All Birds Must Fly: The Experience of Multimodal Hands-free Gaming with Gaze and Nonverbal Voice Synchronization

Authored by:
  Ramin Hedeshy,
  Chandan Kumar,
  Mike Lauer,
 and Steffen Staab.

Introduction

Introduction –Background–

- The speech recognition technology is utilized in the gaming and the another applications.
- Eye gaze is also used as an input system in various fields.

- So, researchers thought that they can make the hands-free input system with combining speech recognition and eye gaze in the game environment.

- In this research, they investigate how these two methods help each other in offering hands-free input system.
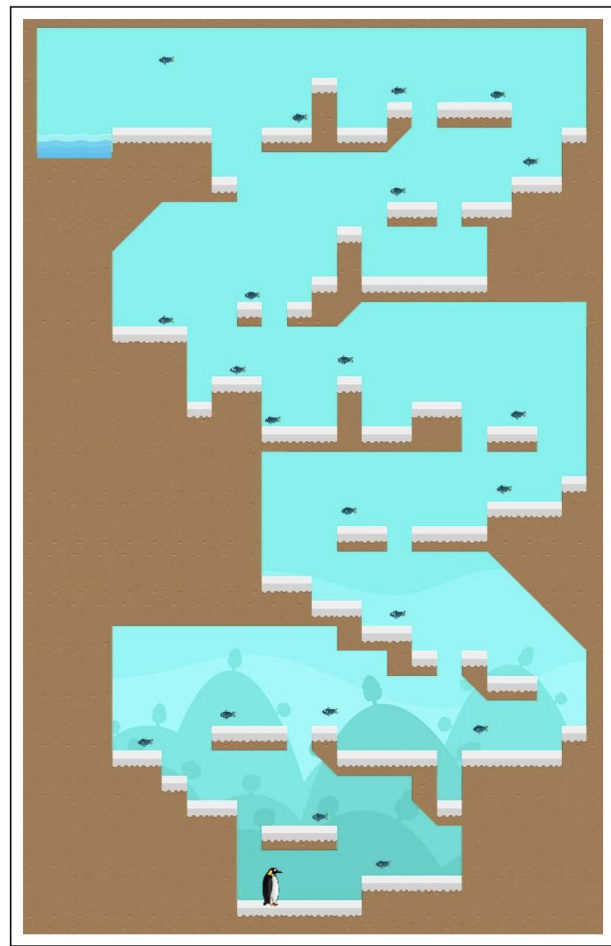
Method

Method – Game Environment –

- The environment is developed with Unity Engine by researchers.

  Title   : All Birds Must Fly

  Genre   : 2D Action Game (Like Super Mario)

  Actions   : Walk (to left, right) and Jump

  Interact : Mouse and Keyboard (MK), or Gaze and Voice (GV)

  Overview: 1. Player is set in the bottom of the stage.

  2. The goal of this game is to reach the top of the stage.

  3. Also, there are fish objects.

  To collect fish objects, player can get extra score.
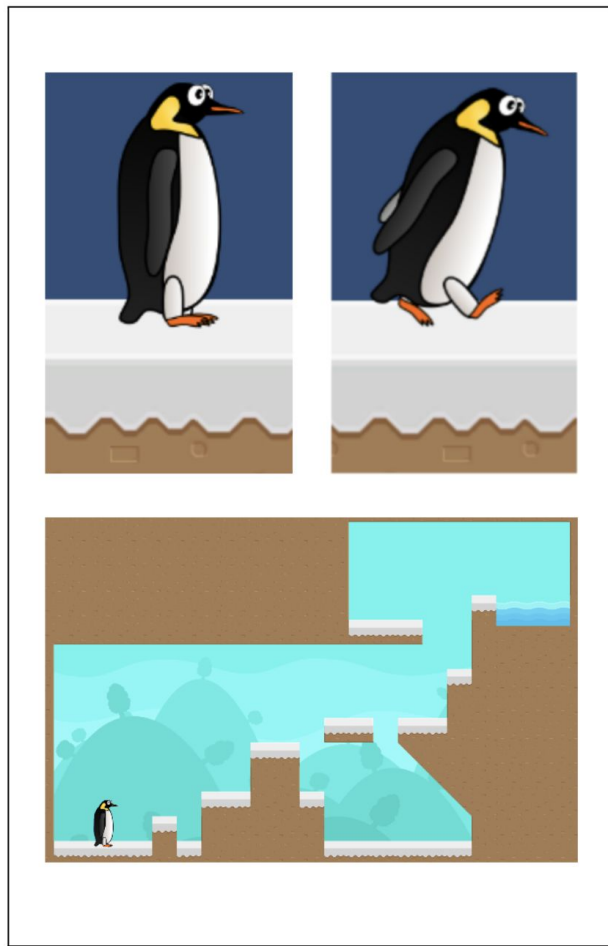
(a) Game Character and Training Level

