https://doi.org/10.53656/phil2024-03-07

Philosophy of Medicine
Философия на медицината

EXPLORING THE NEXUS OF PSYCHOLOGICAL SAFETY AND PHYSICAL HEALTH IN THE WORKPLACE: A MACHINE LEARNING AUGMENTED STUDY

Dr. Satoshi Maeda
Chuo University (Japan)

Abstract. This paper investigates the connection between psychological safety at work and physical health outcomes. Employing data from roughly a thousand respondents in Tokyo in 2021 and utilizing machine learning techniques along with traditional statistics, the study reveals that this popular concept in the field of business management and organization behavior, "psychological safety" at work, can enhance physical health of workers under certain conditions. The finding that its effect remains even when mental stress levels are controlled along with basic ascriptive variables, extends the conventional notion around "psychological safety." The recently developed Causal Forest Double Machine Learning (Causal Forest DML) analysis was used to generate a decision tree, shedding light on the structure of causal inference and indicating the key role of "age" and "personal income" as determinants of the effect of psychological safety on physical health.

Keywords: Psychological safety; workplace stress; Tokyo, Causal Forest DML; machine learning; physical health

1. Introduction

In today's workplace, understanding the effects of workplace dynamics on employee well-being is increasingly important. Psychological safety, which allows employees to share ideas and concerns without fear of negative consequences, is crucial for building a culture that supports innovation, resilience, and teamwork. Beyond these organizational benefits, psychological safety also plays a critical role in an area that has not been as widely discussed: the physical health of employees.

The research uses both traditional statistical methods and advanced machine learning techniques, specifically Ordinal Logistic Regression and Causal Forest Double Machine Learning (Causal Forest DML), analyzing data from nearly a thousand employees. The goal is to provide new insights into how psychological safety relates to physical health and suggest that implementing measures to enhance psychological safety may be an effective way to improve physical health of workers.

2. Literature Review

The concept of psychological safety, first examined by Schein and Bennis in the 1960s, has evolved significantly, becoming central to organizational behavior research. It refers to the perceived freedom to take risks and express oneself without fear of negative consequences at work, and is considered crucial for fostering innovation, collaboration, and resilience among employees.

2.1. Evolution of Psychological Safety

William A. Khan expanded the application of psychological safety in 1990, linking it to employee engagement—a key concept in management and organizational behavior. Amy Edmondson further advanced this area in 1999 with her empirical research, showing that psychological safety enables teams to engage in open dialogue, admit mistakes, and learn, thus enhancing team performance and learning.

2.2. Psychological Safety and Health Outcomes

Although the connection between psychological safety and organizational performance is well-documented, its impact on employee health has received less attention. Studies such as those by Newman et al. (2017) began to explore how psychological safety may reduce work-related stress and improve both physical and mental health.

2.3. Addressing the Gap

This study seeks to examine the unique effect of psychological safety on physical health outcomes in Tokyo's distinctive work culture. Contemporary scholarly research consistently underscores the significance of "psychological safety," highlighting its crucial role in enhancing team performance and mental well-being. While discussions often include how psychological safety might reduce workplace injuries (Edmondson, 2018), its direct link to physical health is not as thoroughly examined.

The impact of psychological safety on employees' physical health presents an intriguing area of study. Common assumptions may not be challenged if psychological safety is thought to benefit physical health through the reduction of mental stress. However, a deeper question emerges: Does psychological safety maintain its beneficial effects on physical health even when mental stress levels are controlled? Discovering that psychological safety contributes independently to better physical health outcomes, separate from mental health, would pave the way for further extensive research.

This nuanced relationship between psychological safety and physical health is the focus of this paper. This insight, despite certain limitations, begins to reveal the potential independent effects of psychological safety on physical health. They suggest that there might be unexplored avenues in the complex interaction between mental and physical health, as well as the impact of psychosocial factors on physical well-being.

3. Methodology

We collected data from nearly a thousand employees in Tokyo and its surrounding areas in 2021. The design and analytical methods of the study were carefully selected to effectively examine the causal relationship between psychological safety and health outcomes, employing both traditional statistical methods and advanced machine learning techniques.

