ソシオネットワーク戦略ディスカッションペーパーシリーズ ISSN 2434-9445
第 130 号　2026 年 4 月
RISS Discussion Paper Series
No.130　April, 2026

Meaning and labor supply: The case of the white jigsaw puzzle

Keiko Yoshida, Katsuhiko Nishizaki, Hiroyasu Yoneda

文部科学大臣認定　共同利用・共同研究拠点

関西大学ソシオネットワーク戦略研究機構

Research Institute for Socionetwork Strategies,
Kansai University
Joint Usage / Research Center, MEXT, Japan
Suita, Osaka, 564-8680, Japan
URL: https://www.kansai-u.ac.jp/riss/index.html
e-mail: riss@ml.kandai.jp
tel. 06-6368-1228
fax. 06-6330-3304

Meaning and labor supply: The case of the white jigsaw puzzle

Keiko Yoshida*†
Katsuhiko Nishizaki*
Hiroyasu Yoneda*

March 28, 2026

Abstract

In Japan, where the birthrate is declining, improving labor productivity is a crucial issue. Since workers' sense of meaning in their work is thought to impact productivity significantly, we focused on its effect. Work meaning can be classified into two categories: the social importance of work and the sense of accomplishment derived from the results themselves. We focused on the latter effect. The purpose of this study is to investigate how minimal perceived meaning influences labor supply and conduct the experiment in Japan, following the methodology of Ariely et al. (2008). In a laboratory environment, the perceived meaning of repetitive tasks was manipulated. In each of the two conditions, participants received a reward of progressively decreasing unit wages for completing a 40-piece white puzzle. Contrary to previous studies, this study's results did not show that perceived meaning influenced labor supply. Possible reasons include participants' young working-age status, which led them to prioritize monetary rewards, and the unique characteristics of the puzzle as a task.

JEL Classification: C91, D91, J22
Keywords: meaningful work, motivation, behavioral economics, experiment

---

† Corresponding author. Email: kyoshida@andrew.ac.jp
＊Momoyama Gakuin University, Faculty of Economics.

1. Introduction

The bubble economy collapsed at the end of the Showa era, and the Heisei recession lasted for 30 years. As a result, there are significant differences in values, especially work ethics, between the Showa and Heisei generations. Hoffman (2022) states that a generational gap exists in the workplace between the Showa and Heisei eras.

The Showa era, born after World War II, built a rapidly growing economy and the accompanying mindset of its people. On the other hand, the younger Heisei generation is considered to have a weaker sense of belonging to the workplace and to be more individualistic than the Showa generation. With the declining birthrate, newly hired employees, whom companies have worked hard to recruit, are leaving. It is important to examine people's work motivations, especially those in the relatively younger generation.

In recent years, a growing literature argues that workers do not only care about their wages but also about whether their jobs are meaningful. Several studies find positive effects of work meaning on workplace productivity. Nikolova and Cnossen (2020) demonstrated why meaningful work matters to labor economists and built on self-determination theory, which specifies the roles of autonomy, competence, and relatedness as preconditions for motivation. Using the European Working Conditions Survey, they found that autonomy, competence, and relatedness explain about 60% of the variation in work meaningfulness perceptions.

Carpenter and Gong (2016) conducted a real-effort experiment with participants who work directly for organizations with clear missions, and they surveyed potential participants for their organizational preferences weeks before the experiment. They show that matched workers produce 72% more than mismatched workers, and that performance pay can increase output by 35% compared to workers who receive only a base wage.

Grant (2008) implemented three field experiments that examined the performance effects, relational mechanisms, and boundary conditions of task significance. His results showed convergent support for the causal effects of task significance on job performance and provided novel insights into the relational mechanisms and boundary conditions for these effects.

Chadi et al. (2017) investigated a causal relationship between the meaning of work and employees' motivation to perform well. They observed a strong decline in exerted effort when they informed workers about the meaninglessness of a job already done.

Kosfeld et al. (2017) manipulated workers' perceived meaning of a job in a field experiment and interacted the meaning of work with both financial and recognition incentives, concluding that workers exert more effort when the meaning is high.

