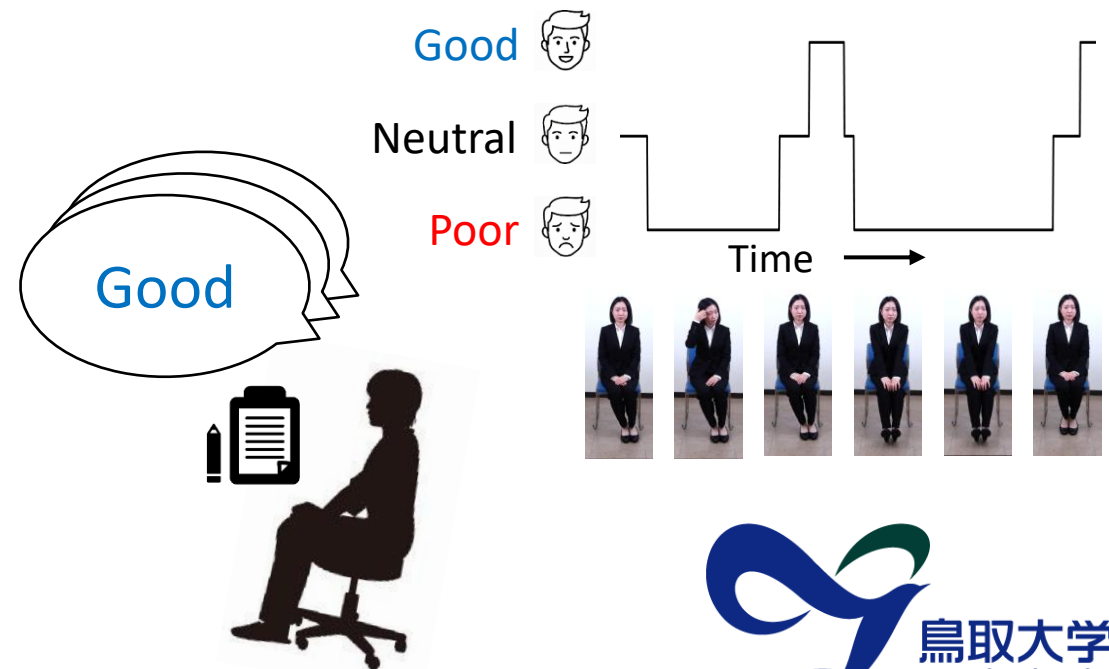


Investigating Temporal Changes of Impression Formed by Interviewers from the Applicant's Behavior

S. Hioki, M. Inoue, M. Nishiyama
Graduate School of Engineering
Tottori University, Japan



Introduction

An applicant's behavior strongly influences an interviewer's impression during recruitment and admission interviews.

Examples of behaviors performed by applicants in interviews



Making hand and body gestures



Time

Eye contact



Time

Nodding



Time



Time



Interviewer
(participant)

Previous studies

Post-viewing evaluation of impression

[Peeters+, International Journal of Selection and Assessment, 2006]

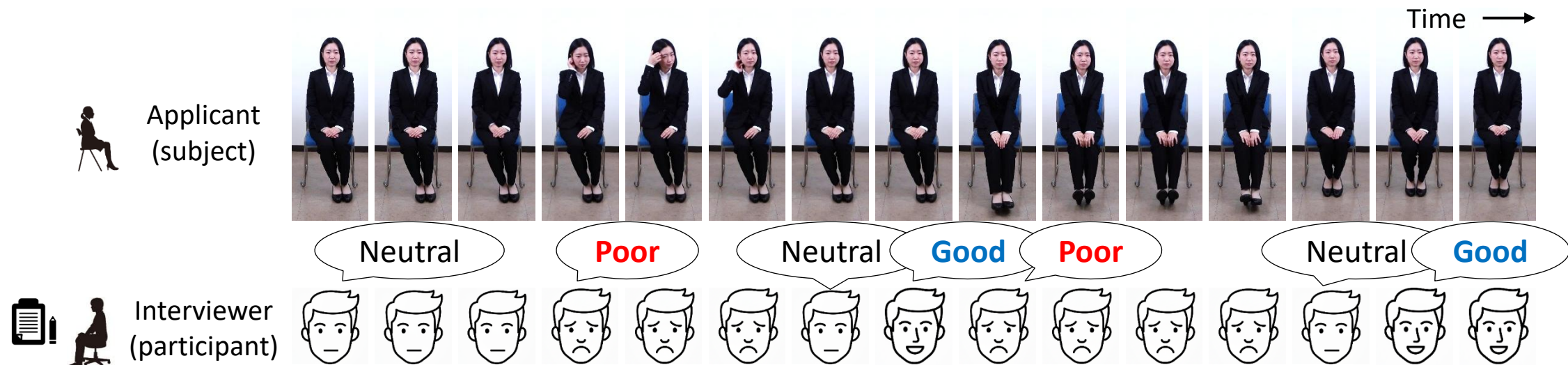
[DeGroot+, Journal of Business and Psychology, 2009]