3.1. Data Collection

We compiled our dataset through a cross-sectional survey, capturing a comprehensive range of variables. These included personal income, age, years of education, employment status, perceived psychological safety, and self-reported physical health. Each variable was chosen for its relevance to the central hypothesis of this study: that psychological safety in the workplace has an influence on the physical health of employees.

3.2. Dataset

From the quantitative survey "5th Survey on Consumption and Lifestyle" conducted by the Global Consumer Culture Research Group in Tokyo and Nagoya in October-November 2021.

Survey Subjects: Individuals aged 20-69 living in 1) the Tokyo area (within a 40km radius of Shinjuku Station), 2) the Nagoya area (within a 25km radius of Nagoya Station).

Sampling Method: Stratified two-stage random sampling using the Basic Resident Register.

Survey Method: Questionnaire survey conducted by mail.

Planned Samples: 4,500 cases (Tokyo area: 3,300, Nagoya area: 1,200).

Valid Responses: 1,237 in the Tokyo area (response rate 37.5%), 444 in the Nagoya area (response rate 37%).

In this particular report, only the data collected from Tokyo and its surrounding areas are used.

3.3. Variables Used for the Analysis

The following demographic and attitude survey items were used in the analysis to represent each variable.

3.3.1. Demographic Survey Items

– Personal Income: Annual personal income before tax deduction in JPY.

– Age: 20 – 69.

– Educational Year: 9 – 18 years of education.

– Male Dummy: 1 for male and 0 for female.

– Employment Status Categories (8 categories): Included only in Causal Forest DML analysis.

– Full-time housewife or househusband (not working part-time or as a family business employee).

– Student (including those working part-time).

– Unemployed (including those who have retired).

– Part-time worker, temporary employee.

– Contract employee, outsourced worker, freelance worker.

– Full-time permanent employee (“seiki-koyou” is the word used in the original questionnaire).

– Self-employed or family business employee.

– Manager, corporate executive, organizational executive.

– Other (specifically:) *Not included in this analysis.

Due to the selection of respondents based on the above criteria, the final valid sample size was reduced to 912.

3.3.2. Attitude Survey Items

– Work-Related Negative Stress: 1 to 4 (4 being the highest level of stress).

– “I’m stressed out from work and job-related issues.”

– Choices: 1. Does not apply, 2. Does not really apply, 3. Somewhat applies, 4. Applies.

– Psychological Safety[1]: 1 to 4 (4 being the highest level of safety).

– “In my current workplace, I feel comfortable reporting or discussing mistakes I’ve made with the people around me.”

– Choices: 1. Does not apply, 2. Does not really apply, 3. Somewhat applies, 4. Applies.

– Self-Reported Physical Health: 1 to 4 (4 being the highest level of physical health).

– For Ordinal Logistic Regression: 3 degrees.

– For Causal Forest DML: 4 degrees.

– “I’m healthy”

– Choices: 1. Does not apply, 2. Does not really apply, 3. Somewhat applies, 4. Applies.

The distribution of the above attributes can be found in the Appendix B at the end of this paper.

3.4. Analytical Methods

To analyze the data, two distinct approaches were employed: Ordinal Logistic Regression and Causal Forest Double Machine Learning. These methods were selected for their complementary insights into the complex nature of the research question. These analytical methods facilitated a thorough examination of the data, identifying significant associations and exploring complex dynamics within the scope of the study, while acknowledging the limitations imposed by the dataset in making definite causal inferences.[2]

3.4.1. Ordinal Logistic Regression (OLR)

This well-established method was used to investigate the associations between various predictors and an ordinal dependent variable that categorizes physical health outcomes into three levels (this variable was originally measured with 4-point scale and later recoded into 3-point scale for OLR). The analysis was performed with careful attention to meeting the assumptions required for ordinal logistic regression, such as the proportional odds assumption. This approach aimed to identify which predictors are statistically significant in relation to the self-reported physical health outcomes.

