Audiovisual integration in depth

Bachelor thesis CKI

11-8-2015

Brechje Bosman, 3634639

Student Liberal Arts and Sciences

Major: Cognitive Neurobiological Psychology

Supervisor: Nathan van der Stoep

2nd supervisor: Maarten van der Smagt

Abstract

There are several reasons to believe that audiovisual integration might be affected by distance. Namely, the retinal image and the perceived intensity will decrease and arrival time of sound will be delayed. In addition, previous research has already pointed out that there exist different regions of 3D-space with respect to the processing of stimuli. We wondered if and in what manner distance affects audiovisual integration. In the present study audiovisual integration was investigated with a detection task at six different distances from 80 cm up to 15 meters. Response times, multisensory response enhancement, and race model violation were analyzed. It seemed that distance did not affect the amount of audiovisual integration, as there were no significant differences in race model inequality violation between the six distances. Our results did indicate a slight change in multisensory integration from peripersonal to extrapersonal space, by distinctive effects around the border of peripersonal space. In addition the far space conditions showed slightly different results then the near space conditions. Given that RTs for the unimodal stimuli were unequal at 80 cm distance and became more equal as distance increased, integration might be enlarged for the distances further away. This effect could have covered a decrease of integration over distance.

Keywords: multisensory integration, stimulus effectiveness, multisensory response enhancement, race model, peripersonal, extrapersonal, space

Introduction

Human beings use different senses to obtain information about the world around them. Each sense is specialized in collecting information in a different way. For example, we obtain visual information with our eyes, collect sound waves with our ears and register odorants in the air with our nose. Sometimes an event stimulates more than one sense simultaneously, in which case the signals are processed rather fast. This speed of processing is substantially faster than that of any similar event that can only be perceived by one of our senses (Stein, Huneycutt & Meredith, 1988; Calvert & Thesen, 2004; Gondan et al, 2005). Strikingly, the processing of these multimodal stimuli is even faster then what could be predicted based on statistical facilitation (Miller, 1982). The information from the different senses is integrated and processed as one event. This phenomenon is called multisensory integration (MSI).

MSI has been widely investigated, both on a cellular level in animal studies as on a functional level in behavioral studies. In these studies it was observed that there exist specific neurons which respond on or are influenced by stimuli from more than one modality. It was also observed that there are certain characteristics of multimodal stimuli that support the effect of integration in these neurons (Stein & Stanford, 2008). These characteristics are known as the principles of MSI (Holmes & Spence, 2005). Some of these principles were replicated by behavioral studies, whilst other principles showed rather conflicting results in humans (Calvert & Thesen, 2004; Holmes & Spence 2005).

One of the principles of multisensory integration that has been identified is the principle of inverse effectiveness (Holmes & Spence, 2005; Holmes, 2007). This principle says that when at least one of the unimodal stimuli is relatively weak, the response of multisensory neurons enhances relative to strong unimodal stimuli (Meredith & Stein, 1986). For behavioral studies this means that when stimuli have low intensities, the difference between the response time of the unimodal and multimodal stimuli is larger compared to the difference between uni- and multimodal stimuli with a high intensities (Senkowski et al, 2011). However some studies have found results that are inconsistent with this principle. That is, larger multisensory enhancement with high intensities (Leone & McCourt, 2013).

Another important aspect of multisensory integration is that it occurs especially when our brain interprets the stimuli as coming from the same event. Hence, the effect of multisensory integration is generally stronger when the content of the stimuli are substantively congruent. This means that the information you will get from both of the stimuli should be the same (Calvert & Thesen, 2004). When the stimuli are incongruent, your brain receives conflicting information at the same time. In the case of conflicting information from different senses there could even be an illusory percept. A famous example of this illusory percept in audiovisual integration is the Mcgruk effect in which people hear ‘dadada’, when they watch a video of a face making the lip movements of ‘gagaga’, while simultaneously hearing a voice saying ‘bababa’ (McGurk & MacDonald, 1976).

Besides content, the location and timing of the stimuli also plays an important role in MSI. When stimuli come from the same location and have the same onset, it is reasonable to assume that the stimuli originate from the same event given that this correlation exists in real life. Consistent with this reasoning, the effect of MSI enhances when stimuli are presented in the same location and at the same time. These principles of integration are referred to as the spatial rule and the temporal rule (Holmes & Spence, 2005). However the spatial rule could not be found in all the experiments conducted. It seems that the location of the stimulus must be relevant for the task to find the effect of the spatial rule (see Spence, 2013 for a review).

Most of the experiments conducted were mainly about the placement of the stimuli in a two-dimensional plane (elevation and laterality, e.g., Meredith & Stein, 1996; Li, Yang & Wu, 2012; Spence & Driver, 1997; Spence, 2013). However, there have been some experiments in which the distance between the stimuli and observer were taken into account for audio and visuo-tactile integration (Canzoneri et al., 2012; Farnè & Làdavas, 2002; Graziano, Hu & Gross, 1995). In these experiments it was observed that the processing of stimuli depends on the spatial location and that the boundaries of peripersonal space are located around a spatial continuum between near and far space. Yet, for audiovisual integration, the placement of stimuli in depth has not been considered in most of the previous research of MSI (see Van der Stoep, Nijboer, Van der Stigchel, & Spence, 2014 for a review of multisensory integration in 3-D space).