1. Variation in interviewers' impressions related to social skills depending on the applicant's behavior
2. Statistically significant inter-rater reliability across participants

They evaluate impressions at the moment when each stimulus video sequence ends and thus can be regarded as investigating the impression formed over the entire time span of the stimulus video sequence.

Issues

However, previous studies using post-viewing evaluation of impression have not addressed how the impressions formed by participants who simulate interviewers change from moment to moment.



To investigate such moment-to-moment changes, we focus on the time series of impressions recorded at each time point of a stimulus video sequence for each participant.

Temporal changes of impression

Issue I1:

It is unclear to what extent the temporal changes of impression are shared across multiple participants.

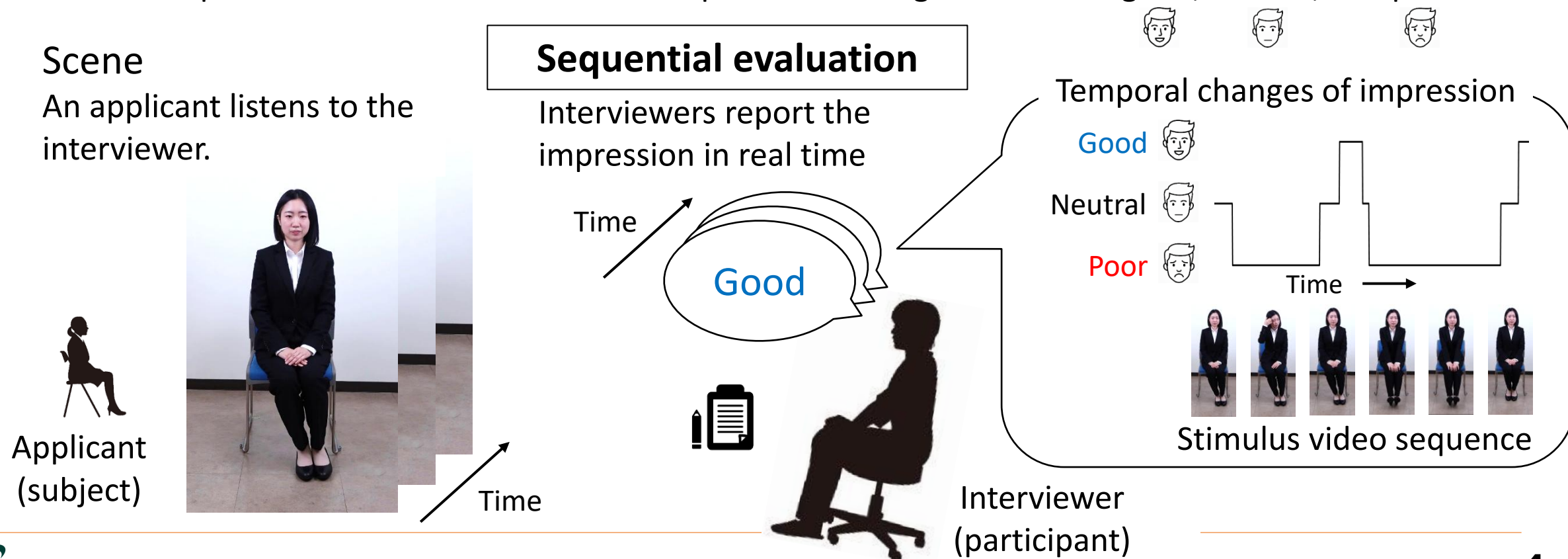
Issue I2:

There is no system that can record the temporal changes of impression while keeping participant burden low.

Purpose

We design a system to record the temporal changes of impression and investigate the extent to which the impression formed from the applicant's behavior is consistent over time across multiple interviewers.

The overall impression related to social skills is represented using three levels: good, neutral, and poor.



Design policy for Issue I1



It is unclear to what extent the temporal changes of impression are shared across multiple participants.

We investigate inter-rater reliability for the temporal changes of impression in the sequential evaluation.

Hypothesis: The temporal changes of impression obtained by the sequential evaluation show statistically significant inter-rater reliability across participants who simulated interviewers.