3.4.2. Causal Forest Double Machine Learning (Causal Forest DML)

Following the associative insights provided by the established method of OLR, a recently developed[3] approach, Causal Forest DML, was employed to further explore the potential impacts of psychological safety on physical health, allowing for proper control of covariates while performing causal inference. This method incorporated "employment status" as an additional independent variable, which was not feasible in OLR for reasons explained later. Additionally, "self-reported physical health" was kept as 4-point scale, unlike in the OLR. This technique employs machine learning to assess the possibility of varying effects of psychological safety across different workforce segments.

In the first step of Causal Forest DML, random forest machine learning algorithm is used to estimate models for both the treatment (in this case, "psychological safety at work") and the outcome ("self-reported physical health"). Random forests are ensembles of decision trees, where each tree is trained on a random subset of the data and features. This ensemble approach helps to reduce overfitting and improve prediction accuracy by averaging the results of multiple trees. Decision trees themselves are simple models that split the data based on feature values to predict an outcome.

In the second step, the outputs from these models are used to calculate the residuals for the treatment and the outcome. This means taking the actual treatment and outcome values for each unit and subtracting the predicted values from them. By doing this, we remove the part of the variation in the outcome and treatment that can be explained by the covariates (in this study, "age," "personal income," "educational attainment," "employment status," "gender," "work-related negative stress"), isolating the effect of the treatment itself.

Finally, regression analysis is performed on these residuals to estimate the treatment effect (CATE: Conditional Average Treatment Effect). The goal of this step is to isolate the impact of the treatment from the effects of other covariates. This involves regressing the residualized outcome on the residualized treatment, allowing for an unbiased estimation of the treatment effect.

In general, machine learning approaches, including Causal Forest DML, are not commonly used in sociology. One reason, in my opinion, is that many conventional machine learning methods focus on prediction, and their models, especially neural networks, are often black boxes, making them difficult to interpret. Sociology emphasizes model visualization and explanation, which these methods typically do not provide.

However, Causal Forest DML differs from conventional machine learning. It allows for partial visualization of analysis results, such as the illustration of average decision tree and the presentation of effect sizes and confidence intervals for intervention variables on outcomes. These features make it more suitable for sociological analysis by offering clearer insights into causal relationships and enhancing interpretability.[4]

Combining Causal Forest DML with techniques like OLR can further validate findings and make results more communicable to a broader audience. This dual approach can leverage the strengths of both methods, offering comprehensive and interpretable insights into sociological phenomena.

3.4.3. Data Integrity and Bias Consideration

Throughout the data collection and analysis phases, rigorous steps were taken to ensure data integrity and address potential biases. These steps included careful survey design, validation of survey responses through randomized sampling, and consideration of confounding variables in the analytical models. The combination of OLR and Causal Forest DML allowed for an analysis that mitigates methodological biases and provides reliable insights into the research questions of the study.

4. Results

4.1. Bivariate Analyses

Before presenting the applications of the aforementioned complex multivariate analyses, it is useful to display the simple bivariate relationships of the key variables as follows. First, a correlation matrix showing Spearman's correlation coefficients. Then, scatter plots visualizing the trends between key variables.

4.1.1. Correlation Matrix

The table below presents Spearman's correlation coefficients among various key variables. "Self-reported physical health" exhibits a minor relationship with "psychological safety" ($p<.01$), after the more obviously related variable, "age" ($p<.01$). The remaining question is whether the relationship remains strong enough after controlling for other variables, which was confirmed in subsequent multivariate analyses.