Many studies focus on the impact of the social importance of labor on labor supply, but there are few analyses of the impact of visualized outcomes on labor supply. Ariely et al (2008) focused on the impact of visualized outcomes. We investigate the meaning and labor supply

with an experiment following Ariely et al (2008). Unlike the research, we could not find the positive effect of perceived meaning on labor supply. Possible reasons include participants' young working-age status, which led them to prioritize monetary rewards, as well as the unique characteristics of the puzzle task.

The remainder of this paper is organized as follows. Chapter 2 explains the experimental design we implemented, and Chapter 3 describes the analysis results. Chapter 4 discusses the analysis results, and Chapter 5 summarizes the conclusions.

2. Experimental design

To examine how meaning affects the relationship between productivity and labor supply, we conducted two types of experiments with young workers as subjects on October 17, 2023, and on October 29 and 30, 2024.[2] All the subjects were male and female workers in their 20s and 30s who appeared to work in Umeda area Osaka, Japan, recruited via flyers distributed within a 1.3-kilometer radius of Umeda, excluding commercial facilities and the subject pool managed by the Research Institute for Socionetwork Strategies, Kansai University.[3] Each subject participated in only one session.[4] To make it easier to participate our experiments after work, they were conducted at Kansai University's Umeda campus.

Following the experimental methodology of Ariely et al. (2008), we conducted two types of experiments: one experiment followed the meaningful condition, while the other followed the Sisyphus condition. Although Ariely et al. (2008) used Bionicle Lego models, we used 40-piece plain white jigsaw puzzles. Under each condition, each subject continued assembling the puzzles as long as the subject wanted. Under the meaningful condition, each subject could track how many puzzles the subject have assembled by sticky notes provided by the experimenter on the partition in front of the subject, whereas under the Sisyphus condition, no such sticky notes were provided.[5] The advantages of using a plain white jigsaw puzzle are that they are less likely to cause differences in preference between men and women, and that

---

[2] The experiments presented in this paper were approved by the Institutional Review Board of the Research Institute for Socionetwork Strategies, Kansai University (Nos. 2023018 and 2024016).

[3] Umeda is one of Japan's leading business districts.

[4] This implies that the data between sessions are mutually independent.

[5] In Ariely et al.'s (2008) experiment, the assembled Bionicles were lined up on the desk in front of the subject under the meaningful condition, whereas they were disassembled by the experimenter.

assembling them can be completed even in a small partitioned area.[6]

All the subjects were randomly divided into two groups, Meaningful (N =28) and Sisyphus (N = 32), and gathered in the same room under identical conditions.[7] Each subject was assigned an identification number randomly and was seated at a desk labeled the same number in the room. Each desk was separated by visual partitions which ensured that each subject could assemble the puzzles anonymously and independently.[8] After receiving an explanation of the experiment, the subjects took a five-question quiz to assess their understanding.[9] After marking the quiz for each subject, the experimenter explained the correct answers to all the subjects to deepen their understanding of the content of the experiment.[10]

Subjects received rewards for completing a puzzle according to a declining unit-wage schedule. Before deciding whether to build each puzzle, each subject was told how much the subject had earned up to that point and how much the subject would earn for completing another puzzle. Each subject was paid 1000 JPY (roughly 6.7 USD at the exchange rate at the time the experiment was conducted) for the first puzzle, 900 JPY (100 JPY less) for the second, and so on, linearly. For the 10th and any subsequent puzzles, the subject received 100 JPY. The only decision each subject made was when to stop assembling another puzzle. At that point, the subject earned the sum of rewards according to the number of completed puzzles, and the experimental session was over. The average was approximately 5803.77 JPY (roughly 38.69 USD). The experimental session lasted approximately 2.25 h, including the explanation of the experiment and the payment of the reward. The average hourly earnings (approximately 2579.45 JPY) exceeded the average hourly wage for a typical young worker working in Umeda at the time of the experiment.

---

[6] In Japan, it is mainly boys who play with Legos.

[7] Subjects were unaware of the other group.

