Conducting Economic Experiments at Multiple Sites: Subjects' Cognitive Ability and Attribute Information

K. Ogawa[1], Y. Osaki[2], T. Kawamura[3], H. Takahashi[4], S. Taguchi[5], Y. Fujii[6], N. Watanabe[7]

[1]Kansai U., [2]Waseda U., [3]Tezukayama U., [4]Hiroshima City U.
[5]Doshisha U. [6]Meiji U. [7]Keio U.

March 18, 2021 at Osaka U. and Kansai U. (on-line)
International WS on Lab and Field Exp.

1. Introduction: Research Question

- **Our Question**: How can we obtain the **homogeneity of subjects' behavior** when they are at **different sites**?
  - Snowberg and Yariv (2021, AER): the observed behavior in university students in the U.S. are bounded in comparison to those in a representative sample of the U.S. population and participants in Amazon MTurk.
  - We concentrate our attention to university students in Japan.

  Why?: increasing needs for large samples of subjects at small or middle-scaled schools in research and education

1. Introduction: How to answer?

(1) propensity score matching with the following covariates

- cognitive ability: correct answer rate of 16 questions excerpted from the Raven's APM test (Raven score)
- general attribute information (age, gender, department)

Why not using regression analysis with dummy variables?
Ans. Our Question is not to estimate fitted values but to confirm the homogeneity of subjects' behavior.

(2) bandit experiment with a context of weighted voting

- not an interdependent decision-making but an individual decision-making $\Leftarrow$ **Our Question**
- weighted voting: difficult to realize the underlying structure of payoffs (Guerci et al., 2017, TD) $\Leftarrow$ **Our Question**

1. Introduction: Main Results

(1) Among schools located in the **same region**, the homogeneity is available relatively **easily**.

- Kansai U-A vs. Osaka SU: homog. with Raven score
- Kansai U-B vs. Doshisha: homog. with Raven score and sci-eng

(2) In the **different regions**, **however**, it is **difficult** to homogenize subjects' behavior.

- Kansai U-B vs. Horoshima CU: not homog. in any ways
- Hiroshima CU vs. Doshisha: homog. with Raven score, econ, and sci-eng.

An Implication: needs for some variables which represent the difference in regional features: subjects' home town?

2. Experimental Design: Binary Choice Problem

- Subjects choose one of two weighted voting games.
  - $[q; v_1, v_2, v_3, v_4]$; q is the quota, v_i is the voting weight.
  - Each subject acts as Member 1, who has v_1.
- clearly informed that the other three members are all fictitious.
- one binary choice problem for the first 40 periods, and a similar but different one in the following 20 periods; A $\rightarrow$ B, B $\rightarrow$ A, C $\rightarrow$ D, D $\rightarrow$C
- payoff generating function: Deagan-Packel index
  - The correct answer is Choice 2.

Table: Binary choice problems and expected payoffs for Member 1

Problem	Choice 1	(expected payoff)	Choice 2	(expected payoff)
A	$[14; \mathbf{5}, 3, 7, 7]$	$(120 \times 2/3 \times 1/3)$	$[14; \mathbf{5}, 4, 6, 7]$	$(120 \times 3/4 \times 1/3)$
B	$[6; \mathbf{1}, 2, 3, 4]$	$(120 \times 1/3 \times 1/3)$	$[6; \mathbf{1}, 1, 4, 4]$	$(120 \times 2/3 \times 1/3)$
C	$[14; \mathbf{3}, 5, 6, 8]$	$(120 \times 2/3 \times 1/3)$	$[14; \mathbf{3}, 6, 6, 7]$	$(120 \times 3/4 \times 1/3)$
D	$[9; \mathbf{1}, 3, 5, 6]$	$(120 \times 1/3 \times 1/3)$	$[9; \mathbf{1}, 2, 6, 6]$	$(120 \times 2/3 \times 1/3)$

- three feedback treatments: (1) no feedback (No-fb), (2) partial feedback (Part-fb), and (3) full feedback (Full-fb).
- a 30-second time limit for the choice stage, and a 10-second limit for the feedback stage
  - no choice within 30 seconds $\rightarrow$ zero point
  - "Wait" message is given in feedback stage for No-fb.
- 1 point is converted 1 JPY.
- zTree (Fischbacher, 2007) is used.