Table 1

	Spearman's	Self-Reported Physical Health	Psychological Safety	Work-Related Negative Stress	Age	p_income /100	Educational Year	Male_Dummy
Self-Reported Physical Health	Correlation	1.000	.149**	-.119**	-.190**	.067*	.087**	-0.031
	Sig. (2-tailed)		0.000	0.000	0.000	0.021	0.002	0.284
	N	1225	936	936	1225	1192	1218	1225
Psychological Safety	Correlation	.149**	1.000	-.227**	-.115**	-0.064	-0.016	-.124**
	Sig. (2-tailed)	0.000		0.000	0.000	0.053	0.620	0.000
	N	936	945	945	945	922	944	945
Work-Related Negative Stress	Correlation	-.119**	-.227**	1.000	-.138**	.158**	.066*	0.050
	Sig. (2-tailed)	0.000	0.000		0.000	0.000	0.043	0.128
	N	936	945	945	945	922	944	945
Age	Correlation	-.190**	-.115**	-.138**	1.000	-0.014	-.204**	.058*
	Sig. (2-tailed)	0.000	0.000	0.000		0.623	0.000	0.042
	N	1225	945	945	1237	1203	1229	1237
p_income /100	Correlation	.067*	-0.064	.158**	-0.014	1.000	.361**	.507**
	Sig. (2-tailed)	0.021	0.053	0.000	0.623		0.000	0.000
	N	1192	922	922	1203	1203	1198	1203
Educational Year	Correlation	.087**	-0.016	.066*	-.204**	.361**	1.000	.228**
	Sig. (2-tailed)	0.002	0.620	0.043	0.000	0.000		0.000
	N	1218	944	944	1229	1198	1229	1229
Male_ Dummy	Correlation	-0.031	-.124**	0.050	.058*	.507**	.228**	1.000
	Sig. (2-tailed)	0.284	0.000	0.128	0.042	0.000	0.000	
	N	1225	945	945	1237	1203	1229	1237

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

4.1.2. Scatter Plots

Figure 1 shows the relationship between "age" and "self-reported physical health". As expected, physical health inversely correlates with age.

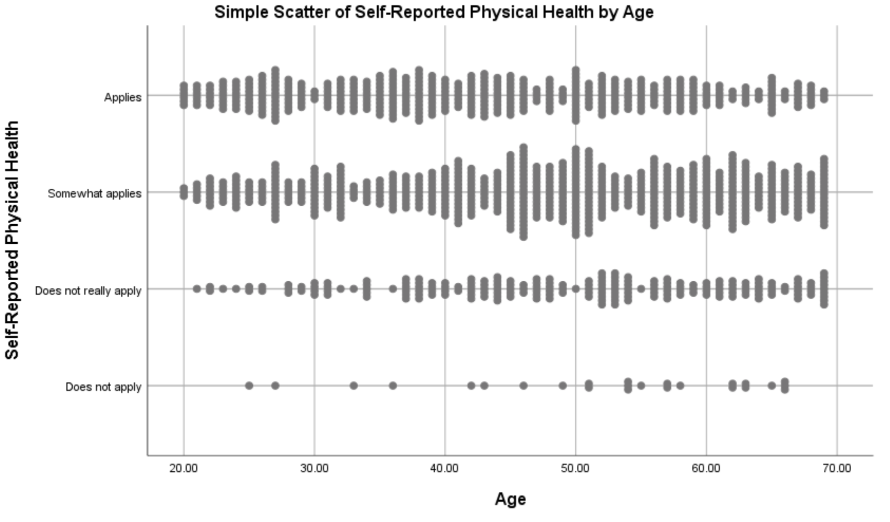


Figure 1

As shown in Figure 2, the correlation between self-reported physical health and work-related negative stress is not necessarily clear.