[8] Although each subject could not see other subject assembling the puzzle, the subject did see other subject finish the task and leave the room. As stated below, note that other subjects leaving the room did not affect the subject's amount of task.

[9] These questions were multiple-choice questions about the tasks the subjects would perform and the rewards they would receive after the experiment. See Appendix for the quiz. Appendix A is the test distributed to subjects, while Appendix B provides an English translation of Appendix A for reference.

[10] The average score was 4.47 out of 5 points. Some subjects missed Question 3 or 5.

3. Results[11]

To conduct a more detailed analysis, data from participants aged 40 or older and those who answered multiple questions incorrectly on the comprehension check test were excluded. Figures 1-3 show the results from each experiment. 1 represents the results under the meaningful condition, and 0 represents the results under the Sisyphus condition. Figure 1 shows the results from the 2023 experiment, with an average number of 15.08 under the Sisyphus condition, and 12.9 under the meaningful condition. Figure 2 shows the results obtained on the first day of one of the two experiments conducted in 2024. The average score was 15.11under the Sisyphus condition and 12.6 under the meaningful condition. Figure 3 shows the results obtained on the second day of one of the two experiments conducted in 2024. The average score was 10.67 under the Sisyphus condition and 14.25 under the meaningful condition. Figure 4 graphs the data from all experiments. The average score was 13.76 under the Sisyphus condition and 13.3 under the meaningful condition.

Table 1 shows the descriptive statistics of the data. "Task" refers to the number of white puzzles completed by the subjects, with an average of 13.6, a minimum of 5, and a maximum of 25. "Meaning" is a dummy variable assigned a value of 1 to subjects who participated in the experiment under a meaningful condition. "Second" is the number of seconds it took each subject to complete the first puzzle. The average was approximately 11 minutes and 55 seconds, the minimum was 5 minutes and 35 seconds, and the maximum was 22 minutes and 41 seconds. "Second" is thought to represent an individual's preference for puzzles or their aptitude for the task of assembling them, because individuals who like puzzles or are good at assembling them may take less time to solve the first puzzle. The average age of the subjects was 29.8 years. "r1" is a dummy variable assigned a value of 1 to subjects who participated in the experiment on October 29, 2024, and "r2" is a dummy variable assigned a value of 1 to subjects who participated in the experiment on October 30, 2024.

"Selfemployed" is a dummy variable assigned a value of 1 to those who answered that they were self-employed in the pre-experiment questionnaire; 3.8% of the participants were self-employed. "Nonregular" refers to contract employees or part-time workers; 9.4% of the participants were non-regular employees. "Manager" is a dummy variable assigned a value of 1 to those who answered that they were managers or company executives in the pre-experiment questionnaire; 5.7% of the participants fell into this category.

Table 2 shows the results of the t-test. The t-value was 0.3361, and we did not find a statistically significant difference in the mean number of tasks between the meaningful

[11] Further inquiries on the date presented in this paper can be directed to the corresponding author. For the analysis, we used STATA 16.

condition and the Sisyphus condition. Therefore, no impact of the visualized results on labor supply was observed.

Table 3 shows the results of the regression analysis. The results show that those who completed the first puzzle faster ultimately solved more puzzles. The coefficient values indicate that a difference of 100 seconds corresponds to a difference of approximately one puzzle. Furthermore, on average, non-regular employees completed four more puzzles than regular employees. This suggests that non-regular employees, due to their relatively lower income, may have needed more compensation, and that their shorter working hours may have allowed them to participate in the experiment without accumulating fatigue.

Table 4 shows the results of the analysis using propensity score matching. Propensity score matching estimation is a statistical method used in observational data to equalize the background characteristics of groups that received treatment or intervention and those that did not, allowing for fair comparison. Each participant's "probability of receiving intervention" is represented by a propensity score between 0 and 1, and individuals with similar scores are paired for analysis. This method is applied to non-randomized observational data; although our experiment followed a randomization procedure, we considered the possibility of some heterogeneity in the data. ATE (1) refers to an analysis that uses only the time taken to complete the first puzzle to calculate the propensity score, while ATE (2) refers to an analysis that includes time, age, experiment participation date, and working conditions in calculating the propensity score. In both results, no impact from the visualized outcomes was observed.