Time →

Stimulus video sequence

Good

Neutral

Poor



Good

Neutral

Poor



Good

Neutral

Poor



Time →

Temporal changes of impression



Participant 1



Participant 2



Participant 3

Similar

Different

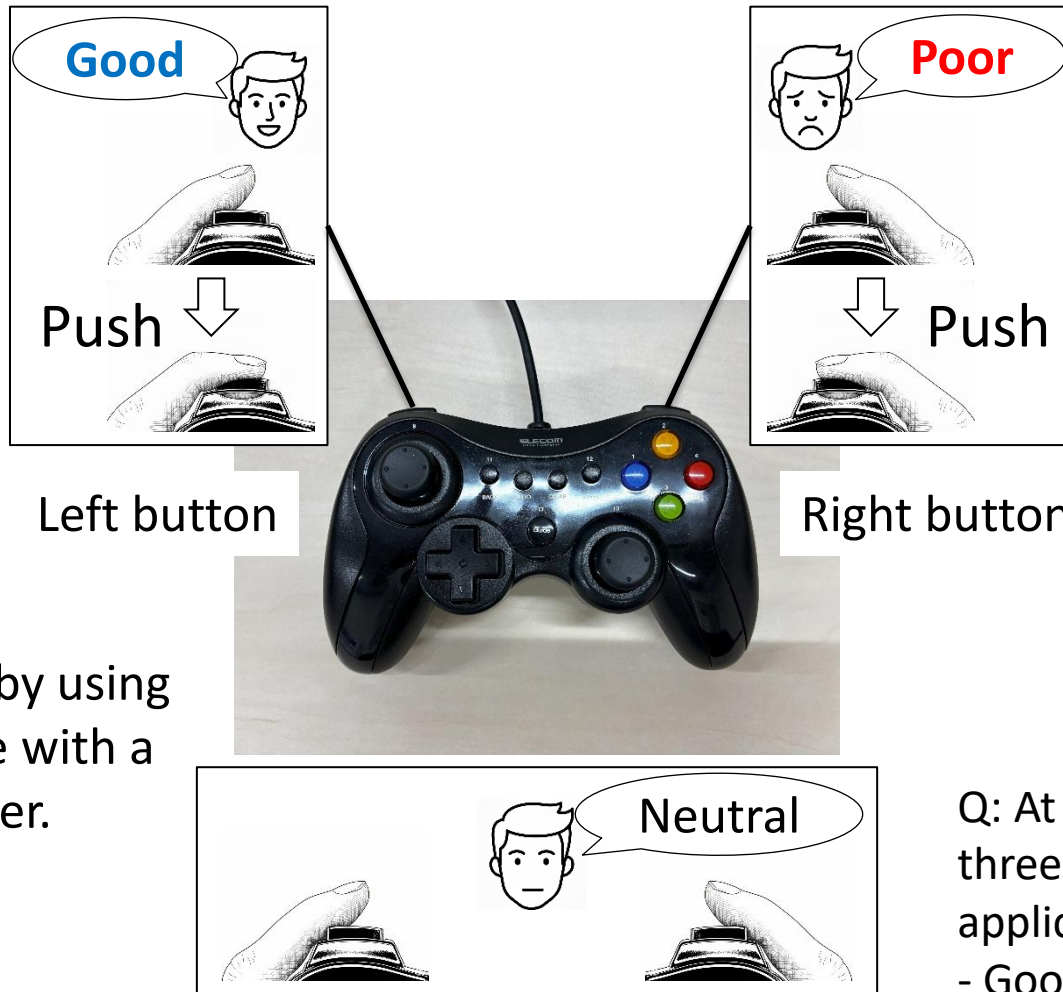
To investigate inter-rater reliability, we compute intraclass correlation coefficients (ICC).