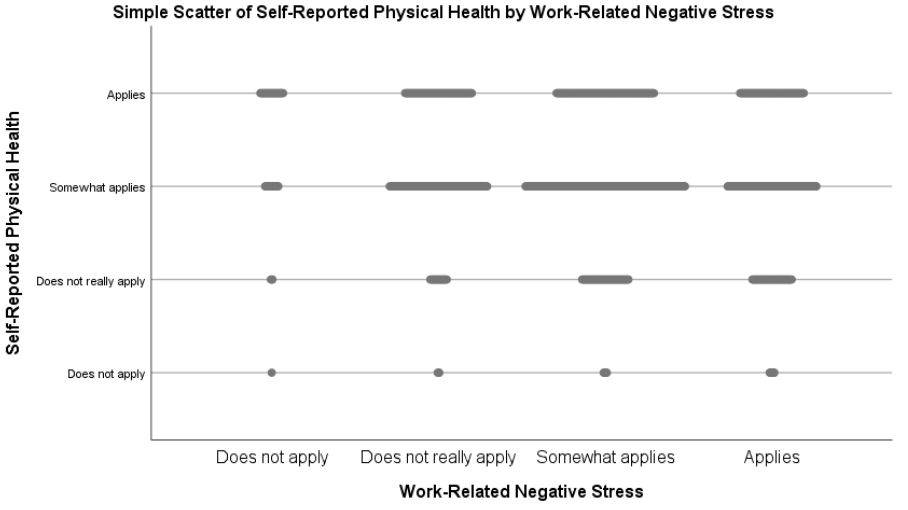


Figure 2

In Figure 3, not clear, however, there may be a positive correlation between psychological safety and self-reported physical health.

Figure 3

Additionally, a minor positive correlation can be observed between psychological safety and work-related negative stress, as shown in Figure 4.

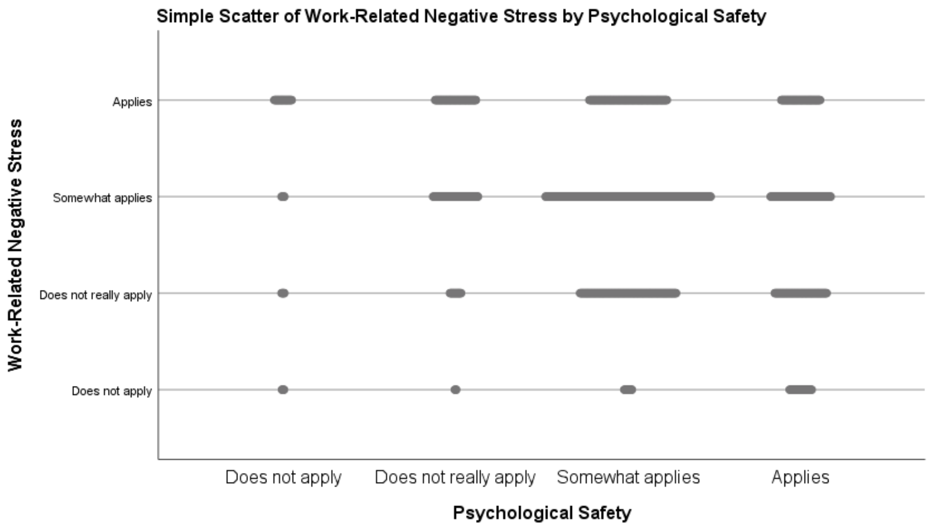


Figure 4

4.2. Multivariate Analyses

As demonstrated by several of bivariate analyses above, the relationships between the key variables are not very strong, if present at all. In the following analyses incorporating multiple variables simultaneously to control confounding effects, we will examine whether these relationships remain.

4.2.1. Ordinal Logistic Regression Result

4.2.1.1. Model Fitness

This analysis initiated with OLR to investigate the variable 'Self-Reported Physical Health' (recoded into 3-point scale for OLR), categorized into three distinct levels: 'Does not apply + Does not apply much', 'Somewhat applies', and 'Applies'. The distribution of responses showed that 16.0% of the participants perceived their physical health as 'Does not apply + Does not apply much,' 49.5% as 'Somewhat applies,' and 34.5% as 'Applies.'

Table 2. Physical Health Outcomes Distribution

		N	Marginal Percentage
Self-R_ PhysicalHealth _3degree	Does not apply + Does not really apply	146	16.0%
	Somewhat applies	451	49.5%
	Applies	315	34.5%
Valid		912	100.0%
Missing		325	
Total		1237	

The aim is to assess the influence of psychological safety and other predictors on physical health outcomes, utilizing data from 912 valid respondents. As indicated by the results below, the analysis suggests that psychological safety, even when controlling for other significant predictors, remains effective for physical health outcomes, although the strength of this effect is not very strong.

The modest value of the Nagelkerke pseudo-R-squared (.068) is not necessarily a significant concern in this case, as the goal is not to establish a model that predicts the outcome. However, it does suggest the possibility that there are other important variables not included in this study. Additionally, the large number of cells with zero frequencies could be a concern and might warrant model adjustments.