4. Discussion

Unlike the results obtained by Ariely et al. (2008), we couldn't observe any effect of perceived meaningfulness on labor supply. Several factors can be considered for these results. Firstly, the subjects are relatively young workers in Japan. Henrich et al. (2010) point out that many behavioral science analyses use university students in Western countries as subjects. In contrast, the subjects in our experiment were workers in their 20s and 30s. Since work experience can significantly change people's views on work, the results of our experiment may differ from those of experiments using university students. These differences may have contributed to the differences in our results. Furthermore, it is possible that the work values of the younger generation differ from those of other generations. A public opinion survey on national life conducted by the Cabinet Office in August 2025 included a question about the purpose of work, with 81.3% of 19-29-year-olds responding that they work to earn money. The overall average for this question was 63.5%, suggesting that younger generations prioritize monetary rewards and place less emphasis on psychological rewards such as job satisfaction and meaning in their work. Furthermore, according to a "Survey on

Intergenerational Gaps" conducted by SMBC Consumer Finance Co., Ltd. in 2025, 19.6% of younger generations selected "job satisfaction" as what they seek in work, compared to 33.0% of older generations. This survey also suggests differences in the importance of job satisfaction across generations. It is possible that these values of younger generations influenced the results.

Secondly, we must consider whether white puzzles played an appropriate role. We adopted the white puzzle as an appropriate task because it was accepted as the same task regardless of gender or preference. However, even with a plain white puzzle, the task can have different meanings for individuals who enjoy solving puzzles and those who do not. Data analysis included the time taken to assemble the first puzzle, which revealed that this time coefficient was statistically significant. In Ariely et al.'s (2008) experiment, the task was to assemble a robot. For the subjects, seeing their assembled robot taken apart might be emotionally painful, but seeing a white puzzle taken apart might not be as emotionally painful. Also, the situation of having their assembled robots lined up in front of them might have different meanings than receiving colorful sticky notes each time they complete a puzzle.

Thirdly, in the experiment by Ariely et al. (2008), each subject participated alone, without other subjects present. In our experiment, subjects participated alongside other participants in the same group, albeit separated by partitions. Therefore, the presence of other participants may influence the participants' behavior. Regarding this possibility, we asked participants to complete an online questionnaire after the experiment. This questionnaire asked about when they stopped the task, and many chose options such as "because I got bored with assembling the puzzle" or "because I felt it was not worth it compared to the income." Therefore, the presence of other participants may not have a significant influence.

5. Conclusion

This study examined, through economic experiments, whether visualized outcomes improve people's labor supply. Regression analysis suggests that while visualization of achievement is unrelated to the amount of work completed, whether participants found the task enjoyable or had an aptitude for it may improve task completion. Furthermore, on average, non-regular employees completed 4 more puzzles than regular employees. Our results showed that visualized outcomes did not stimulate participants' motivation to work.

The reasons for the results are as follows. The participants were young workers who may have prioritized monetary rewards. Alternatively, the unique nature of the puzzle task, the presence of assembled robots, and the provision of colorful sticky notes for completing each puzzle may have had different meanings, meaning the sticky notes may not have influenced the participants' behavior.

Future challenges include conducting experiments with subjects aged 40 and older and conducting analyses that consider the time taken to solve the second and subsequent puzzles. While this study targeted workers under 40, comparing the results with those of subjects aged 40 and older would allow for a more comprehensive discussion. Such comparisons are highly significant because there are concerns that differences in values between generations are causing problems in the workplace in Japan.

Acknowledgements

The authors thank Yuko Arayama, Yasuharu Ukai, Naoki Watanabe, and workshop participants at Kansai University for helpful suggestions. The authors also thank Kazuhito Ogawa and the staff of RISS and Kansai University's Umeda Campus for supporting the experiments presented in this paper. This research was supported by the Joint Usage and Research Center, Research Institute for Socionetwork Strategies, Kansai University.

References

Ariely, Dan, Emir Kamenica, & Dražen Prelec (2008). "Man's search for meaning: The case of Legos," Journal of Economic Behavior and Organization, 67(3–4), 671–677.