Design policy for Issue I2

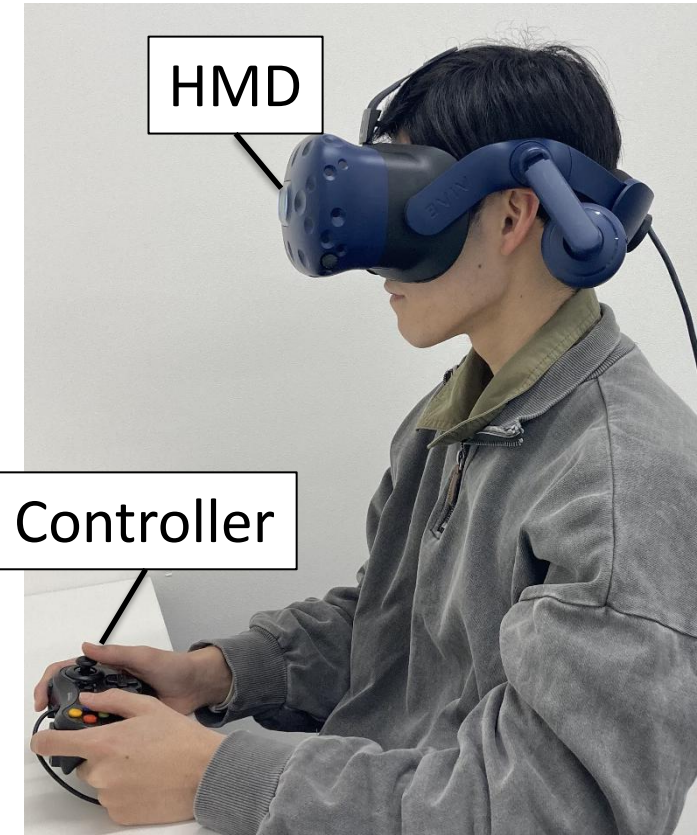


There is no system that can record the temporal changes of impression while keeping participant burden low.

We develop a system that allows participants to report the impression at each time point of a stimulus video sequence.



We aim to reduce cognitive burden by using a three-level scale with a handheld controller.



Q: At each time point, please select one of the following three options to report the impression you form of the applicant.
- Good, Neutral, or Poor.

Experimental conditions

20 participants who simulated interviewers (Japanese; 15 male and 5 female; mean age: 22.6 years) participated in the sequential evaluation.

The subjects in the stimulus video sequences were 20 university students simulating applicants (10 male and 10 female).

18 stimulus video sequences in $\tilde{S}_g$ in which the subject exhibited behavior intended to give a **good** impression.

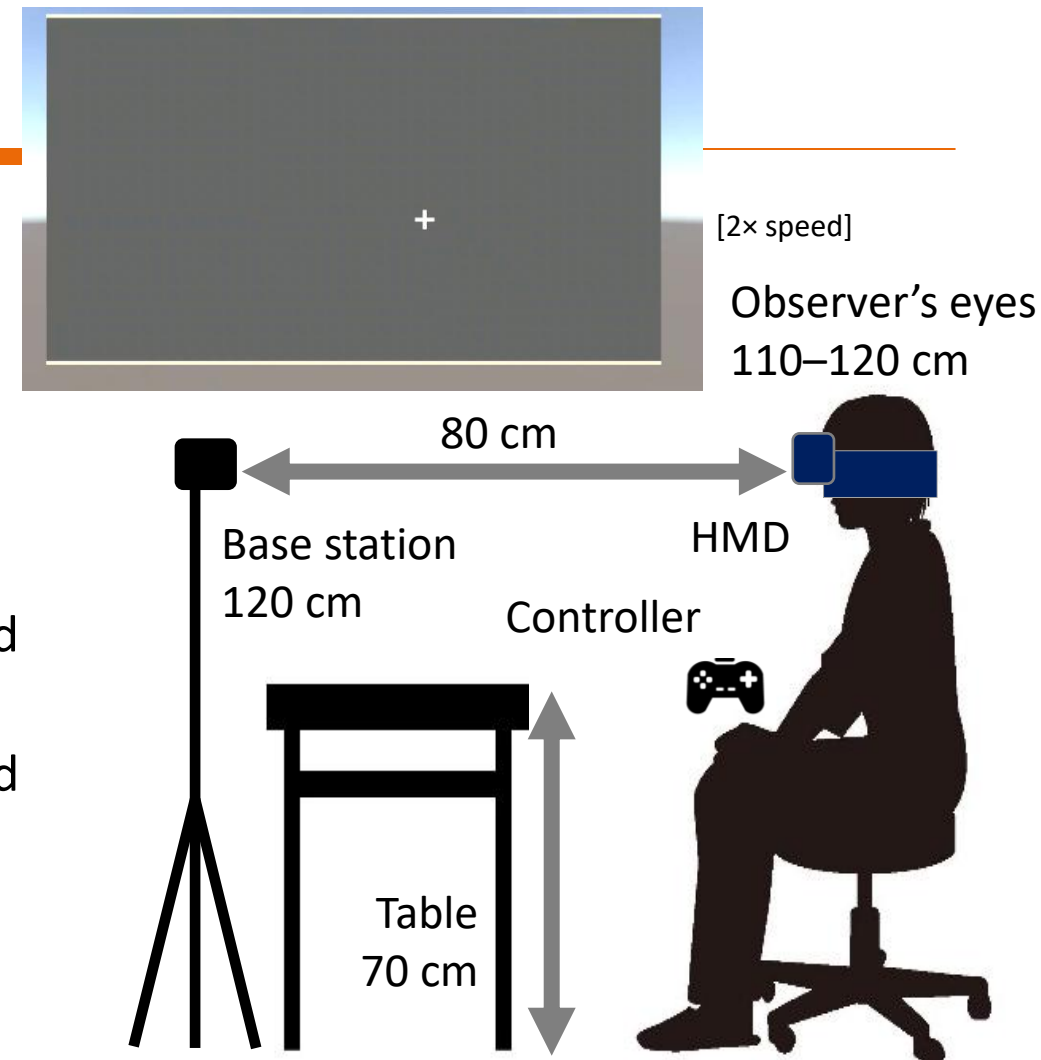
18 stimulus video sequences in $\tilde{S}_b$ in which the subject exhibited behavior intended to give a **poor** impression.