Despite these caveats, other indicators demonstrate that this model still adequately fits the data. Using a logit link function, the final model significantly improved from the intercept-only model, with a Chi-Square of 55.392 (df = 6, p < .001). The goodness-of-fit tests also indicated adequate model fit, with a Pearson Chi-Square of 1766.604 (df = 1750, p = .386) and a Deviance Chi-Square of 1736.277 (df = 1750, p = .588).

4.2.1.2. Key Predictive Factors

The analysis identified several key factors that are statistically significant in relation to physical health outcomes:

– Age: For every one-year increase in age, there is a decrease in the likelihood of reporting better health outcomes (Estimate = -.024, $p < .000$), suggesting that younger respondents tend to report better health.

– Income: Higher personal income is associated with slightly better health outcomes (Estimate = .045, $p = .026$), although the effect size is minimal.

– Psychological Safety: Individuals who feel comfortable discussing mistakes at work are more likely to report better physical health outcomes (Estimate = .255, $p = .003$), highlighting the importance of a psychologically safe environment.

– Work Related Stress: Higher reported stress levels are associated with worse health outcomes (Estimate = -.299, $p < .000$), underscoring the negative impact of stress on physical health.

“Male dummy” and “education year” were not found to significantly predict physical health outcomes in this model, indicating that the effect of psychological safety on physical health is relatively consistent across genders and educational levels.

Table 3

Parameter Estimates

		Estimate	Std. Error	Wald	df	Sig.	Odds Ratio	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Threshold	[Self_R_ PhysicalHealth 3dgr = 1.00]	-2.477	0.776	10.188	1	0.001		-3.998	-0.956
	[Self_R_ PhysicalHealth 3dgr = 2.00]	-0.069	0.771	0.008	1	0.929		-1.581	1.442
Location	PsychologicalSafety	0.255	0.087	8.654	1	0.003	1.29	0.085	0.426
	p_income_ dividedBY100	0.045	0.020	4.931	1	0.026	1.05	0.005	0.085
	age_	-0.024	0.006	18.297	1	0.000	0.98	-0.034	-0.013
	Education_Year	0.022	0.038	0.335	1	0.563	1.02	-0.052	0.096
	Work_RelatedStress	-0.299	0.078	14.563	1	0.000	0.74	-0.453	-0.145
	Male_Dummy	-0.227	0.145	2.476	1	0.116	0.80	-0.511	0.056

Link function: Logit.

4.2.1.3 Test of Parallel Lines

The test of parallel lines (Chi-Square = 12.211, df = 6, p = .057) implied that the relationships between outcome groups are consistent. These findings demonstrate that psychological safety, along with other demographic factors, plays a significant role in predicting physical health outcomes within the workplace.