Subject's Monitor

Please choose one out of the following two committees (Choice 1 or Choice 2). Each committee decides a distribution of 120 points among four members. You are Member 1.
In both committees, 22 votes are apportioned to those members and you have 5 votes.
Any proposals of point distributions need 14 votes in favor to be adopted.

$$\text{Choice 1} \quad [14; \mathbf{5}, 3, 7, 7], \quad \text{Choice 2} \quad [14; \mathbf{5}, 4, 6, 7]$$

When subjects choose Choice 2 and MWC $(5, 6, 7)$ appears, they see, for instance, the following results on their monitor, regardless of any treatments.

You chose the following committee.

$$\text{Choice 2: } [14; \mathbf{5}, 4, 6, 7].$$

Next, in the full-feedback treatment, subjects see

The committee decided to distribute 120 points this time as follows.
You obtained 40 points this time.

(**40**, 0, 40, 40)

on their monitors. In the partial-feedback treatment, the payoff distribution is not shown, but rather the following note is shown on their monitors:

You have obtained 40 points this time.

In the no-feedback treatment, the payoff distribution is not shown and simply

Please wait for a while.

is shown on the subjects' monitors.

- Raven test: measuring subjects' ability of pattern recognition, which fits our bandit experiment.
  - In each question of the test, eight patterns are drawn, and the subject selects a pattern that matches those visual patterns from the options.
  - Our subjects are asked to learn the pattern of payoffs
- We used 16 questions excerpted from the Raven's APM test.
  - The APM version is composed of 48 questions in total.
  - Subjects answered 16 questions within 10 minutes after completing the binary choice problems.
  - Colored Progressive Matrices (CPM), Standard Progressive Matrices (SPM), and Advanced Progressive Matrices (APM), in ascending order of difficulty.
    - Gill and Prowse (2016, JPE) and Basteck and Mantovani (2018, GEB) used the SPM version, while Proto et al. (2019, JPE) used the APM version.
  - Guerci et al. (2017, TD) used the same 16 questions of the Raven APM test.

2. Experimental Design: Session Information

- March 2, 2018 to October 17, 2019
- Kansai University, Osaka Sangyo University, Doshisha University, Hiroshima City University
- undergraduate students who do not know indices of voting power. In total, 816 subjects participated.
- the average amount paid as a reward was 2534 JPY (1 USD was about 110 JPY at that time).
- rewards paid for participation (not incentive payment) was different at Kansai U-A and Osaka SU (Group A) from those of Kansai U-B, Doshisha, and Hiroshima CU (Group B).
- The experimenter is the same for Group A.
  A text-to-speech software was used for Group B.

Table: subjects' attribute information (p-val for Fisher exact test)

site	# of subj.	male	female	p-val	econ	sci-eng	others
Kansai U-A	240	128	112		21	53	166
Osaka SU	120	99	21	**<0.001**	47	38	35
Kansai U-B	197	98	99		24	41	132
Doshisha	135	76	56	0.828	21	0	114
Hiroshima CU	124	49	75	**0.047**	0	24	100

Table: Raven scores of subjects: basic statistics

site	# of subj.	mean	std.dev.	p-value	min	max
Kansai U-A	240	11.208	2.170		3	16
Osaka SU	120	10.625	3.041	**0.027**	3	16
Kansai U-B	197	11.518	2.398		2	15
Doshisha	135	11.578	2.300	0.890	5	16
Hiroshima CU	124	10.976	2.441	**0.038**	3	15

difference between two sites

- Kansai U-A vs. Osaka SU: Raven, gender
- Kansai U-B vs. Doshisha: sci-eng
- Kansai U-B vs. Hiroshima CU: region, Raven, gender, econ
- Hiroshima CU vs. Doshisha: region, Raven, gender, econ, sci-eng

