defined by the level of audio-visual coherence. Auditory motion either went in the same direction as the Gabor patch (Congruent), in the opposite direction (Incongruent), did not move

(Static) or was not displayed at all (NoSound). The contrast of the Gabor patch as well as the auditory velocity were individually retrieved from session1. Again, a fixation cross indicated participants that they could initiate a trial by pressing a designated button on the controller. The Gabor patch appeared in the screen center, either on its own (NoSound) or simultaneously with a head-centered non-moving sound (Congruent, Incongruent and Static). After a randomized time between 500 and 1000ms, the Gabor patch started moving with a constant velocity (target: 12deg/s, distractor: 8 or 16deg/s) for a randomized time between 750 and 1250ms. Only in the congruent and incongruent condition, the auditory stimulus started moving simultaneously and for the same duration. To ensure that subjects pay attention to the auditory stimulus, they had to report the directional coherence auf audio and visual stimulus after each block.

2.1.5. Parameter estimation

The aim of session1 was to determine individual parameters for Gabor patch contrast as well as the judged auditory velocity. For each subject, we averaged the gain for each contrast and fitted a Weibull function to these values. From this Weibull function, we retrieved the contrast value that corresponds to a gain of 0.7.

For the auditory velocity of each target contrast value, we fitted psychometric functions for the proportion of “auditory motion faster” judgments over different auditory velocities and retrieved PSE and JND values. All auditory velocities were fitted in a log scale. The delogarithmized PSE determines the auditory velocity that is perceived as equally fast as the corresponding Gabor patch. We fitted a weighted linear regression to the 5 (logarithmized) PSE values over the different target contrasts. Residuals of each PSE were weighted by their JNDs. The value of the previously set contrast was entered into the resulting linear equation, which results in the log scaled auditory velocity applied in session2.

To set the auditory velocity for both of the two distractors, we applied the same procedure for the high contrast PSEs (8, 12, 16deg/s). We fitted a weighted regression over the different velocities and extracted the values for 8 and 16deg/s.

2.2 Results

Figure 1 shows an example trial data from session1. Whereas the oculomotor responses to a low contrast visual stimuli varied between low gain pursuit interleaved by saccades and high gain pursuit (Fig 1B), participants could not initiate pursuit with auditory motion only and thus made a sequence of saccades.

Figure 2 shows individual parameters from session1 for the set contrast (Fig 2A) as well as for the judged auditory velocity in log-coordinates (Fig 2B). A table with all individually set parameters can be found in the appendix. Contrary to our expectations, the judged auditory velocity decreases with increasing contrast (Fig 2B). Indeed, 6 out of 8 subjects showed a negative slope. However, the slopes did not significantly differ from zero ($t(7) = -1.03$, $p = .336$). We also fitted a weighted linear regression over the different velocities (high contrast) to determine the auditory velocities for the distractors. Here, the judged auditory velocity increased with increasing visual velocity for 7 out of 8 participants. Still, these slopes did not differ from zero either ($t(7) = 1.78$, $p = .118$).

The average gain in session2 was 0.71. Data from session2 has been converted into dichotomic oculometric data and fitted to psychometric functions. Example data from one subject (A) and across all subjects (B) is shown in Figure 3. Compared to the NoSound condition, PSEs in the sound conditions cluster around the same value with more PSEs shifted to higher values (Fig 3B). To test the hypothesis that audio-visual integration is beneficial for smooth pursuit, we performed a one-way repeated measures ANOVA on the individual PSEs across participants with the four different conditions as within-subject factor. The ANOVA did not show any significant differences between the conditions, $F(3,21) = 1.6$, $p = .22$.

We split up eye traces in time bins of 100ms and computed oculometric functions for each time bin (Fig 4). To test whether any effect of audio-visual