[1× speed]



$\tilde{S}_g$

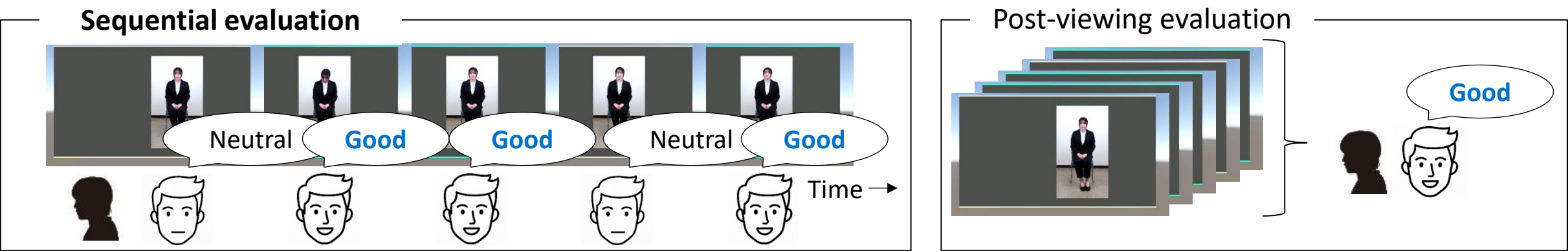
$\tilde{S}_b$



1. Convert the good option to 1, the neutral option to 0, and the poor option to -1.
2. Compute ICC(2,1) for the analysis of inter-rater reliability.

Results of inter-rater reliability

- I1:  It is unclear to what extent the temporal changes of impression are shared across multiple participants.
- Hypothesis: The temporal changes of impression obtained by the sequential evaluation show statistically significant inter-rater reliability across participants who simulated interviewers.

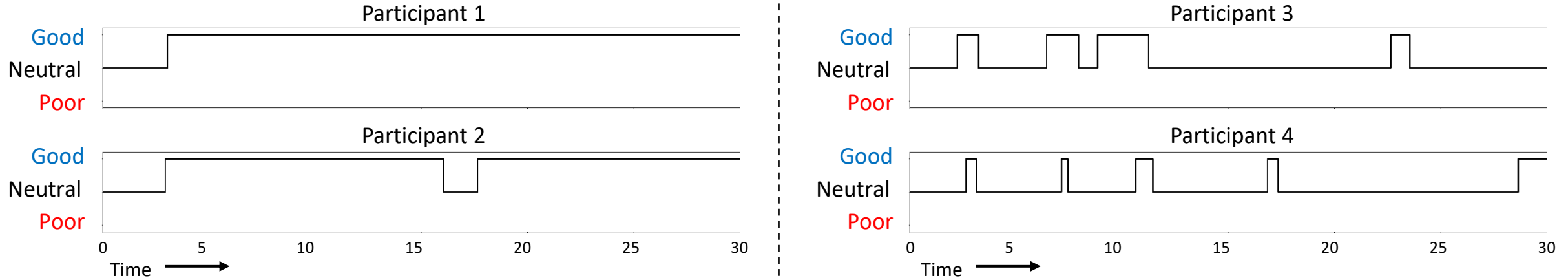


	ICC(2,1)	95% CI Lower	95% CI Upper	p-value	Interpretation of ICC
Sequential evaluation	0.568	0.553	0.578	< .001	Moderate
Post-viewing evaluation	0.885	0.817	0.923	< .001	Good

Although inter-rater reliability in the sequential evaluation was lower than that in the post-viewing evaluation, the temporal changes of impression formed from the applicant’s behavior showed moderate inter-rater reliability across participants who simulated interviewers.

Purpose of clustering

Inter-rater reliability does not explicitly reveal how impressions change over time.