Table: Raven scores of general public: basic statistics

test site	# of subj.	mean	std. dev.	min	max
EEL at Kansai U	1,057	7.986	3.326	0	16

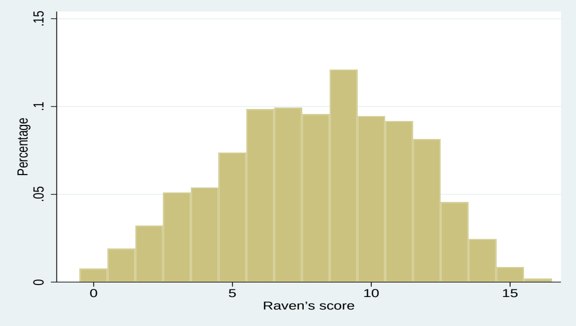

Figure: Histogram of Raven scores of non-student general public

Hypothesis: In a two-armed bandit experiment in the context of weighted voting, there is no significant difference in subjects' choices made at different experimental sites, after we control for subjects' cognitive ability scores in addition to their attribute information (age, gender, whether they were economics students, and whether they were science or engineering students) collected in the course of standard subject management.

3. Results: Criteria

We say that subjects' behavior was homogenized by introducing subjects' Raven scores if the following criteria are satisfied.

- The null hypothesis is that the rates of correct answers are the same after the covariate adjustment.
- The answers to all binary choice problems are pooled at each period, because of small samples in Group B.

Criterion 1: For the data that integrate all feedback treatments, we cannot reject the null hypothesis for every block of 10 consecutive periods when subjects' Raven scores are introduced to all covariates regarding their attribute information.

Criterion 2: For every feedback treatment, (1) the number of the rejection of the null hypothesis does not increase as compared to the case where subjects' Raven scores are not introduced, and (2) there are at most two treatments in which we cannot reject the null hypothesis for every block of 10 consecutive periods.

3. Results: Kansai U-A vs.Osaka SU

Result 1: For subjects at Kansai U-A and Osaka SU, who differed in terms of Raven and gender, their behavior was homogenized by Raven.

Table: p-values for z-test: Kansai U-A vs. Osaka SU, integrated data

Covariates	1 — 10	11 — 20	21 — 30	31 — 40	41 — 50	51 — 60
age	0.095	0.302	0.728	0.871	**0.015**	**0.048**
gender	0.734	0.680	0.534	0.148	**0.050**	0.157
age, gender	0.640	0.975	0.492	0.564	0.069	0.228
age, econ, sci-eng	0.517	0.281	0.750	0.740	**0.004**	0.066
gender, econ, sci-eng	0.929	0.719	0.267	0.366	0.087	0.423
age, gender, econ, sci-eng	0.948	0.635	0.662	0.965	**0.005**	**0.043**
Raven	0.372	0.252	0.529	0.759	**0.014**	**0.013**
Raven, age	0.062	0.977	0.952	0.535	**0.016**	0.180
Raven, gender	0.483	0.914	0.695	0.789	**0.025**	0.145
Raven, age, gender	0.822	0.532	0.776	0.275	0.079	0.471
Raven, age, econ, sci-eng	0.591	0.949	0.957	0.423	0.065	0.297
Raven, gender, econ, sci-eng	0.934	0.646	0.622	0.229	0.407	0.209
Raven, age, gender, econ, sci-eng	0.937	0.648	0.532	0.323	0.295	0.387