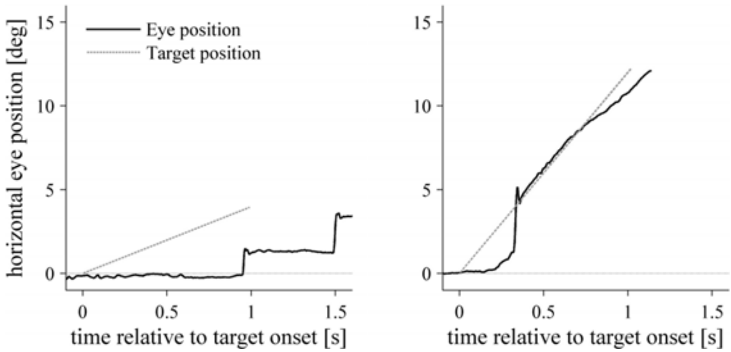

Figure1. Horizontal eye position relative to target onset from an example trial from session1. The solid black line represents the eye position, the dotted gray line the target position. (A): Ocular response to an auditory target moving at 4.7deg/s. Target position is the corresponding coordinate on the screen. (B): Horizontal eye position after onset of a Gabor patch moving with 12deg/s and a Michelson contrast of 0.01

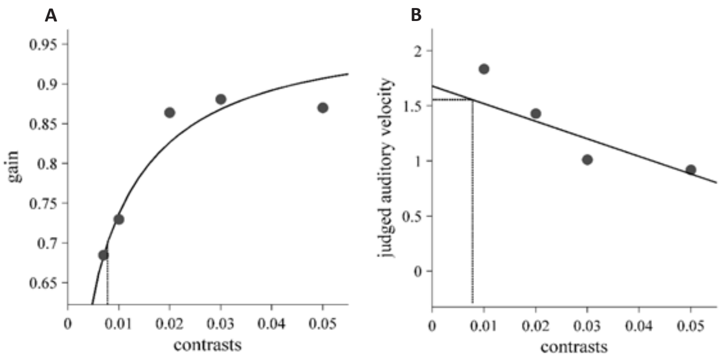

Figure2. Set contrast for session2 (A) and auditory velocity in log scale judged as equally fast for the contrast given in (A). Both data are from subject04. (A): Gain as a function of target Gabor contrast. Data points are arithmetic means for each contrast condition. The solid line represents the fitted Weibull function and the dashed line indicates the contrast which corresponds to a gain of 0.7. (B): Judged auditory velocity as a function of contrast. Data points are PSE values in log coordinates, the solid line represents the weighted linear regression, dashed lines indicate both parameter chosen for session2. One data point is out of the plotted range.

integration might be restricted in time, we performed a two-way repeated measures ANOVA on the PSE values for the within-subject factors condition and time. Only information from 100 to 700ms after pursuit onset was considered, resulting in 6 levels. The results reveal a main effect of time, $F(5,35) = 9.99$, $p = .005$, as well as a significant interaction between the two, $F(15,105) = 2.07$, $p = .017$. The main effect of condition did not yield a significant effect. To pursue the nature of the interaction, we performed post-

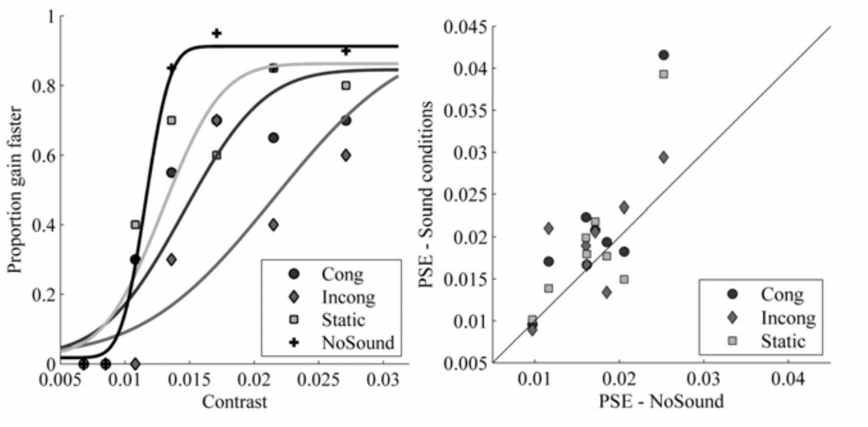

Figure3. Oculometric results for a single (A) and across all participants (B). (A): Data points represent the proportion of eye movement traces faster than the median over different contrast levels for the four different conditions (dark gray: Congruent, medium gray: Incongruent, light gray: Static, black: NoSound). Solid lines represent fitted oculometric functions. Some data points are occluded by data from other conditions. (B): PSEs from the different sound conditions relative to the NoSound condition.

hoc t-tests by comparing the PSEs of the three sound conditions against the NoSound condition in every time bin. We applied a Bonferroni-corrected alpha-level ($\alpha = 0.05/18 = 0.0028$). None of the post-hoc tests revealed a significant result. Still, at pursuit onset, participants tend to be slightly worse in the NoSound condition, whereas they tend to be better in the NoSound condition during the later periods of pursuit (Fig 4).