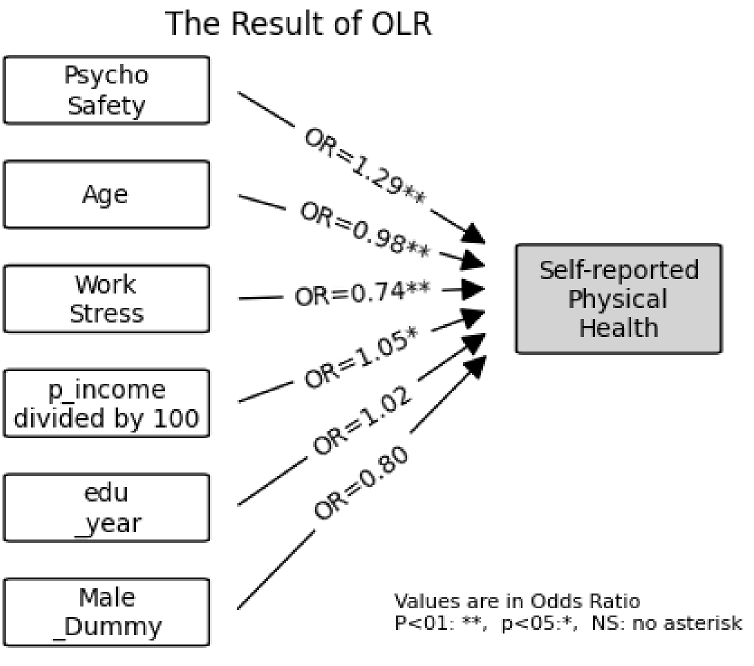


Figure 5

The diagram above is a Directed Acyclic Graph (DAG) illustrating the OLR model and its results mentioned earlier. These insights lay the groundwork for the subsequent analysis with Causal Forest DML and inform the discussions on the potential implications of psychological safety within the workplace.

4.2.2. Causal Forest Double Machine Learning (Causal Forest DML) Results

The following are the results of the Causal Forest DML analysis[5], which assessed how psychological safety influences physical health outcomes while controlling for covariates (with "employment status" added on top of the covariates used in the OLR model). This method is well-suited for this research as it effectively estimates varying treatment effects (in this particular model, “psychological safety at work”) across diverse individuals.

4.2.2.1. Average Treatment Effects and The Decision Tree

The histogram in Figure 6 illustrates a wide distribution of treatment effects with an average of 0.136, highlighted by a dashed line, suggesting that psychological safety, to some extent, have a beneficial impact on the physical health of respondents on average. However, the 95% confidence interval for the treatment effect ranges from -0.04 to 0.32, encompassing both negative and positive values. This indicates that while the average treatment effect is positive, there is a small amount of uncertainty about the uniformity of the treatment effect’s direction across different groups within the population. Nonetheless, as illustrated in Figure 6, for the majority of respondents, the intervention of psychological safety exhibited a positive effect. The question, then, is not whether the intervention is effective, but for whom it is effective. Figure 7 through 9 provide insights into this question.