Table: p-values for z-test: Kansai U-B vs. Osaka SU, No-fb

Covariates	1 − 10	11 − 20	21 − 30	31 − 40	41 − 50	51 − 60
age	0.223	0.529	0.717	0.970	0.272	0.197
gender	0.805	0.637	0.120	**0.031**	0.531	0.313
age, gender	0.710	0.383	**0.001**	$<$**0.001**	0.660	0.596
age, econ, sci-eng	0.255	0.791	0.870	0.587	0.189	0.169
gender, econ, sci-eng	0.979	0.649	0.081	**0.011**	0.424	0.461
age, gender, econ, sci-eng	0.816	0.567	**0.007**	**0.002**	0.386	0.323
Raven	0.483	0.393	0.893	0.965	0.072	**0.013**
Raven, age	0.061	0.523	0.920	0.790	0.111	0.055
Raven, gender	0.723	0.729	0.237	0.059	0.653	0.376
Raven, age, gender	0.842	0.377	**0.001**	$<$**0.001**	0.511	0.402
Raven, age, econ, sci-eng	**0.028**	0.268	0.805	0.506	0.746	0.582
Raven, gender, econ, sci-eng	0.952	0.457	0.458	0.081	0.319	0.340
Raven, age, gender, econ, sci-eng	0.994	0.665	**0.003**	$<$**0.001**	0.932	0.994

Table: p-values for z-test: Kansai U-B vs. Osaka SU, Part-fb

Covariates	1 — 10	11 — 20	21 — 30	31 — 40	41 — 50	51 — 60
age	0.074	0.082	0.493	0.706	0.201	0.947
gender	0.391	0.529	0.288	0.588	0.094	0.329
age, gender	0.322	0.340	0.560	0.619	0.120	0.669
age, econ, sci-eng	0.816	0.390	0.268	0.825	0.252	0.791
gender, econ, sci-eng	0.689	0.411	0.229	0.851	0.175	0.848
age, gender, econ, sci-eng	0.198	0.145	0.186	0.263	0.232	0.915
Raven	0.086	0.148	0.130	0.087	0.128	0.917
Raven, age	0.546	0.375	0.251	0.914	0.153	0.351
Raven, gender	0.380	0.551	0.233	0.251	0.152	0.521
Raven, age, gender	0.457	0.223	0.203	0.581	**0.044**	0.291
Raven, age, econ, sci-eng	0.962	0.637	0.202	0.989	**0.024**	0.129
Raven, gender, econ, sci-eng	0.829	0.758	**0.016**	0.678	0.124	0.502
Raven, age, gender, econ, sci-eng	0.817	0.843	0.207	0.846	0.303	0.887

Table: p-values for z-test: Kansai U-B vs. Osaka SU, Full-fb

Covariates	1 — 10	11 — 20	21 — 30	31 — 40	41 — 50	51 — 60
age	0.784	0.870	0.520	0.481	**0.040**	0.061
gender	0.835	0.571	0.843	0.890	0.310	0.611
age, gender	0.901	0.324	0.966	0.992	0.069	0.225
age, econ, sci-eng	0.625	0.966	0.410	0.170	0.455	0.296
gender, econ, sci-eng	0.788	0.645	0.734	0.293	0.294	0.147
age, gender, econ, sci-eng	0.630	0.950	0.401	0.164	0.467	0.307
Raven	0.759	0.964	0.475	0.425	0.173	0.654
Raven, age	0.452	0.660	0.615	0.813	0.284	0.812
Raven, gender	0.367	0.963	0.807	0.730	0.080	0.437
Raven, age, gender	0.490	0.562	0.997	0.993	0.101	0.375
Raven, age, econ, sci-eng	0.644	0.906	0.849	0.404	0.422	0.129
Raven, gender, econ, sci-eng	0.984	0.412	0.445	0.962	0.284	0.101
Raven, age, gender, econ, sci-eng	0.488	0.980	0.825	0.319	0.466	0.126

3. Results: Kansai U-B vs.Doshisha

Result 2: For subjects at Kansai U-B and Doshisha, who differed in terms of sci-eng, their behavior was homogenized by Raven and sci-eng.