Recently, an increase in audiovisual integration in far as compared to near space was observed (Van der Stoep, Van der Stigchel, Nijboer, & Van der Smagt, 2015). Two explanations were given for this finding. On the one hand audiovisual integration might be of greater use in far space. Both seeing and hearing are dominant in far space, whereas in near space other senses are generally used, such as touch and taste (see Previc, 1998 for a review). That is why audiovisual integration might be especially helpful in the processing of information from far space whilst for near space other senses may dominate. Besides, the further the stimuli are presented the weaker their intensity. According to the rule of inverse effectiveness, the effect of audiovisual integration has to enhance when stimuli are presented further away and thus have weaker intensities. The results of Van der Stoep et al. (2015) led to the conclusion that the increase of integration in far space was a result of a combination of both changes in the region of space and the intensity of the stimuli.

In the experiment of Van der Stoep et al. (2015) only two distances were compared; respectively 80 cm and 200 cm. This means that instead of enhancing gradually over distance, the results could also be due to a distinction between 80 cm (near space) and 200 cm (far space). This distinction was already observed in audio- and visuotactile integration by identifying the area of peripersonal space (Canzoneri et al., 2012; Farnè & Làdavas, 2002; Graziano, Hu & Gross, 1995). In addition, a distinction has been observed between far and near space in the visual domain. That is, some patients who suffered from visuospatial neglect in far space could still locate objects perfectly well in near space and the other way around (Aimola et al, 2012; van der Stoep et al., 2013; Halligan

& Marshall, 1991; Vuilleumier et al, 1998). This indicates that visual information from far space is processed separately from visual information from near space.

However, in real life we obtain auditory and visual information from all kind of distances and much of this information comes from far space. An interesting side effect of information that originates from far space is that due to the difference in speed of sound and speed of light, the arrival times of the information will differ from each other. At a distance of 15 meter this means that auditory stimuli reach the brain about 43 ms later than visual stimuli. Stimuli that are presented at 15 m distance will however, still be perceived as originating from the same event (Alais & Carlile, 2005) because the stimuli fall within the temporal binding window (Colonius & Diederich, 2010; Stein & Stanford, 2008). Nevertheless, this time difference might still have an influence on the integration effect.

In general, studying audiovisual integration in different depth planes has a lot of interesting aspects, which are not yet well understood: changes in stimulus effectiveness, temporal alignment, and regions of space. In this study it was investigated how audiovisual integration changes across six different distances from 80 cm up to 15 meter. On the one hand one could expect that the effect of MSI will keep enhancing as the distance increases because intensity will gradually decrease further and the benefit of audiovisual integration is steadily getting larger. On the other hand there could be a distinction of MSI between near and far space with a turning point at which MSI will not further enhance. This would be in line with the distinction between near and far space found in the visual domain and in audio- and visuotactile integration.

Methods

Participants

Twenty-eight participants were tested in all conditions of this study. The data of five participants was removed from further analysis based on their accuracy or RTs (see *Data analysis*). Hence, the data of twenty-three people was analyzed (thirteen were female). Mean age was 22.91 years (SD = 2.09). Everyone was right-handed. All the participants had proper normal or corrected to normal vision. In addition none of the participants had relevant hearing problems for this task. Before starting with the experiment everyone received instructions on the task and signed an informed consent. If something was not clear participants were free to ask for further explanation of the task. The participants received a monetary reward or study credits for taking part in this study.

Setup

In order to investigate the effect of audiovisual integration at different distances, a redundant target effect task was used (Miller, 1982). The task included three different target stimuli: an auditory (A), visual (V), and audiovisual (AV) stimulus. Furthermore the task was divided in six distance conditions, which entailed a different distance between the origin of the stimuli and the participant. During the task participants had to press the spacebar as soon as they detected any of the target stimuli mentioned above. Their response time for every stimulus was recorded to investigate the difference between the response time (RT) of unimodal and multimodal stimuli and to compare the RTs at various distances with each other.

The experiment took place in an area of 15 meters long and 1.40 meters width. The demarcation of this space on the right and on the backside originated from the walls of the room. On the left side a temporary wall of 1.40 meters high was created. At one end of the area an Iiyama vision master pro 454 monitor (resolution: 1024 x 768, screen refresh rate: 100 Hz) and Harman/Kardon HK206 speakers (frequency response: 90-20000 Hz) were placed. The monitor was placed in the center of the corridor, at a distance of 47.5 cm from each side. The speakers were placed next to the screen on both sides, so that the sound appeared to originate from the monitor. The speakers were positioned at a 37.5 cm distance from the wall. The chair was also placed exactly in the center of the hallway, but varied in distance from the screen for the different conditions.

In every condition the task was conducted at a different location, by moving the chair further or closer to the screen. The distances used for the conditions were 80 cm, 140 cm, 200 cm, 500 cm, 1000 cm and 1500 cm. This distance was measured from the monitor to the middle of the chair, as this was where the eyes of the participants were. See Figure 1 for an overview of the setup.

Figure 1. A schematic bird’s-eye view of the setup of the experiment. A chair was placed at each of the distances shown depending on the specific condition.

Experiment

The six conditions of the experiment all consisted of the same task. Every participant had to perform this task in all six conditions. To avoid any effect of the sequence of conditions, the order of the distances at which the task had to be performed was counterbalanced. Hence, the first distance on which the task was performed differed between participants. Before a condition started the participant was seated at the correct distance from the screen and loudspeakers. Consequently the participant had a small break between conditions while moving the chair. Each condition started with 12 practice trials so that the participant got familiar with the task and stimuli. After the practice trials of the first condition people were asked if they fully understood the task. Once the first condition (i.e., first distance) was completed, results were checked on accuracy and response time. When the results of the participant showed that the participant understood the task, the rest of the conditions were conducted. If not, the participant got feedback on the results and had to perform the task again. If no improvement was observed, the participant was excluded from the experiment.