To examine such temporal patterns, we consider applying clustering to the temporal changes of impression of individual participants.

By assigning each participant to a cluster, we aim to capture dominant patterns of the temporal changes of impression within each cluster.

We focus on the number of participants assigned to each cluster to identify dominant patterns.

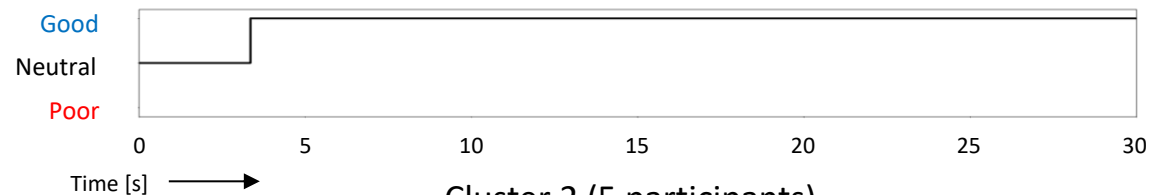
For clustering, we computed pairwise Manhattan distances between participants' temporal changes of impression and performed hierarchical clustering.

Clustering results

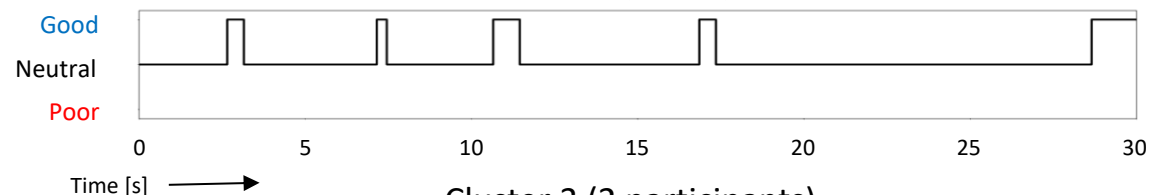
Stimulus video sequence with the **highest** total subjective score in the sequential evaluation



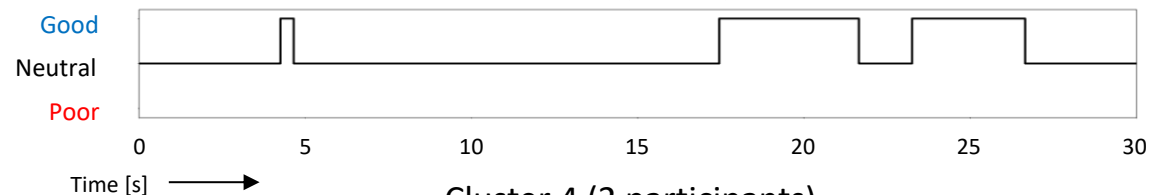
Cluster 1 (9 participants)



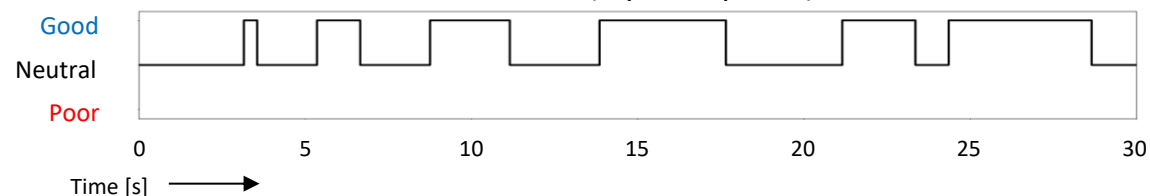
Cluster 2 (5 participants)



Cluster 3 (2 participants)



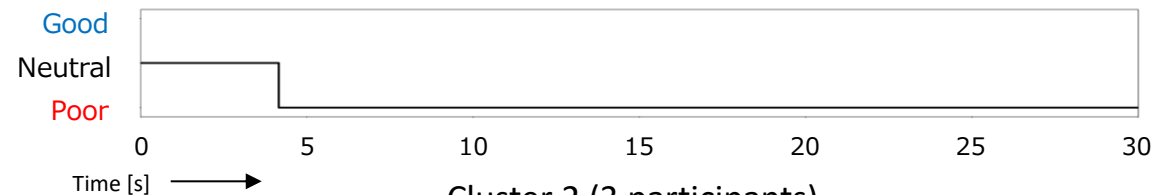
Cluster 4 (2 participants)



Stimulus video sequence with the **lowest** total subjective score in the sequential evaluation



Cluster 1 (13 participants)