Table: p-values for z-test: Kansai U-B vs. Doshisha, integrated data

Covariates	1 − 10	11 − 20	21 − 30	31 − 40	41 − 50	51 − 60
age	0.648	0.364	**0.024**	0.339	0.834	0.432
gender	0.793	0.866	0.157	0.467	0.691	0.100
age, gender	0.570	0.745	0.888	0.649	0.517	0.741
age, econ	0.676	0.268	0.055	0.319	0.664	0.441
gender, econ	0.758	0.891	0.138	0.597	0.727	0.109
age, gender, econ	0.972	0.906	0.430	0.622	0.606	0.368
Raven	0.878	0.939	0.135	0.472	0.513	0.050
Raven, age	0.539	0.996	0.093	0.661	0.656	0.133
Raven, gender	0.509	0.659	0.134	0.770	0.753	0.132
Raven, age, gender	0.243	0.320	0.355	0.377	0.957	0.087
Raven, age, econ	0.745	0.887	0.341	0.731	0.959	0.122
Raven, gender, econ	0.292	0.773	0.174	0.961	0.431	**0.044**
Raven, age, gender, econ	0.769	0.924	0.371	0.697	0.537	0.051

Table: p-values for z-test: Kansai U-B vs. Doshisha, No-fb

Covariates	1 — 10	11 — 20	21 — 30	31 — 40	41 — 50	51 — 60
age	0.693	0.421	0.470	0.515	0.765	0.563
gender	0.580	0.629	0.779	0.657	0.876	0.885
age, gender	0.775	0.722	0.791	0.924	0.496	0.317
age, econ	0.533	0.496	0.596	0.549	0.689	0.508
gender, econ	0.370	0.973	0.904	0.809	0.892	0.971
age, gender, econ	0.745	0.657	0.860	0.752	0.576	0.317
Raven	0.344	0.608	0.667	0.464	0.936	0.582
Raven, age	0.318	0.806	0.942	0.705	0.383	0.308
Raven, gender	0.287	0.993	0.868	0.901	0.964	0.889
Raven, age, gender	0.920	0.479	0.961	0.941	0.841	0.926
Raven, age, econ	0.154	0.417	0.665	0.591	0.801	0.792
Raven, gender, econ	0.277	0.983	0.959	0.995	0.813	0.981
Raven, age, gender, econ	0.307	0.321	0.993	0.692	0.774	0.842

Table: p-values for z-test: Kansai U-B vs. Doshisha, Partial-fb

Covariates	1 — 10	11 — 20	21 — 30	31 — 40	41 — 50	51 — 60
age	0.127	0.886	0.194	0.989	0.865	0.415
gender	0.396	0.236	0.144	0.582	0.754	0.274
age, gender	0.453	0.780	0.607	0.604	0.854	0.414
age, econ	0.097	0.988	0.119	0.815	0.879	0.282
gender, econ	0.259	0.362	0.065	0.466	0.671	0.226
age, gender, econ	0.574	0.516	0.548	0.678	0.898	0.312
Raven	0.549	0.235	0.235	0.880	0.710	0.272
Raven, age	0.377	0.488	0.095	0.338	0.324	0.100
Raven, gender	0.490	0.216	0.237	0.429	0.580	0.132
Raven, age, gender	0.200	0.134	0.077	0.717	0.572	0.201
Raven, age, econ	0.210	0.519	0.144	0.524	0.665	0.205
Raven, gender, econ	0.838	0.106	0.363	0.730	0.576	0.321
Raven, age, gender, econ	0.898	0.232	0.159	0.519	0.456	0.101

Table: p-values for z-test: Kansai U-B vs. Doshisha, Full-fb

Covariates	1 − 10	11 − 20	21 − 30	31 − 40	41 − 50	51 − 60
age	0.898	0.671	0.145	0.801	0.667	0.095
gender	0.568	0.810	0.260	0.768	0.521	0.060
age, gender	0.950	0.507	0.230	0.228	0.984	0.078
age, econ	0.398	0.916	0.132	0.608	0.637	0.134
gender, econ	0.376	0.842	0.257	0.832	0.514	0.071
age, gender, econ	0.526	0.743	0.452	0.376	0.716	0.060
Raven	0.761	0.871	0.485	0.980	0.685	**0.033**
Raven, age	0.347	0.406	0.129	0.803	0.395	**0.049**
Raven, gender	0.824	0.453	0.108	0.926	0.644	0.170
Raven, age, gender	0.838	0.461	0.295	0.892	0.969	0.120
Raven, age, econ	0.959	0.450	0.091	0.981	0.342	**0.047**
Raven, gender, econ	0.414	0.925	0.297	0.845	0.115	**0.012**
Raven, age, gender, econ	0.464	0.159	**0.035**	0.516	0.932	0.188