(b) Main Level

Figure 1: Example images from the All Birds Must Fly game.

Method – Voice Input–

- In this research, especially they used the <u>non-verbal</u> voice detection.
  ＊non-verbal: the voice that does not include the meaning.
  (example: Humming, Volumes, Pitches, Buzzing, etc…)
- They used Humming for input because following reasons:
    1. It is easy to utter regardless in disability.
    2. It is easy to detect and fast.
- Humming input is used as the jump trigger.
- They used the microphone built-in the laptop as the device for the sound detection.
- They used the Autocorrelation for detecting humming.

Autocorrelation

- This is the technique that is used in the digital signal processing.

$$R_{ff}(\tau) = \bar{f}(-\tau) * f(\tau) = \int_{-\infty}^{\infty} f(t+\tau)\bar{f}(t)\,dt$$

$$= \int_{-\infty}^{\infty} f(t)\bar{f}(t-\tau)\,dt$$

- τ : lag

  f : it signify the wave of signal.

  f_bar: conjugate complex number function

Autocorrelation

$$R_{ff}(\tau) = \overline{f}(-\tau) * f(\tau) = \int_{-\infty}^{\infty} f(t+\tau)\overline{f}(t)\,dt$$

$$= \int_{-\infty}^{\infty} f(t)\overline{f}(t-\tau)\,dt$$

- In this research, the way to detect humming in game is not clearly mentioned, they only mentioned about the use of the autocorrelation.
  It is just my guess, they first correct humming data and calculate Rff on it, and in game playing they calculate Rff on real-time, then judge if it is humming or not with comparing these two results of Rff.

Method – Eye Gaze –

- They used Eye Tracker 4C developed by Tobbi.
- The frequency of tracking is 90Hz.
- Tracker is set on the bottom of screen.

Experiment

Experiments – Overview –

- They conducted two experiments:
  1. The comparison experiment between 'Gaze and nonverbal Voice (GV)' performance and 'Mouse and Keyboard (MK)' one.
  2. The investigation about the feasibility of the suggested input system (GV) for people who have disability in speech or physical.

Experiment 1 – Participants –

- There are 15 participants:

  Male   : 6

  Female : 9
- The mean of age is 24.4 years.
- 5 of them have game experience, the other have little experience.
- No participants have experience of the eye tracking and the voice detection.

Experiment 1 – Procedure –

- Place   : The laboratory of the university (Stuttgart, in Germany)
- Flow   : 1. The adjustment for each devices.
            2. Select one method as first they use.
            3. See the tutorial video of a method (MK or GV) .
            4. Play the game with a method in Training stage.
            5. Play it twice with a method in main stage.
            6. Take the questionnaire sheet on one of methods.
            7. Change the method to use and do one more time from 3 to 6.

- Time   : About 45 minutes

Experiment 2 – Participants –

- There are 10 participants from the special education schools or the facility for who have any disability:

    Male   : 5
    Female : 5

- Two of them had Ducheme Muscular Dystorophy disorder, two of them had Cerebral Palsy, and two of them had Spastic Tetraparesis.
  These disability cause trouble in motor skills (s.t. walk, jump, etc.)
- Three of them had speech disability. (they can humming)
- In the experiment, 4 participants had difficulty in playing cause of such as the need of devices for breathing, eye twitching, etc.

Experiment 2 – Procedure –

- Place   : The facility for who have disability.

- Flow   : It is almost same as the flow in experiment 1.
              The only difference is the flow about the method of MK.