A single condition contained 150 trials consisting of 30 catch trials (20%) and 120 target present trials: 40 A, 40 V, and 40 AV stimuli. The order of the trials was randomized so that the participant had to pay attention to detect a stimulus. The participant could decide when he wanted to start with the experiment after the practice trials by pressing the spacebar. Every condition was divided in three parts by pauses, one after 50 and one after 100 trials. The participant could continue with the experiment after a pause by pressing the spacebar whenever he was ready.

A trial started with a black fixation cross of 3.2 cm presented on a gray background (19.94 cd/m^2). Participants were instructed to focus on the fixation cross throughout the experiment. The fixation cross was displayed for a random time between 700 and 1300 ms. Then the cross disappeared and a blank screen was presented for a random period between 200-250 ms. After this blank screen a stimulus was shown for 100 ms. In case of a catch trial, no target appeared. When the participant detected a stimulus he had to press the spacebar. The time between the onset of the stimulus and the press was recorded. The onset of the stimulus was followed by an inter stimulus interval of 1500 ms, whereupon the next trial begun. The background color remained equal during the entire experiment.

The stimuli used in the experiment were either, visual, auditory or audiovisual. The visual stimulus consisted of a white circle of 3.2 cm diameter (102cd/ m^2). The auditory stimulus consisted of a 100 ms white noise burst. The audiovisual stimulus consisted of the simultaneous presentation of the visual and auditory stimuli that were used in the unimodal trials. The various distances used in the conditions caused the loudness of the auditory stimuli and the retinal image size of the visual stimuli to vary across the distance conditions. These distance-dependent changes occur in real life as well, and are relevant as shown by the findings of van der Stoep et al. (2015) that both distance as inverse effectiveness play a role in the difference in MSI between far and near space. In Table 1 the specific values of the perceived intensity for each condition are reported.

Table 1. Stimuli characteristics

Distance	Sound intensity of auditory target (db)	Light intensity of visual target on background. (cd/m²)	Retinal image size(mm)	Visual angle (º)
80	63.77	109.5	0.68	2.29
140	59.43	109	0.39	1.32
200	58.17	108	0.27	0.92
500	52.73	33.3	0.11	0.37
1000	46.50	22.7	0.05	0.18
1500	41.67	21	0.04	0.12

Data analysis

Go trials with RTs between 100 ms and 1000 ms and no-go trials without a response were considered correct. Participants were not included in the analysis when they did not take part in every condition. Participants were also excluded when they showed median RTs > 450 ms or/and an accuracy lower than 80% for at least one of the modalities at any of the distances. These participants were regarded as unable to perform the task as instructed. Based on these requirements, seven out of thirty people were excluded from the analysis. For the other participants RTs < 100 ms and > 1000 ms were removed. This led to the removal of 0.04% of the data in Go conditions in the 80 cm condition, 0.07% in the 140 cm condition, 0.04% in the 200 cm condition, 0.11% in the 500 cm condition, 0.07% in the 1000 cm condition and 0.07% in the 1500 cm condition. Note that the percentages are calculated from the 23 participants who were included in the analysis.

For the response time analysis the median RTs were used since RT distributions are generally skewed and the median is less affected by outliers than the sample mean. For each participant the median RT's per modality per distance were obtained. A 6 x 3 repeated measure ANOVA with the factors Distance (80 cm, 140 cm, 200 cm, 500 cm, 1000 cm and 1500 cm) and Modality (A, V and AV) was used to analyze differences between conditions. Whenever the assumption of sphericity was violated the Greenhouse Geisser correction was used. To investigate which conditions differ from each other two-tailed paired samples t-tests with a Bonferroni correction were used.

To analyze the amount of quicker detection of AV stimuli compared to the fastest unimodal stimuli, the multisensory response enhancement (MRE) was calculated. Both the absolute MRE (aMRE) and the relative MRE (rMRE) were calculated for each participant in each distance:

$$\mathrm{aMRE} = \min(\mathrm{RT_A}, \mathrm{RT_V}) - \mathrm{RT_{AV}}$$

$$\mathrm{rMRE} = \frac{\min(\mathrm{RT_A}, \mathrm{RT_V}) - \mathrm{RT_{AV}}}{\min(\mathrm{RT_A}, \mathrm{RT_V})} \times 100\%$$

The absolute MRE gives information about the amount of enhancement in milliseconds. The relative MRE was also calculated to control for differences in unimodal baselines between distances. Note that for calculating the MRE's, median RT's were used. Both aMRE and rMRE were analyzed with a repeated measure ANOVA with the factor distance (80 cm, 140 cm, 200 cm, 500 cm, 1000 cm and 1500 cm) to investigate if there was any difference in enhancement between conditions (distances).

To test if the MRE was a result of statistical facilitation alone, or that there must have been multisensory integration to require these amounts of enhancement, the cumulative distribution functions (CDF) of the RT's to A, V and AV stimuli were calculated. The CDF of the audiovisual stimuli was compared with the upper bound of statistical facilitation predicted by the race model (Raab 1962; Miller 1982, 1986; Ulrich et al. 2007):

$$P(RT_{AV} < t) \leq P(RT_A < t) + P(RT_V < t)$$

This formula provides the probabilities to get certain RT's on audiovisual stimuli based on chance alone. The curve obtained from this formula is very steel because it is corrected for a maximal possible negative correlation between A and V. This would be the case when A is processed faster if V is processed at the same time, without any integration occurring. When the probability on a certain RT in the AV condition is higher than the probability calculated with the race model, a reasonable conclusion is that multisensory integration was present. In this case we can say that the race model has been violated.