Cabinet Office (2025). "The Public Opinion Survey on the Life of the People"

Carpenter, J., & Erick Gong (2016). "Motivating agents: How much does the mission matter?," Journal of Labor Economics, 34(1), 211–236.

Chadi, Adrian, Sabrina Jeworrek, & Vanessa Mertins (2017). "When the meaning of work has disappeared – Experimental evidence on employees' performance and emotions," Management Science, 63(6), 1696–1707.

Grant, A. (2008). "The significance of task significance: Job performance effects, relational mechanisms, and boundary conditions," Journal of Applied Psychology, 93(1), 108–124.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). "The weirdest people in the world?" Behavioral and brain sciences, 33(2-3), 61–83.

Hoffman, Michael. (2022, June 23). Generation gap at work in the office between Showa and Heisei employees. Japan Today. https://japantoday.com/category/features/kuchikomi/generation-gap-at-work-in-the-office-between-showa-and-heisei-employees (Accessed March 31, 2026)

Kosfeld, M., Susanne Neckermann, & Xiaolan Yang (2017). "The effects of financial and recognition incentives across work contexts: The role of meaning," Economic Inquiry, 55(1), 237–247.

Nikolova, M., & Femke Cnossen. (2020). "What makes work meaningful and why economists should care about it." Labour Economics 65: 101847.

SMBC Consumer Finance Co., Ltd. (2025). "Survey on Intergenerational Gaps"

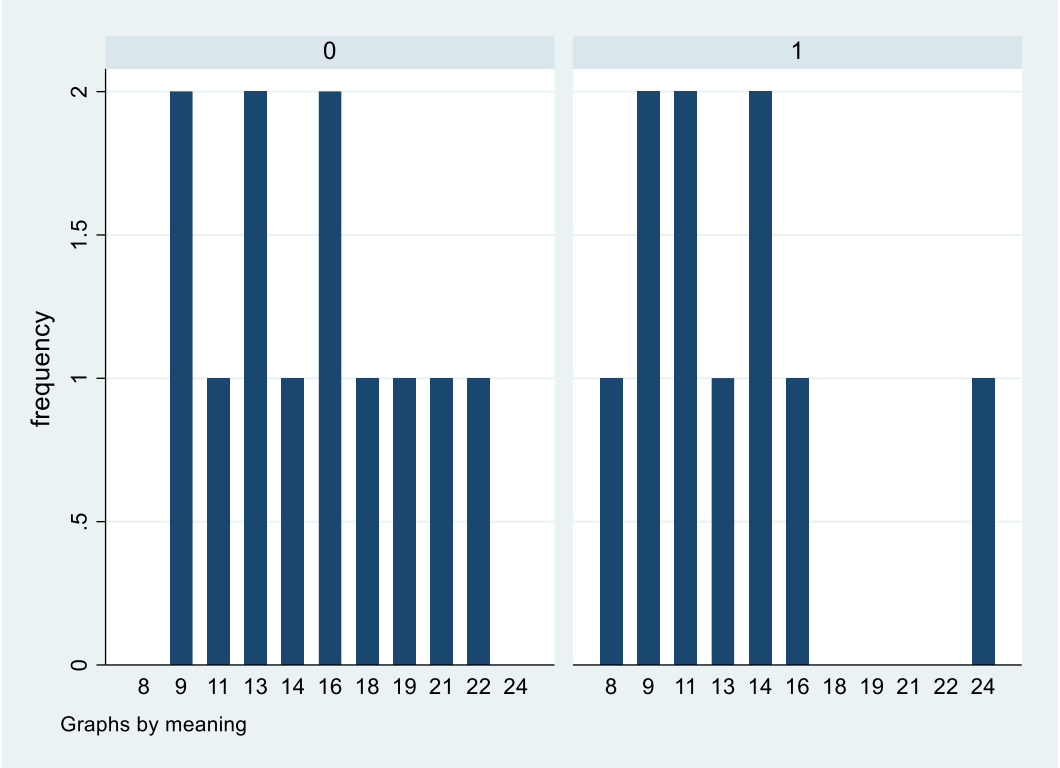

Figure 1. Number of white puzzles completed in the experiment on October 17, 2023.