Experiments – Questionnaire–

- The questionnaire is same between in experiment 1 and experiment 2.
- There are three sections:
    1. Participant Information
    2. Experience
    3. Control

- The factor of evaluation is scored from 1 to 5.
  1 is the lowest point, and 5 is the highest point.
- Some examples of questions are shown in next slide.

Experiments – Questionnaire–

- Participant Information

  Age, gender, job, and experience on game, technology in this research(GV).
- Experience

  How well do you think you played?

  Was it challenging, exciting, fun?
- Control

  Did you run out of breath?

  Did the penguin move exactly the way you wanted it to?

  The evaluation for each input (mouse, eye tracking, keyboard, and voice).

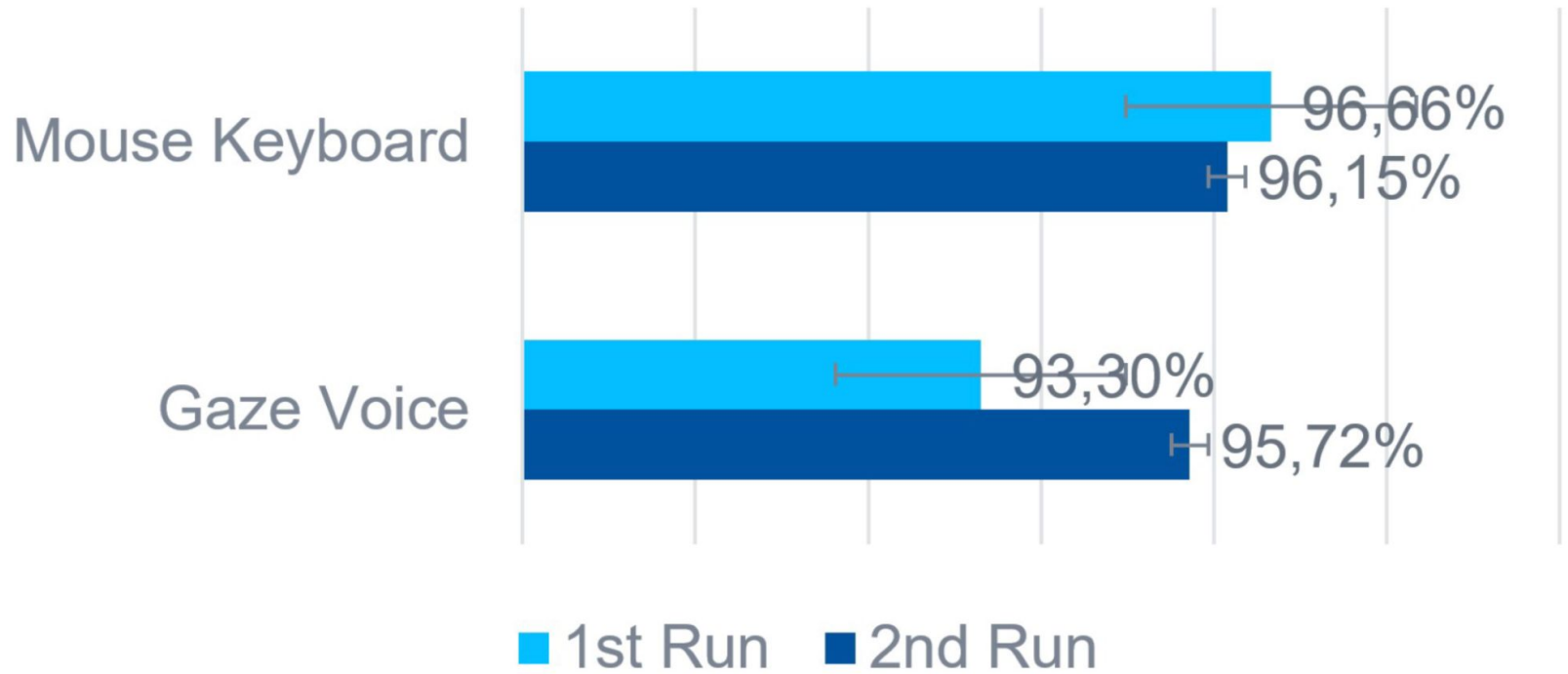 Was it exhausting?, How fast did you play?