The average and median amount of race model violation in ms was compared between conditions with a repeated measure ANOVA with the factor distance (80 cm, 140 cm, 200 cm, 500 cm, 1000 cm and 1500 cm) In addition the difference between the CDF of AV and the race model curve was calculated at nine percentiles of the curve. To test if there was a significant violation, one sample t-tests were used. Because there had to be nine tests for every distance the *p*-values were corrected for nine tests using the Bonferroni method. The range of percentiles that showed significant violation was compared between the various distances.

Results

Accuracy

The 23 participants were very accurate performing the task. Overall accuracy was 0.991 (SD = 0.024). The accuracy for Go trials of 0.997 (SD = 0.009) indicates that participants were very capable of detecting the stimuli. In addition participants were also able to suppress their response in the Catch trials showed by an accuracy of 0.974 (SD = 0.041) for the no-go trials.

Response times

The RT's showed a significant main effect for Modality [$F(1.453, 31.966) = 89.860$, $p < .001$, $\eta p^2 = .803$]. RTs for AV stimuli ($M = 264$ ms, SD = 5 ms) were significant shorter than the RT's for A stimuli ($M = 296$ ms, SD = 8 ms) [$t(22) = 8.600$, $p < .001$, d = 4.789] and the RT's for V stimuli ($M = 318$ ms, SD = 7 ms) [$t(22) = 17.592$, $p < .001$, d = 8.67]. The difference between RT's on A stimuli and V stimuli was also significant [$t(22) = 4.357$, $p = .001$, d = 3.04].

For the distance from the stimuli there was also a significant main effect observed [$F(3.336, 73.386) = 12.763$, $p < .001$, $\eta p^2 = .367$]. The RT's for distances which were quite close, respectively 80 cm ($M = 276$ ms, SD = 8 ms), 140 cm ($M = 279$ ms, SD = 8 ms) and 200 cm ($M = 278$ ms, SD= 8 ms) were significantly different from the conditions in the real far space, respectively 1000 cm (M= 307 ms, SD= 7 ms) and 1500 cm ($M = 319$ ms, SD= 6 ms) [All *t*-values > 3.924 and *p*-values < .011].

There was also an interaction effect observed [F(5.705, 125.508) = 6.819, p < .001, ηp^2 = .237]. In figure 2 the average median RT's are displayed. In the graph it is shown that RT's for A increase more than RT's for V over distance. There is indeed an interaction effect between A and V when tested separately with a 2 x 6 repeated measures ANOVA. [F(5, 110) = 9.968, p < .001, ηp^2 = .312], but when corrected for the travel time of sound no interaction is observed [F(5, 110) = .549, p = .739]. This indicates that the difference in the increase in RT between A and V can be explained by the delayed arrival of sound compared to light. After correcting for speed the difference between A and V remains significant with higher RT's for V than for A.

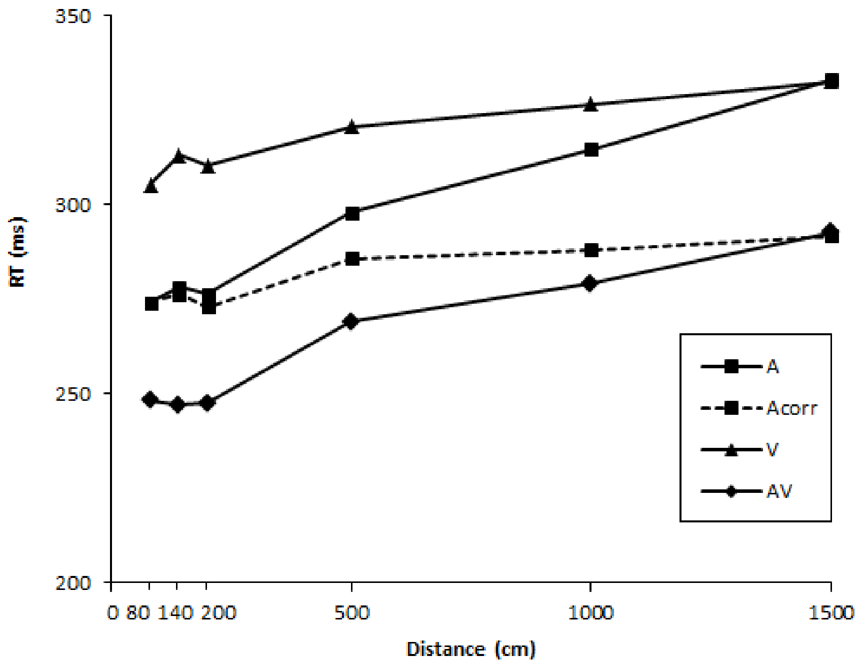

Figure 2. Average median response times for A, V and AV in all six distances.

Multisensory response enhancement

All absolute and relative MRE's were tested significant with one sample t-tests [all t-values > 6.1, all p-values < .001], indicating that responses on AV stimuli were faster than responses to the unimodal stimuli at every distance. In Figure 3 the average absolute MRE's for all distances are shown in the left panel and average relative MRE's are shown in the right panel.

There was no main effect of Distance for the absolute MRE [F(5,110) = .842, p = .523]. To control for the dissimilarities between the baselines of the various conditions, the relative MRE was also tested for differences between conditions. However also in relative terms there were no significant differences observed [F(5,110) = .760, p = .581].