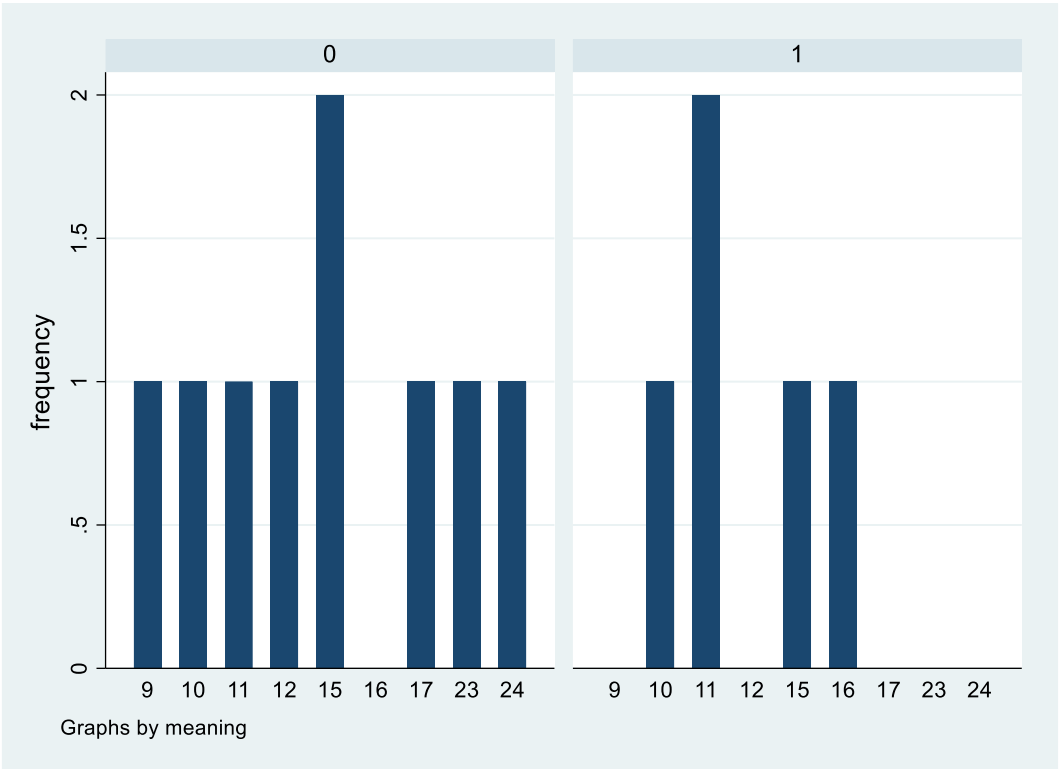

Figure 2. Number of white puzzles completed in the experiment on October 29, 2024.