Results (Ex 1)

Figure 3: Jump success rate for each method and session in Study 1, plotted as bars. Error margins indicate the standard deviation.

Figure 4: Mean completion time for each method and session in Study 1, plotted as bars. Error margins indicate the standard deviation.
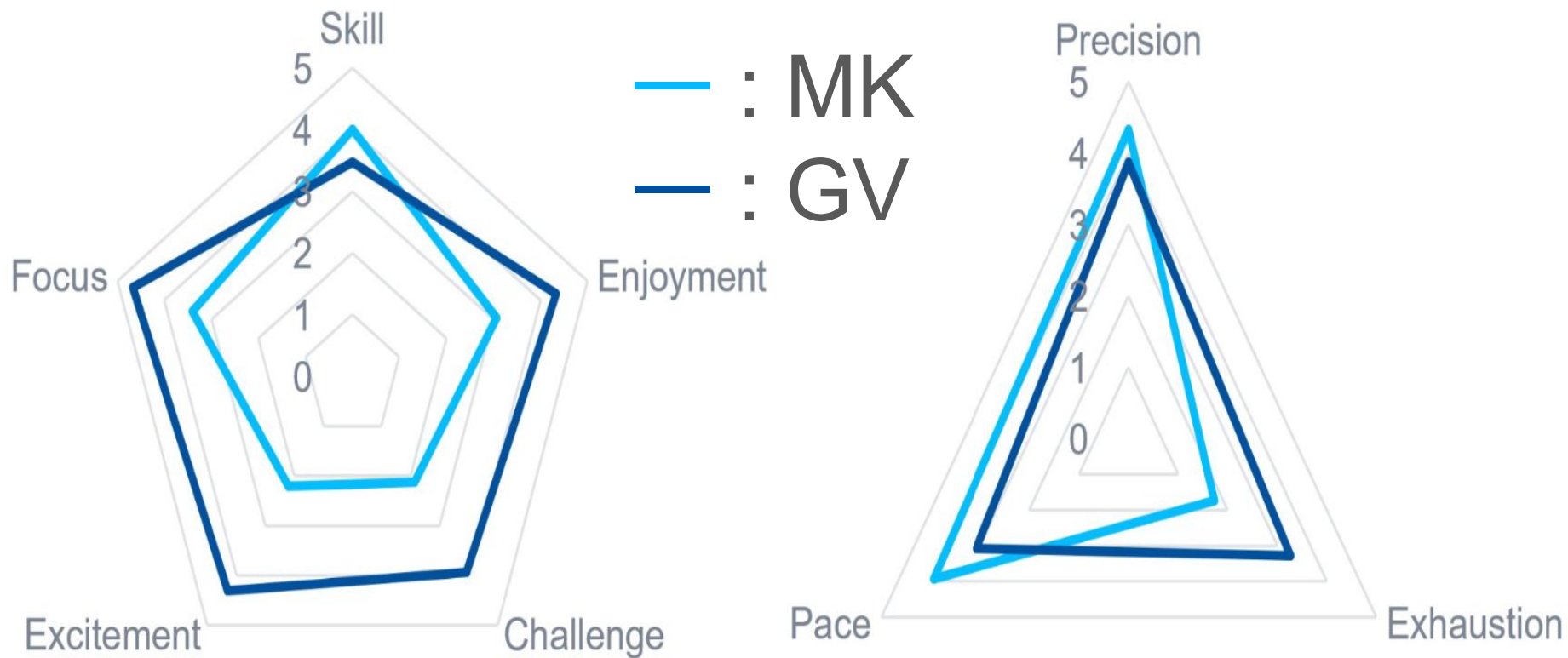
Figure 5: Subjective response (1 to 5) by method and questionnaire item in study 1.