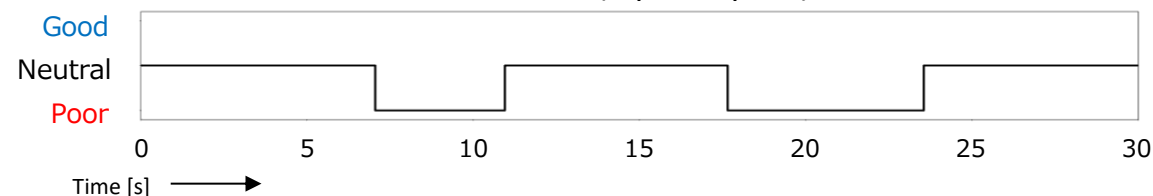
Cluster 2 (3 participants)



Cluster 3 (1 participant)



Cluster 4 (1 participant)

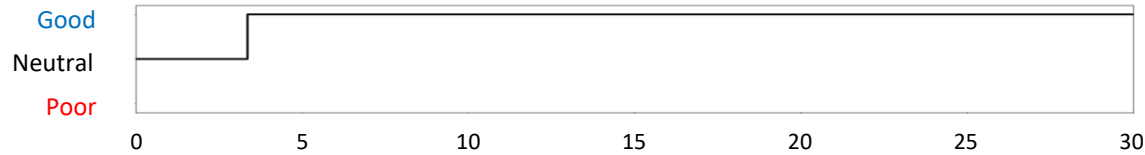


Dominant temporal patterns

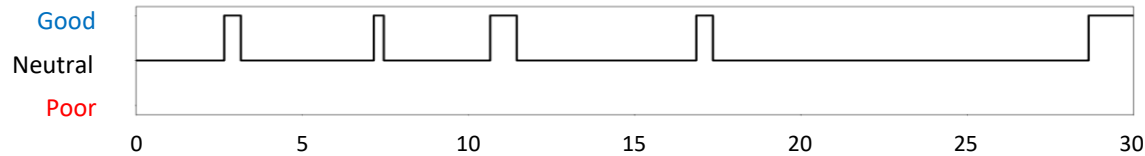
Stimulus video sequence with the **highest** total subjective score in the sequential evaluation



Cluster 1 (9 participants)



Cluster 2 (5 participants)



Cluster 3 (2 participants)



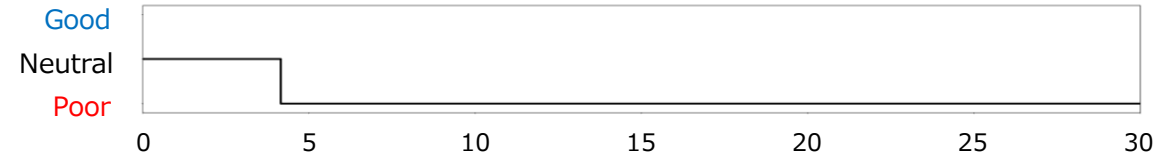
In Cluster 1, participants formed a good impression from the beginning and maintained a good impression until the end.

In Cluster 2, participants formed a good impression each time the applicant nodded.

Stimulus video sequence with the **lowest** total subjective score in the sequential evaluation



Cluster 1 (13 participants)

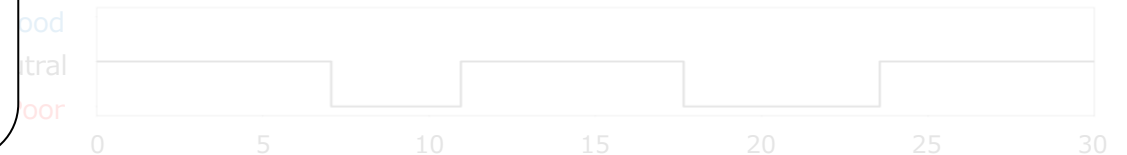


Cluster 2 (3 participants)



In Cluster 1, participants formed a poor impression from the beginning and maintained a poor impression until the end.

Cluster 4 (1 participant)



Qualitative evaluation of the developed system

I2:  There is no system that can record the temporal changes of impression while keeping participant burden low.

- We conducted a questionnaire survey with participants after the sequential evaluation and the post-viewing evaluation were completed.
- We asked participants to provide free-text comments on the system's usability and perceived burden.