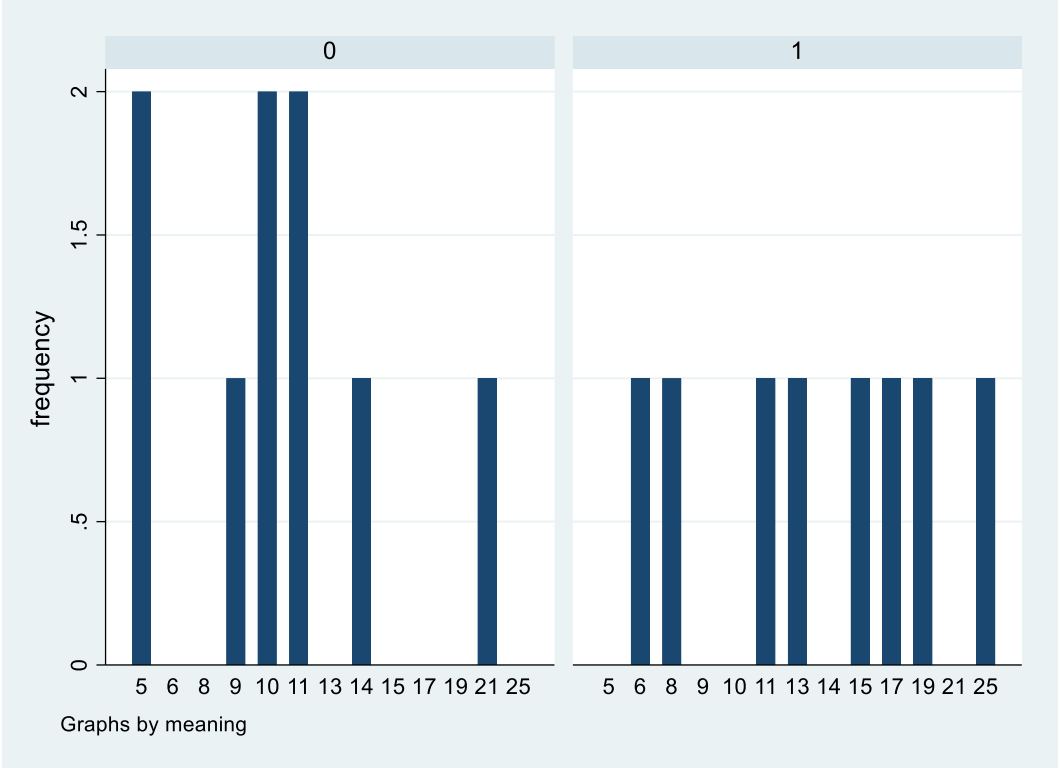

Figure 3. Number of white puzzles completed in the experiment on October 30, 2024.