Other result (there are no figure)

- The time between their movement.
  It is measured by measure the time they do not move.
- The mean time for action input.
- The proportion of motion (to all play time).
- The proportion of Walking action (to all motion time).
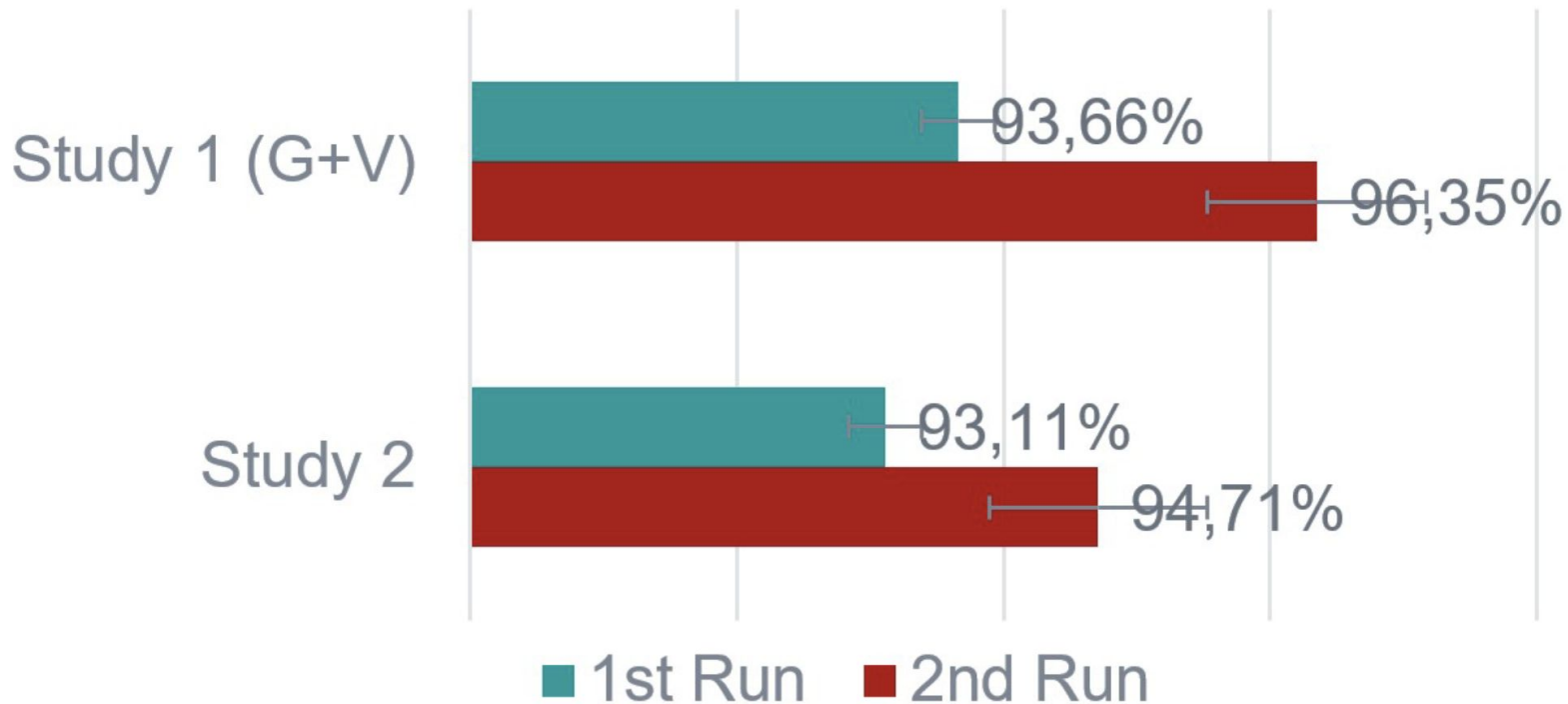- The question "which system do you prefer."

Results (Ex 2)