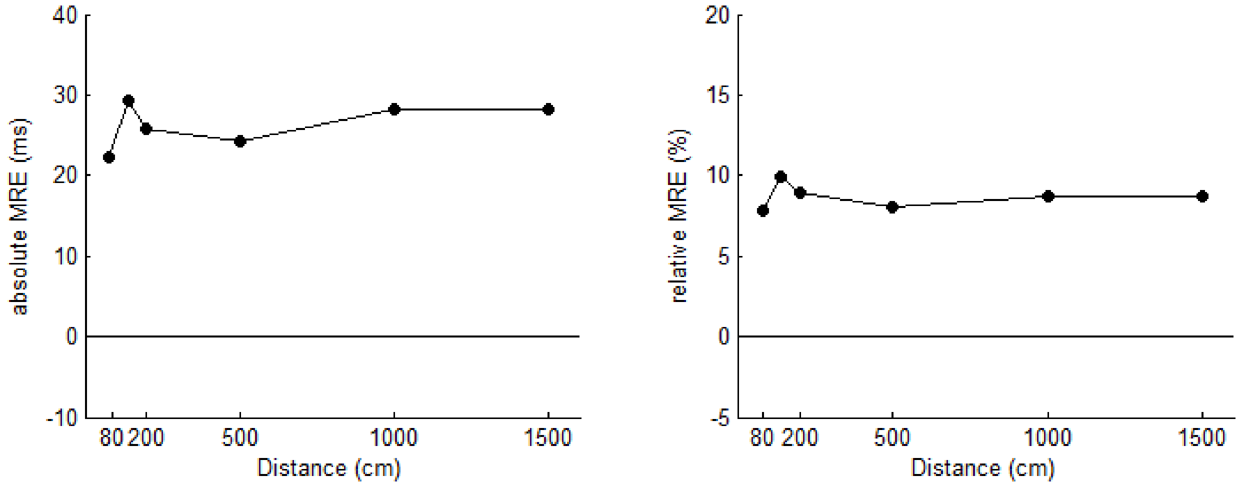

Figure 3. Left side: Average absolute MRE's for the six distances. Right side: Average relative MRE for the six distances.

Race model inequality violation

The CDF of the RT on AV stimuli was compared with the upper bound of statistical facilitation predicted by the race model. If the CDF of AV was higher than the race model prediction the model was violated. The average Race model violation for every distance is shown in the left panel of Figure 4. When the average RMVs were compared no significant difference was observed [$F(5,110) = 1.258$, $p = .287$]. In addition we compared the median RMVs for dissimilarities because they are less affected by the negative violation at the high percentiles, but again there were no significant results observed [$F(5,110) = 1.254$, $p = .289$]. The median RMV is shown in the right panel of Figure 4.

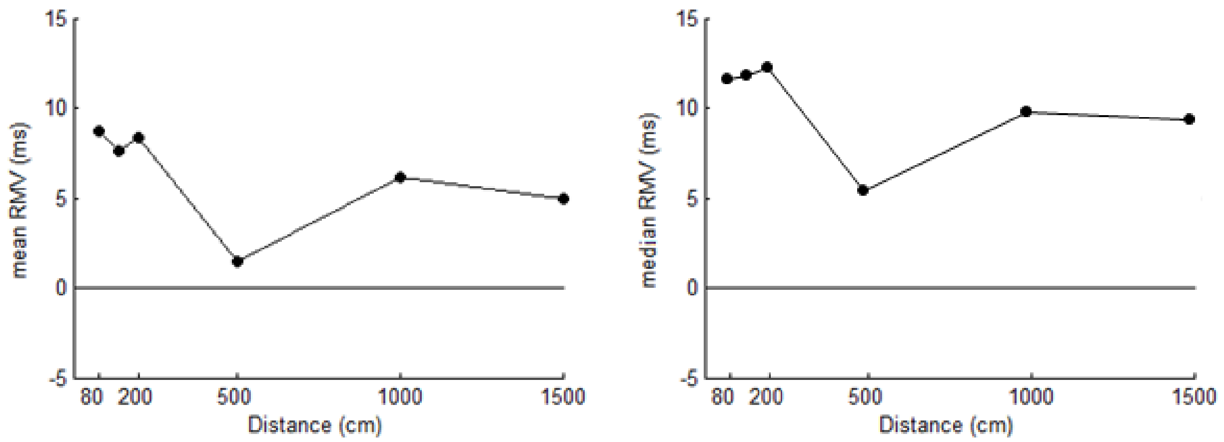

Figure 4. Left panel: The average race model violation in ms for every distance. Right panel: The median race model violation in ms for every distance.

Figure 5 shows the amount of race model violation for nine percentiles for every distance. For every distance it was tested at what range of percentiles the race model violation was significant. In the 80 cm condition a significant effect from the 2nd till the 6th percentile was observed significant [all t-values > 4.1, all p-values < .003]. The same range was observed for the 140 cm condition [all t-values > 3.2, all p-values < .034]. In the 200 cm condition the range was even wider, from the 2nd till the 7th percentile [all t-values > 4.2, all p-values < .003]. The 500 cm condition showed a different pattern, only the third percentile was significant [$t = 3.9$, $p = .005$]. At a distance of 1000 cm, again a range from the 2nd till the 6th percentile was observed [all t-values > 4.0, all p-values < .004]. For the condition of 1500 cm the range from the 1st till the 6th percentile was significant [all t-values > 3.3, all p-values < 0.02].