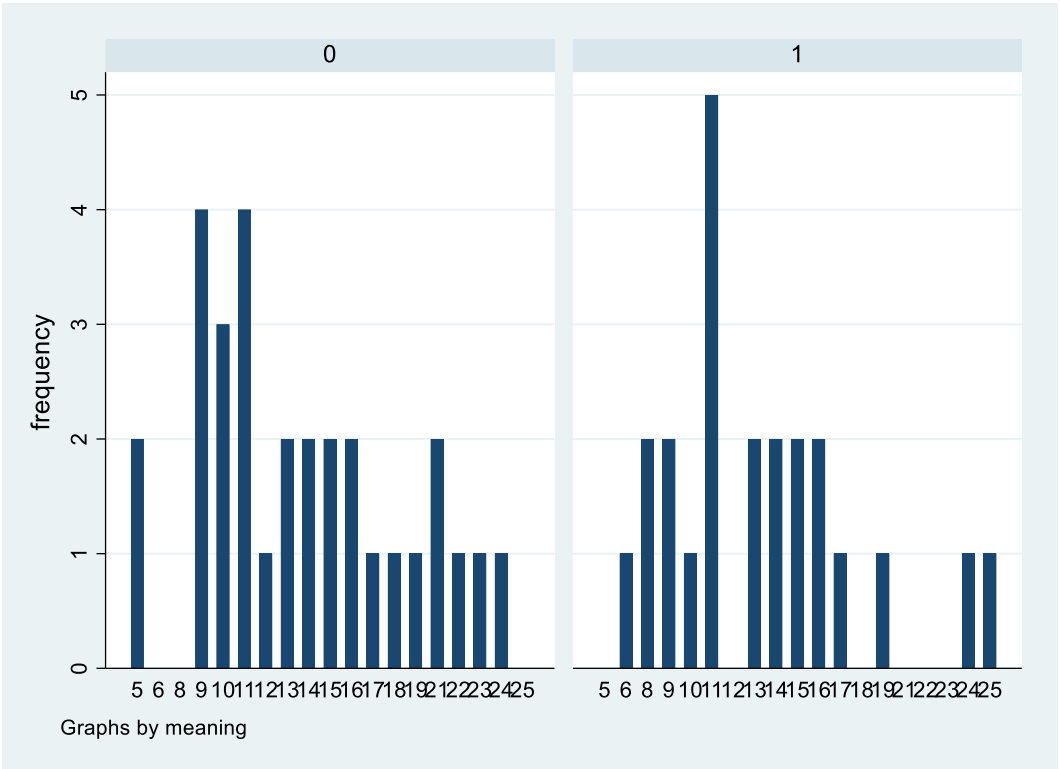

Figure 4. Number of white puzzles completed

Table 1. Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
task	53	13.566	4.921	5	25
meaning	53	0.434	0.500	0	1
second	53	715.434	283.987	335	1361
age	53	29.755	4.247	21	39
r1	53	0.264	0.445	0	1
r2	53	0.321	0.471	0	1
selfemployed	53	0.038	0.192	0	1
nonregular	53	0.094	0.295	0	1
maneger	53	0.057	0.233	0	1

Table 2. t-test

Group	Obs	Mean	Std. Err.	Std. Dev.
Sisyphus	30	13.767	0.931	5.097
Meaningful	23	13.304	0.997	4.781
combined	53	13.566	0.676	4.921
			t = 0.3361	

Table 3. Regression analysis

	(1) task	(2) task	(3) task	(4) task
meaning	-0.462	-0.150	-0.272	-0.198
	(-0.34)	(-0.14)	(-0.25)	(-0.19)
second		-0.0112***	-0.0121***	-0.0121***
		(-6.00)	(-6.14)	(-6.01)
age			0.158	0.191
			(1.18)	(1.44)
r1			1.148	0.945
			(0.88)	(0.71)
r2			-1.224	-1.768
			(-0.99)	(-1.45)
selfemployed				1.018
				(0.34)
nonregular				4.031*
				(2.28)
maneger				-1.553
				(-0.68)
_cons	13.77***	21.66***	17.69***	16.63***
	(15.19)	(14.56)	(4.71)	(4.42)
N	53	53	53	53

t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table 4. Propensity score matching

task	Coef.	Std. Err.	z	P>|z|
ATE (1)				
meaning (1 vs 0)	1.170	0.910	1.29	0.199
ATE (2)				
meaning (1 vs 0)	-0.226	1.123	-0.2	0.84

Appendix A: Comprehension Test (Original Version)

ID 番号＿＿＿＿＿＿＿＿

2023 年 10 月 17 日

実験理解度確認テスト

●このテストは実験の内容を理解しているかを測定するものです。

●解答の際は□の中にチェックをしてください。

●ID 番号を右上に記入してください。

●制限時間は 3 分です。

問 1．この実験で被験者が行うことを選んでください。

□40 ピースの無地のパズル

□レゴブロックで作るロボット

□千羽鶴

問 2．パズルを 2 枚完成させた場合の報酬合計額を選んでください。

□1000 円　□1500 円　□1900 円

問 3．パズルの報酬の 1 枚ごとの減少額を選んでください。

□200 円　□100 円　□50 円

問 4．15 枚目のパズルを完成させた場合の報酬額を選んでください。

□0 円　□50 円　□100 円

問 5．パズルを 2 枚完成させた場合と 5 枚完成させた場合の報酬合計額を比べて、
1 枚当たりの報酬額が高い場合を選んでください。

□2 枚完成させた場合　□5 枚完成させた場合

Appendix B: Comprehension Test (English Version)

ID. _____

October 17, 2023

Experiment Comprehension Test

This test measures your understanding of the experiment.

Please check the box in the appropriate place when answering.

Please write your ID number in the upper right corner.

Time limit: 3 minutes.

Question 1. Choose what the subject will do in this experiment.

☐ A 40-piece blank puzzle

☐ A robot made from LEGO blocks

☐ A thousand paper cranes

Question 2. Choose the total reward amount for completing two puzzles.

☐ 1000 yen ☐ 1500 yen ☐ 1900 yen

Question 3. Choose the amount by which the reward decreases with each puzzle completed.

☐ 200 yen ☐ 100 yen ☐ 50 yen

Question 4. Choose the reward amount for completing the 15th puzzle.

☐ 0 yen ☐ 50 yen ☐ 100 yen

Question 5. Compare the total reward amount for completing 2 puzzles versus completing 5 puzzles, and choose the option with the higher reward per puzzle.

☐ Completing 2 puzzles ☐ Completing 5 puzzles