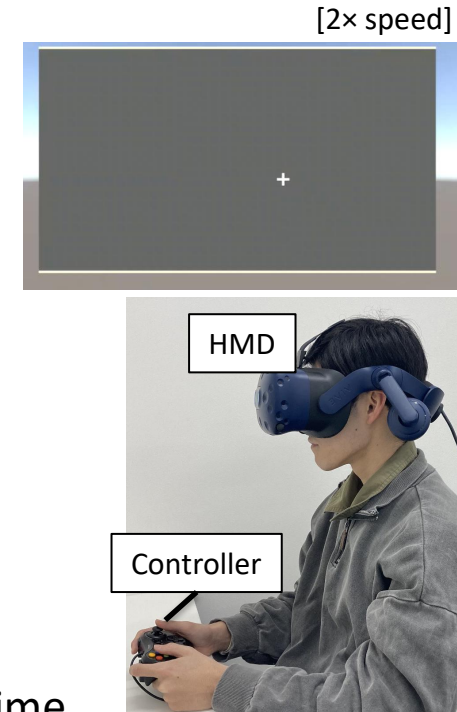
Usability

- 😊 A1: I quickly became familiar with which controller button corresponded to a good impression and which corresponded to a poor impression.
- 😊 A2: It was easy because I simply reported my impressions intuitively using the buttons, rather than thinking deeply.
- 😞 A3: Because it was my first time holding a game controller, I was slightly confused about the operation.

Burden

- 😊 A4: Because the video sequence did not stop partway through, I could stay focused and report my impressions as natural reactions.
- 😊 A5: I did not feel any particular burden compared with the post-viewing evaluation, and I felt that I could continue even during a long experiment.
- 😞 A6: The HMD was heavy, and my head became tired when using it for a long time.

Our system shows strong potential to record the temporal changes of impression while keeping participant burden low.



Conclusions

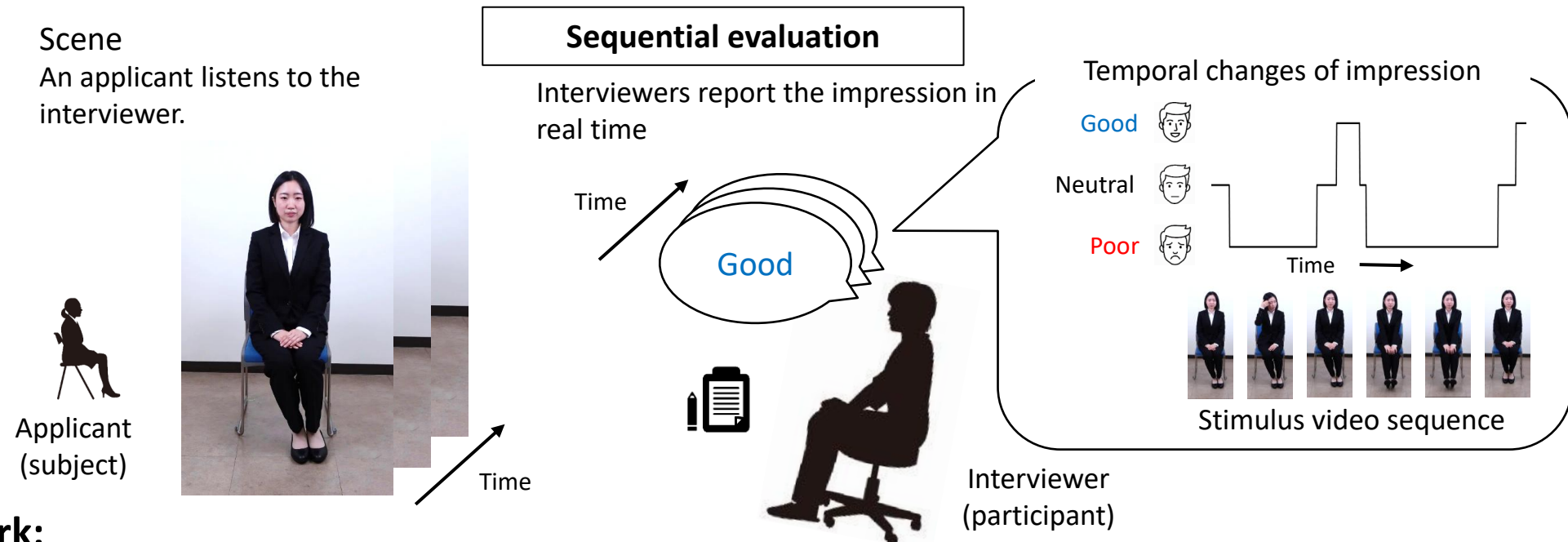
- We have developed a system that records the temporal changes of impression that interviewers form from the applicant's behavior during viewing of a stimulus video sequence.
- We have also investigated inter-rater reliability across participants who simulated interviewers.



Moderate inter-rater reliability ($ICC = 0.568, p < .001$)



Many participants reported no substantial burden



Future work:

- Analyze the relationship between the temporal changes of impression and the applicant's behavior
- Increase the diversity of participants to improve the reliability of the findings