Figure 6: Jump success rate for each method and session in Study 2, plotted as bars. Error margins indicate the standard deviation.
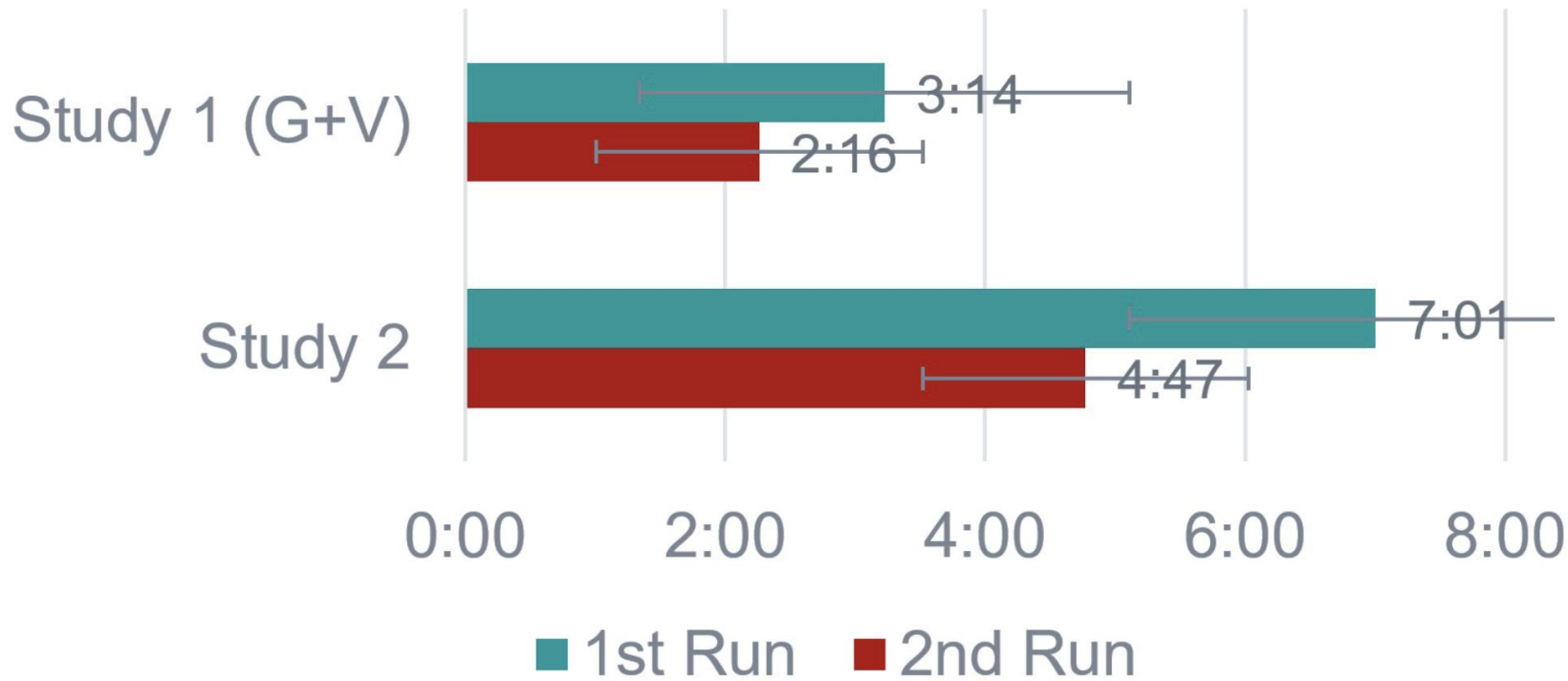
Figure 7: Mean completion time for each method and session in Study 2, plotted as bars. Error margins indicate the standard deviation.

Figure 8: Subjective response (1 to 5) compared to Study 1.

Other result (there are no figure)

- The time between their movement.
  It is measured by measure the time they do not move.
- The mean time for action input.
- The proportion of motion (to all play time).
- The proportion of Walking action (to all motion time).

- One of participants commented as following:
  "These hands free interaction methods is very useful for our (who have disability) life, including not only game environment."

Other result (there are no figure)

	MK	GV (Result 1)	GV (Result 2)
The time between movement	0.5 sec	1 sec	1.6 sec
The mean time for action input	650 ms	724 ms	755 ms
The proportion of motion	65%	49%	38%
The mean proportion of walking	21%	25%	18%
Preference	5 / 15 people	10 / 15 people	—-

Conclusion

Conclusion

- MK is more higher score in performance and precision than GV.
- Participants felt GV is more fun and immersive than MK, so the majority of them prefer GV.

- They stated that for future work, they want to use FPS game as experiment environment, and they want to create multiple actions corresponding to the length, pitch, or tones of voice.