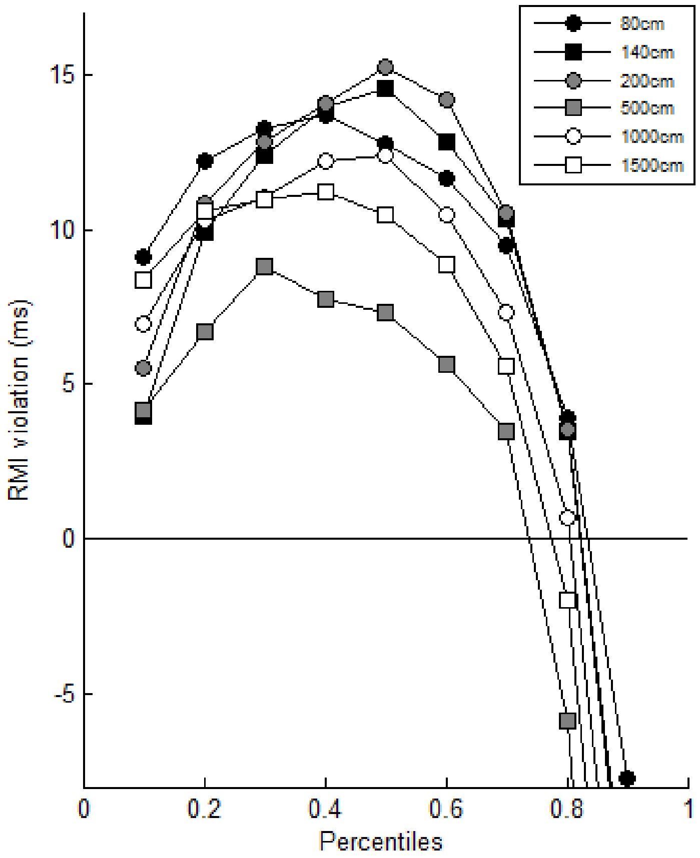

Figure 5. Amount of race model violation in ms for each distance.

Discussion

The intent of this experiment was to gain more insight into how audiovisual integration changes as the distance between stimuli and the observer increases. Participants had to respond as fast as possible to auditory, visual and audiovisual stimuli that were presented at different distances. The data revealed no major differences in the amount of integration between the distances but did show some small but interesting variations and patterns.

First of all, we observed a slight increase of RTs over distance for all modalities (A, V and AV). This overall increase of RTs could be a result of the decrease of perceived intensity due to the increase of distance (Piéron, 1914; Kohfeld, 1971; Niemi, 1979). The RTs of AV did not show a different pattern than the pattern of the fastest of the unimodal condition (A), indicating that distance did not have a large effect on the amount of enhancement. The increase of the RTs for the auditory stimuli was steeper than for visual stimuli. However, when we corrected the RTs for the auditory stimuli for the delayed arrival time of sound, a similar pattern for A as for V emerged. The difference between the patterns of A and V was entirely due to the difference between speed of light and speed of sound.

Secondly, multisensory response enhancement was observed at all distances both in absolute and relative terms. There were no large differences in the amount of MRE between the various distances. This means that participants responded faster on AV stimuli than unimodal stimuli, regardless of the distance at which the stimuli were presented. Nonetheless, it is striking that at a distance of 140 cm the enhancement was larger than at 80 cm and 200 cm. The additional analysis of the multisensory integration by comparing our data to the race model revealed a similar pattern. The race model inequality was violated for a wide range of percentiles at almost all distances and in both average and median RMVs there were no large differences between conditions. However, the 140 cm condition stands out again when looking at the average race model violation, with a lower race model violation than for both the 80 cm and 200 cm condition.

Even though the 140 cm condition was not significantly different form the others, it is still interesting to see that it deviates from the distances close to it. One could expect that the results of the 140 cm condition (which is exactly between 80 cm and 200 cm) lie between the results of the 80 cm and 200 cm conditions, because of a possible gradual enhancement or reduction. This is not the case, so something seems to happen in this area of space. It might have to do with the transition between peripersonal and extrapersonal space that probably lies somewhere around this area (see Previc, 1998 and Van der Stoep, Nijboer, Van der Stigchel, & Spence, 2015 for a review). Some researchers suggested that the shift between peri- and extrapersonal space is related to arm length, which causes individual differences in the size of the peripersonal space (Halligan, Fink, Marshall, & Vallar, 2003; Longo & Lourenco, 2007). In addition to variances between persons, the transition can also differentiate within a person. That is, the transition is not located at a fixed absolute distance, as there have been found ways to extend (Longo & Lourenco, 2006) or reduce (Longo & Lorenzo 2009) the area of peripersonal space. Even the available information can affect the area of peripersonal space, as different brain areas use different sensory information to estimate the peripersonal space (Makin, Holmes & Zohary, 2007).

The amount of audiovisual integration observed here at 140 cm might be affected by being near this transition of peripersonal to extrapersonal space. In previous research it was shown that people have difficulties estimating the correct distance of visual stimuli that were presented around the area of the border (Gabbard, Cordova & Ammar, 2007). These difficulties might be a result of different processing of stimuli that are located around the transition area. Investigation of a combination of the location of the transition and audiovisual integration in a wider spectrum of distances between 80 cm and 200 cm, will be useful to obtain more information about the role of the transition in integration.

The 140 cm condition is not the only condition that is noteworthy. In both the average RMV, the median RMV, and the width of percentiles showing significant violation, the 500 cm condition also stands out. At this distance the race model violation is less than at any of the other distances. Something must have influenced multisensory integration at just this particular distance. As far as we know, there are no findings from previous research that could explain this observation. Therefore we checked the test room for any oddities that could have influenced the integration. For example, there could be features in the room that caused different reflections of sound at 500 cm than at the other distances. We did not observe anything in the test room that could explain the results. However, the remarkable results of the 500 cm condition could be due to insufficient counterbalancing which was a consequence of the elimination of the data from some participants. The 500 cm condition has been conducted as the first condition and as the last condition one time less than the other distances. Both the first and the last condition are sensitive to acquire less adequate data, due to respectively experience and fatigue. This may have caused the race model violation to be lower at 500 cm distance.

