Multimodal Perception and Action

Nelson, J.S.

2020

Our perception of the world around us depends on input from many sense organs working together. In particular, our senses of vision, hearing, and touch are crucial for sensing the present state of our environment. Each sense is particularly suited for certain kinds of information, which we compile in order to perform simple and complex motor tasks. For example, normally sighted individuals rely in large part on the visual system to determine the locations of, and interact with objects in our environment, but we are also able to reach for and manipulate objects in the dark by relying on haptic inputs. Our sense of timing of brief events, which for example is necessary for actions such as singing, dancing, and playing musical instruments, is largely informed by auditory and haptic input, but in cases such as orchestral performances the players must synchronize with the visual signals of the conductor. Interception, a common action in sports, is commonly thought of as a visually guided task, though many animals use olfaction, echolocation, or electroreception to hunt. It is therefore not immediately clear if, or why, interception should be only possible in the visual domain for humans. These questions led us to wonder how each of our senses informs our ability to perform certain tasks, and how generalizable these tasks are to other, less common sensory modalities.

We conducted four studies to understand the influence of our senses on space and time, as well as how we combine spatial and temporal information in an interception task. The study conducted in **Chapter 2** concerns our perception and ability to act in the space within reach of our own bodies. We hypothesized that blind individuals would perform differently in a line-drawing task from blindfolded sighted individuals, due to a lack of development of visual experience. We set up a booth at the ZieZo Beurs in Utrecht and collected data from 132 participants, all of whom were attendees of the convention. These participants, whose visual capacity ranged from being fully blind to being normally sighted, were instructed to perform a series of tasks, wherein they matched the location of a pen in their hand to that of their other hand under the table, and then drew lines of a specific length in a specific direction. We were surprised to find no differences in performance, with regard to positioning of the pen, or the length or angle of the drawn lines, between our groups. We conclude that the development of visual experience has no effect on people's sense of the space around their own bodies.

Chapter 3 explores the precision with which we can determine the timing of brief events, across differing sensory modalities. It is well known that we have better temporal precision for auditory and haptic stimuli than for visual ones, but the extent to which motor actions play a role in our perception is not yet fully understood. We developed a novel method to track the position of participants' index finger and predict when they would perform a finger tap, before said tap actually occurred. Using this method, we set up a temporal order judgment task, wherein participants had to determine which of two brief stimuli occurred first. The experiment consisted of three conditions, each of which explored a different stimulus pair: a flash-tap condition, a beep-tap condition, and a beep-flash condition. We found that participants were most precise in the beep-tap condition, followed by the flash-tap condition. Participants' precision in the beep-flash condition was much poorer, to such an extent that our experiment could not accurately capture it. Our findings contribute to growing evidence that self-generated motor actions benefit precise temporal perception.

In **Chapter 4** we discuss a rhythmic tapping task, meant to explore the effect of perturbed feedback on participants' ability to synchronize with a simple guide rhythm across different sensory modalities. Most rhythmic tapping studies involve an auditory guide rhythm, and provide participants no feedback about their own performance (except for the haptic information given by tapping). Our study considered the effect of both visual and auditory guide stimuli. Each time the participant tapped their finger, they also caused a visual or auditory stimulus to occur. We used the same tap predictor as in the previous chapter to modify the exact timing of this feedback stimulus, such that it could occur too early or too late, relative to the finger tap. We found that in conditions where the modality of the guide and feedback differed, participants showed a greater bias and worse precision in trying to follow the guide than in conditions where the two modalities were the same. Particularly in the visual-auditory condition, the timing of participants' taps relative to the guide was dragged along with the offset of the feedback. Taken together, this indicates that participants rely more on the feedback in the cross-modal conditions than in unimodal conditions, and that they synchronize best with the guide when external stimuli are limited to a single sensory modality.

The thesis concludes with a task which requires a sense of both space and time to execute successfully. In **Chapter 5**, we carried out a study on interception in the tactile domain. Interception in humans is usually a visual endeavor, as our visual system is best suited to track fast and distant objects. It has previously been shown that people's ability to hit or catch a moving object can be predicted from how precisely they can determine both the object's position at any given time, as well as its speed. In our study, we sought to determine whether interception could also be executed on the basis of tactile information, and if so, whether people's performance similarly depended on a combination of position and velocity information. In each of our experiments we blindfolded participants and used a robotic device to deliver stationary or moving tactile stimuli to the participants' left arm. The study consisted of three experiments: one to see how precisely participants could locate a stationary target they felt on their arm, one to test their ability to move their hand at the same speed as a moving target, and finally, one wherein they had to hit the target as it moved along their arm. Our results showed that participants could locate the static target on the arm, albeit with relatively poor precision; that they were able to distinguish between the various velocities of the moving object; and that performance in the Interception experiment could be predicted based on performance in the Position and Velocity experiments, similarly to what is found in visual interception experiments.

Our studies have shed new light on how we perceive the state of the world around us. We have shown in **Chapters 2** and **5** that spatial perception and interception, while often informed via the visual system, are at the same time not always dependent on it. We found in **Chapter 3** that people are far more precise at comparing the timing of a self-generated motor action and that of an external stimulus than they are at comparing the timing of two external stimuli. In **Chapter 4** we showed that participants rely more on feedback during cross-sensory rhythmic tapping tasks than in unisensory ones, and that people's ability to maintain rhythm is hindered when they are asked to pay attention to external stimuli of more than one sensory modality.