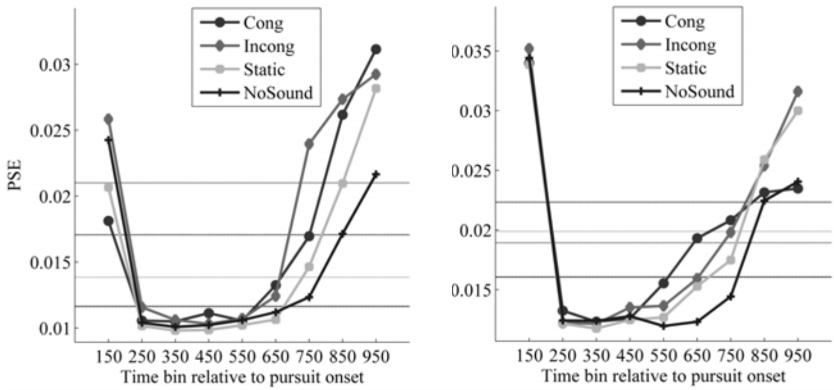

Figure4. PSEs across different time bins [ms] and conditions relative to pursuit onset for two different observers. Time bins have a width of 100ms and are referred by their central position. Data points are PSEs in the respective time bin, dashed lines represent the PSE computed from the complete trace.

2.3. Discussion experiment 1

We investigated whether auditory information is integrated to form an audio-visual smooth pursuit command. We combined an unreliable visual signal with auditory motion signals. Therefore, we first identified for each subject the target contrast which elicited a pursuit gain of 0.7 and the auditory velocity which was perceived as equally fast. Participants were then presented audio-visual signals that had different directional coherences, i.e. compared to the visual target, the auditory motion either moved in the same direction, in the opposite direction, was stationary or was not displayed. We hypothesized to find lower PSE values when visual target and auditory motion move coherently.

We measured the judged auditory velocity for different visual contrasts and different visual velocities. We observed the tendency that perceived velocities increase with increasing stimulus velocity but decrease with increasing contrast. This is contrary to the findings by Thomspon (1982) that the perceived velocity increases with contrast. We cannot rule out the possibility that there is no or a minor effect of the visual contrast on perceived velocity for our stimulus range and that the negative slopes are a result of chance. Another explanation might be that the relationship between perceived velocity and contrast might be fundamentally different during pursuit. The perceived velocity (Thompson, 1982) as well as pursuit gain (Spering et al., 2005) increase with increasing contrast. However, neither Spering and colleagues nor any other study did investigate the contrast-dependent perceived velocity during pursuit. It requires future research to solve this issue.

Overall, we did not find any beneficial effect of an auditory signal on smooth pursuit. This could be due to different reasons: either (I) auditory signals are not integrated for smooth pursuit, (II) our setup is not suitable to test audio-visual integration, (III) participants first have to learn the integration or (IV) any beneficial effect is counteracted by other factors (e.g. the distraction by a sound). The first two points will be discussed in the general discussion (see section 4). The assumption that participants first have to learn the integration seems justified as the audio-visual stimulus is rather artificial. In this case, participants would have to be repeatedly exposed to this kind of audio-visual stimulus. Furthermore, the mere presence of a sound can have an influence on smooth pursuit. Although our experimental conditions were not statistically different from one another, PSE values were mostly lower in the NoSound condition. Moreover, the auditory stimulus was reported to be perceived as unpleasant by two subjects. Thus, the negative effect of a present sound on pursuit could cast an effect that counteracts any audio-visual benefit and is comparatively larger.

To exclude these two possibilities, we performed experiment 2, in which participants were consistently exposed to an audio-visual target which moved in the same direction for three consecutive session taking place on different days. Instead of varying the presence or the directional coherence with the visual target, we varied the velocity of each component.

http://www.springer.com/978-3-658-08310-6

Audio-Visual Integration in Smooth Pursuit Eye
Movements
Wolf, C.
2015, X, 64 p. 11 illus., Softcover
ISBN: 978-3-658-08310-6