Another salient observation in our data was the difference between close distances (80 cm, 140 cm and 200 cm) and far distances (500 cm, 1000 cm and 1500 cm). In both RTs, MRE, and RMV we observed a different pattern for the first three distances compared to the last three. For RTs we observed roughly the same RTs for distances in near space whilst in far space they were slowly rising as distance increased. The MREs showed a parabolic shape in near space but stayed more or less the same in far space. When analysing RMV we observed higher RMV for the three distances in near space than for the three distances in far space. Taken together, these patterns could indicate that stimuli that are closer to the observer are processed in a different way than stimuli presented in far space. Specifically, it is assumed that multisensory integration is more thorough for information from peripersonal space because we have more senses that are capable of processing information from near space (touch, taste etc.) than senses that are capable of processing information from far space (vision and hearing) (Holmes & Spence, 2004). Yet, auditory and visual information can be obtained from far as well as near space. However the processing of audiovisual stimuli could be different for near space compared to far space. Namely, in far space no additional senses could provide information about the event, whilst in near space there could be more information from one of our other senses. The opportunity of more available information might change the way our brain treats stimuli from near space.

Furthermore, it has been suggested that multisensory integration is larger when the unimodal RTs are the same (i.e., they are physiologically simultaneous; Hershenson, 1962; Nickerson, 1973; Leone & McCourt, 2013). Equal RTs suggest that the auditory and visual stimuli must have arrived more or less simultaneously at the cortex, despite the natural differences in cortical latency of sound and light. In other words, larger differences in unimodal RTs must provide less integration than equal unimodal RTs. However, our data appears not to be consistent with this principle. Whereas the unimodal RTs are almost equal at the distance of 15 m and showed a difference of > 25 ms at 80 cm, the absolute MRE is just slightly larger for 15 m and almost the same in terms of relative MRE. When looking at the race model violation we even observed more violation at the distance of 80 cm than for the 15 m condition.

Remarkably, in a recent study of audiovisual integration at 1 and 15 m, no enhancement in the 15 m condition was observed at all (unpublished data) indicating that integration would decrease over distance. It could be that both these principles are true. That is, audiovisual integration is stronger with equal RTs on the unimodal stimuli but at the same time decreases when the distance increases. If this is the case the two principals were working against each other because the difference between unimodal RTs was larger in near space compared to far space. For example, at 15 m the integration will be enhanced by the higher equality of the unimodal RTs but at the same time reduced by the larger distance, resulting in no change in the amount of integration. It would be interesting to analyze the data from a similar experiment where the unimodal RTs will be more or less equal at the 80 cm distance and presumably, due to the speed of sound, have a bigger unimodal RT-difference in the far space conditions. If our hypothesis is correct, there will be no, or at least less, integration in the far space conditions compared to the near space conditions.

In sum, although RTs decreased as the distance from the observer increased, there were no large effects of distance on multisensory integration observed. Nevertheless there were some interesting patterns and small deviations present in our data. First of all, the amount of multisensory integration was different for the 140 cm distance, which could be a consequence of a transition between peripersonal and extrapersonal space. We also observed a distinction between the distances from 80 cm till 200 cm and 500 cm up to 1000 cm, which could indicate that stimuli from near space are processed in a different manner than stimuli from far space. Finally, we point out that the difference in unimodal RTs was not the same in all conditions. We reasoned that an effect in which audiovisual integration would decrease when distance increases, could be invisible due to the effect of physiological simultaneity at large distances.

Relevance of current study for Artificial Intelligence

Artificial intelligence studies the human brain in relation to computers. It uses knowledge about how the brain operates, to create computers that are ‘intelligent’ and tries to figure out how the brain works with the help of computer models. The current study provides information on how fast people process audiovisual stimuli when presented at different distances. This information can be used to develop computers that are more capable of localizing stimuli than they are now. If a computer uses a similar system of processing stimuli as the human brain, you can use response times to calculate how far the stimulus was located.

Besides, knowing how the brain accomplishes faster responses to multisensory stimuli than to unimodal stimuli, and possible differences of this process between distances, can help to create computers that are faster for multimodal stimuli at any distance. Nowadays, a computer slows down when processing more information at the same time, whereas humans, as observed in this study, are faster when processing information from more than one modality at the same time. If we could implement this idea in a computer, computers could reach higher speed rates for complex information than any speed rates they are currently achieving. This can be used for example in security devices that register both visual information and sound.

In addition to creating better and faster computers, there are some other practical applications as well. Features of information in everyday life that require attention can be adjusted to the knowledge that was obtained in this study. For example, traffic signs might be safer when they have an auditory component as well because we observed that there was still multisensory enhancement in far space. Not only could the knowledge obtained from the far space conditions, be of use. Having more knowledge about the ideal combination of distance and stimuli features such as intensity, can be at help when thinking about at what distance an observer should be to processes multisensory information optimally. How far do you have to be from your computer screen to gain the information best and what seat in the lecture hall will be optimal for obtaining the content? These are just a few of the questions in which the knowledge of this study is necessary for solving.

Literature

Alais, D., & Carlile, S. (2005). Synchronizing to real events: Subjective audiovisual alignment scales with perceived auditory depth and speed of sound. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(6), 2244-2247.

Aimola, L., Schindler, I., Simone, A. M., Venneri, A. (2012). Near and far space neglect: task sensitivity and anatomical substrates. *Neuropsychologia*, *50*, 115–1123.