Table: p-values for z-test: Kansai U-B (excl. sci-eng) vs. Doshisha, Full-fb

Covariates	1 — 10	11 — 20	21 — 30	31 — 40	41 — 50	51 — 60
Raven	0.792	0.777	0.447	0.757	0.373	**0.024**
Raven, age	0.448	0.912	0.312	0.968	0.578	0.979
Raven, gender	0.979	0.730	0.329	0.944	0.711	0.136
Raven, age, gender	0.928	0.802	0.658	0.611	0.384	0.272
Raven, age, econ	0.913	0.989	0.123	0.529	0.397	0.076
Raven, gender, econ	0.650	0.725	0.250	0.953	0.503	0.084
Raven, age, gender, econ	0.845	0.948	0.393	0.667	0.902	0.193

3. Results: Hiroshima CU vs. Kansai U-B

Result 3: For subjects at Hiroshima CU and Kansai U-B, who differed in terms of region, Raven, gender, and econ, their behavior was **NOT** homogenized in any ways.

Table: p-values for z-test: Kansai U-B vs. Hiroshima CU, integrated data

Covariates	1 − 10	11 − 20	21 − 30	31 − 40	41 − 50	51 − 60
age	**0.035**	0.747	**0.013**	0.123	0.145	0.201
gender	0.116	0.414	**0.018**	0.067	0.062	0.131
age, gender	0.062	0.838	**0.024**	0.105	0.320	0.811
age, sci-eng	0.136	0.415	**0.004**	0.068	0.173	0.318
gender, sci-eng	0.208	0.440	0.056	0.133	0.158	0.344
age, gender, sci-eng	0.086	0.953	**0.025**	0.378	0.146	0.970
Raven	0.160	0.905	0.154	0.252	0.162	0.095
Raven, age	0.392	0.999	0.179	0.092	0.114	0.163
Raven, gender	0.100	0.947	0.142	0.200	0.113	0.339
Raven, age, gender	0.091	0.642	0.102	0.227	0.550	0.744
Raven, age, sci-eng	0.192	0.621	**0.015**	0.098	0.368	0.815
Raven, gender, sci-eng	0.087	0.664	0.126	0.059	0.400	0.276
Raven, age, gender, sci-eng	0.367	0.903	0.104	0.279	**0.030**	0.339

Table: p-values for z-test: Kansai U-B vs. Hiroshima CU

	Covariates	1 – 10	11 – 20	21 – 30	31 – 40	41 – 50	51 – 60
No-fb	age, gender, sci-eng	0.453	0.992	0.920	0.729	0.449	0.179
	Raven, age, gender, sci-eng	0.074	0.232	0.333	**0.031**	0.896	0.233
Part-fb	age, gender, sci-eng	0.566	0.304	0.197	0.079	**<0.001**	**0.016**
	Raven, age, gender, sci-eng	0.483	0.605	0.605	0.534	**0.002**	0.052
Full-fb	age, gender, sci-eng	0.481	0.905	**0.041**	0.752	0.474	0.625
	Raven, age, gender, sci-eng	0.208	0.754	0.060	0.481	0.529	0.098

Table: p-values for z-test: Kansai U-B (excl. econ) vs. Hiroshima CU

	Covariates	1 – 10	11 – 20	21 – 30	31 – 40	41 – 50	51 – 60
No-fb	age, gender, sci-eng	0.374	0.747	0.612	0.796	0.471	0.325
	Raven, age, gender, sci-eng	0.106	0.260	0.222	**0.040**	0.911	0.902
Part-fb	age, gender, sci-eng	0.782	0.682	0.103	0.251	**0.002**	0.141
	Raven, age, gender, sci-eng	0.135	0.943	0.600	0.777	**0.028**	0.421
Full-fb	age, gender, sci-eng	0.638	0.829	**0.004**	0.542	0.102	0.500
	Raven, age, gender, sci-eng	0.539	0.633	0.058	0.591	0.500	0.343

3. Results: Hiroshima CU vs.Doshisha

Result 4: For subjects at Hiroshima CU and Doshisha, who differed in terms of, region, Raven, gender, econ, sci-eng, their behavior was homogenized by Raven, econ, and sci-eng.