Calvert, G. A., & Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology-Paris*, *98*(1), 191-205.

Canzoneri, E., Magosso, E., & Serino, A., (2012). Dynamic sounds capture the boundaries of peripersonal space representation in humans. *PLoS ONE, 7*(9), e44306.

Colonius, H., & Diederich, A. (2010). The optimal time window of visual–auditory integration: a reaction time analysis. *Frontiers in integrative neuroscience*, *4*.

Farnè, A., & Làdavas, E. (2002). Auditory peripersonal space in humans. *Journal of Cognitive Neuroscience. 14*, 1030–1043.

Gabbard, C., Cordova, A., & Ammar, D. (2007). Estimation of reach in peripersonal and extrapersonal space: A developmental view. *Developmental Neuropsychology*, *32*(3), 749-756.

Graziano, M. S., & Gross, C. G. (1995). The representation of extrapersonal space: a possible role for bimodal, visual-tactile neurons. *The cognitive neurosciences*, 1021-1034.

Gondan, M., Niederhaus, B., Rösler, F., & Röder, B. (2005). Multisensory processing in the redundant-target effect: A behavioral and event-related potential study. *Perception & Psychophysics*, *67*(4), 713-726.

Halligan, P. W., Fink, G. R., Marshall, J. C., & Vallar, G. (2003). Spatial cognition: Evidence from visual neglect. Trends in Cognitive Sciences, 7, 125 – 133.

Halligan, P. W., & Marshall, J. C. (1991). Left neglect for near but not far space in man. *Nature*, *350*(6318), 498-500.

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of experimental psychology*, *63*(3), 289.

Holmes, N. P., & Spence, C. (2004). The body schema and multisensory representation (s) of peripersonal space. *Cognitive processing*, *5*(2), 94-105.

Holmes, N. P., & Spence, C. (2005). Multisensory integration: space, time and superadditivity. *Current Biology*, *15*(18), R762-R764.

Holmes, N. P. (2007). The law of inverse effectiveness in neurons and behaviour: multisensory integration versus normal variability. *Neuropsychologia*, *45*(14), 3340-3345.

Kohfeld, D. L. (1971). Simple reaction time as a function of stimulus intensity in decibels of light and sound. Journal of Experimental Psychology, 88, 251–257.

Leone LM, McCourt ME (2013) The roles of physical and physiological simultaneity in audiovisual multisensory facilitation. *i-Perception 4*, 213–228.

Li, Q., Yang, J., & Wu, J. (2012, July). Effects of spatial location on bisensory audiovisual integration in horizontal meridian. In *Complex Medical Engineering (CME), 2012 ICME International Conference on* (pp. 302-305). IEEE.

Longo, M. R., & Lourenco, S. F. (2007). Space perception and body morphology: Extent of near space scales with arm length. Experimental Brain Research, 177, 285 – 290.

Longo, M. R., & Lourenco, S. F. (2006). On the nature of near space: Effects of tool use and the transition to far space. Neuropsychologia, 44, 977 – 981.

Lourenco, S. F., & Longo, M. R. (2009). The plasticity of near space: Evidence for contraction. *Cognition*, *112*(3), 451-456.

Makin, T. R., Holmes, N. P., & Zohary, E. (2007). Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *The Journal of neuroscience*, *27*(4), 731-740.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746–748.

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

Meredith, M. A., & Stein, B. E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *Journal of Neurophysiology*, *75*(5), 1843-1857.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive psychology*, *14*(2), 247-279.

Nickerson, R. S. (1973). Intersensory facilitation of reaction time: energy summation or preparation enhancement?. *Psychological review*, *80*(6), 489.

Niemi, P. (1979). Stimulus intensity effects on auditory and visual reaction processes. *Acta Psychologica*, *43*(4), 299-312.

Piéron, H. (1914). Recherches sur les lois de variation des temps de latence sensorielle en fonction des intensit´es excitatrices. Anne´e Psychologique, 20, 17–96.

Previc, F. H. (1998). The neuropsychology of 3-D space. *Psychological Bulletin*, *124*(2), 123-164.

Senkowski, D., Saint-Amour, D., Höfle, M., & Foxe, J. J. (2011) Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *Neuroimage*, *56*(4), 2200-2208.

Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences*, *1296*(1), 31-49.

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & psychophysics*, *59*(1), 1-22.

Stein, B. E., Huneycutt, W. S., & Meredith, M. A. (1988). Neurons and behavior: the same rules of multisensory integration apply. *Brain Research*, *448*(2), 355-358.

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, *9*(4), 255-266.

Van der Stoep, N., et al. (2014). Multisensory interactions in the depth plane in front and rear space: A review. *Neuropsychologia*, *70*, 335-349.

Van der Stoep, N., Van der Stigchel, S., Nijboer, T. C. W., & Van der Smagt, M. J. (in press). Audiovisual integration in near and far space: Effects of changes in distance and stimulus effectiveness. *Experimental Brain Research*. DOI: 10.1007/s00221-015-4248-2

Van der Stoep, N., Visser-Meily, J. M. A., Kappelle, L. J., De Kort, P. L. M., Huisman, K. D., Eijsackers, … & Nijboer, T. C. W. (2013). Exploring Near and Far Regions of Space: Distance Specific Visuospatial Neglect After Stroke. *Journal of Clinical and Experimental Neuropsychology*, *35*(8), 799-811.

Vuilleumier, P., Valenza, N., Mayer, E., Reverdin, A., & Landis, T. (1998). Near and far visual space in unilateral neglect. *Annals of Neurology*, *43*(3), 406-410.