Table: p-values for z-test: Hiroshima CU vs. Doshisha, integrated data

Covariates	1 − 10	11 − 20	21 − 30	31 − 40	41 − 50	51 − 60
age	0.090	0.408	0.371	0.475	0.374	0.996
gender	0.103	0.281	0.459	0.465	0.113	0.998
age, gender	0.060	0.218	0.308	0.494	0.228	0.780
Raven	0.054	0.189	0.354	0.414	0.066	0.678
Raven, age	0.080	0.092	0.404	0.266	0.319	0.616
Raven, gender	0.082	0.173	0.463	0.424	0.106	0.694
Raven, age, gender	0.112	0.224	0.742	0.476	0.804	0.493

Table: p-values for z-test: Hiroshima CU vs. Doshisha

	Covariates	1 — 10	11 — 20	21 — 30	31 — 40	41 — 50	51 — 60
No-fb	age, gender	0.170	0.984	0.474	0.610	0.245	0.626
	Raven, age, gender	0.270	0.471	0.977	0.933	0.554	0.391
Part-fb	age, gender	0.474	0.058	0.599	0.169	**0.022**	0.183
	Raven, age, gender	0.652	**0.017**	0.333	0.097	**0.013**	0.857
Full-fb	age, gender	0.475	0.142	0.902	0.281	0.199	0.464
	Raven, age, gender	0.522	0.697	0.192	0.907	0.737	0.457

Table: p-values for z-test: Hiroshima CU (excl. sci-eng) vs. Doshisha (excl. econ)

	Covariates	1 — 10	11 — 20	21 — 30	31 — 40	41 — 50	51 — 60
No-fb	age, gender	0.573	0.626	0.620	0.981	0.100	0.369
	Raven, age, gender	0.116	0.645	0.363	0.323	0.944	0.242
Part-fb	age, gender	0.416	**0.003**	0.470	0.101	0.195	0.483
	Raven, age, gender	0.415	**0.009**	0.547	0.231	0.152	0.560
Full-fb	age, gender	0.971	0.240	0.797	0.380	0.999	0.854
	Raven, age, gender	0.138	0.518	0.263	0.551	0.680	0.677

4. Concluding Remarks

(1) Today's talk

- **Hypothesis**: No significant difference in subjects' choice at different experimental sites, after we control for Raven in addition to age, gender, econ., and sci-eng.
- **Observation**: The hypothesis was affirmatively verified among schools located in the same region.

(2) The remaining half of this paper: Meaningful learning was not observed, whereas it was observed at Osaka University and the University of Tsukuba.

- higher Raven scores and longer thinking time to choose $\Rightarrow$ there is another latent factor to homogenize subject's behavior

Appendix: Propensity Score Matching

Definition: Denote by $\boldsymbol{x}_i$ the vector of covariates for subject i and by z_i the variable that stands for assignment to an experimental site. The probability of subject i with his or her covariate vector $\boldsymbol{x}_i$ being assigned to site 1, $e_i = p(z_i = 1|\boldsymbol{x}_i)$, is called the propensity score for subject i.

In practice, many researchers estimate the propensity scores of subjects by using binomial logistic regression:

$$e_i = \frac{1}{1 + exp(-(b_0 + b_1 x_{i1} + \cdots + b_p x_{ip}))}.$$

Let subject i at site 1 be paired with the subject j at site 2 who has the closest propensity score to his or her propensity score. Their cognitive ability scores and attribute information do not differ greatly, because e_i and e_j have similar values.

The paired subjects are thus considered to be assigned randomly to one of the two experimental sites in a pseudo manner. Consider e as a random variable. By using the distribution of e, we have

$$E(y_1 - y_2) = E_e E(y_1|z = 1, e) - E_e E(y|z = 2, e),$$

where z denotes the assignment variable.