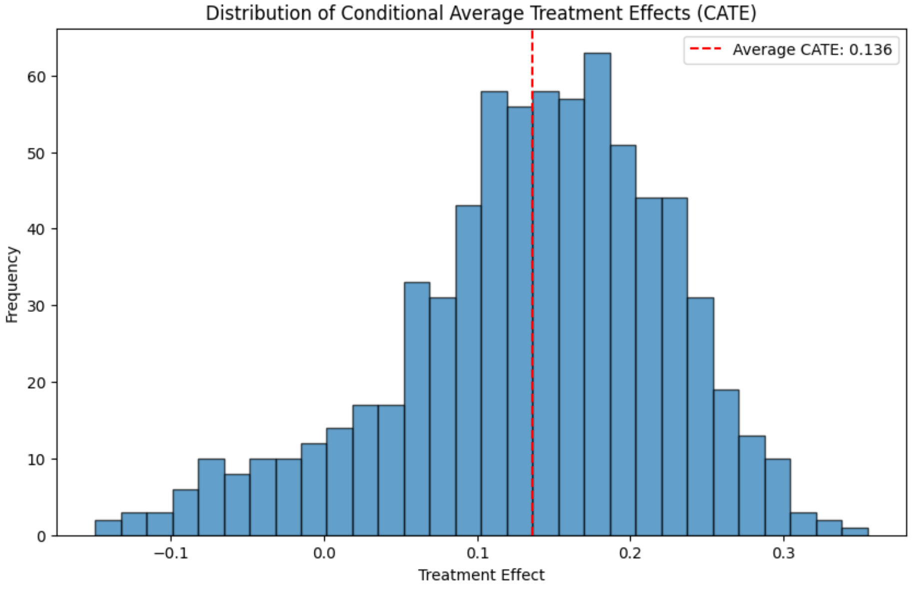


Figure 6

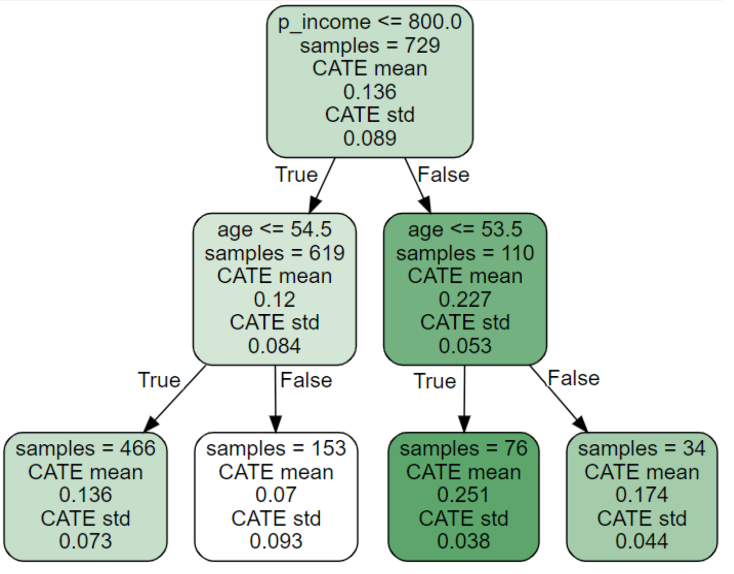


Figure 7. Averaged Decision Tree Generated from Ensemble Approach

Causal Forest DML generates many decision trees in its process using the bootstrap method. The decision tree in Figure 7 is a concise summary of the overall effect of the entire forest to illustrate conditional average treatment effect (CATE) estimated. This way, it is used to understand complex causal relationship more clearly, unlike other machine learning approaches. This particular tree reveals significant heterogeneity in treatment effects based on personal income and age. While psychological safety generally benefits respondents, the magnitude and consistency of this effect vary:

– Higher-income respondents (more than 8,000,000 yen/year), especially those younger than 53.5 years old, benefit the most from the intervention, with the highest and most consistent positive treatment effects (CATE = 0.251, standard deviation = 0.038).

– Lower-income respondents still experience a positive if they are 54.5 years old or younger. However, the treatment effect is trivial for those older than 54.5.

"Age" and "personal income" are the two key variables shaping the effect of psychological safety at work on physical health. The following two figures illustrate conditional treatment effects in relation to these 2 variables by plotting each individual's data.

4.2.2.2. Subgroup Analysis and Implications

In Figure 8, CATE by age is displayed, showing the relationship between age and the treatment effect of psychological safety. Note that in the process of computing the values of CATE, the confounding effects of covariates are controlled.

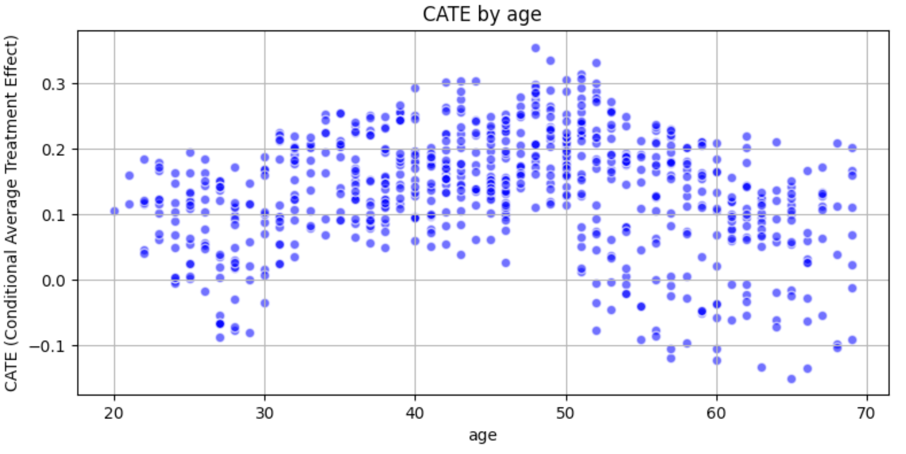


Figure 8

A further analysis focused on a specific age subgroup (ages 31 to 51) revealed a higher CATE of 0.174 with 95% confidence interval (CI) of (-0.00, 0.35) compared to overall ATE. This suggests that the beneficial effects of psychological safety are more pronounced within this age group.

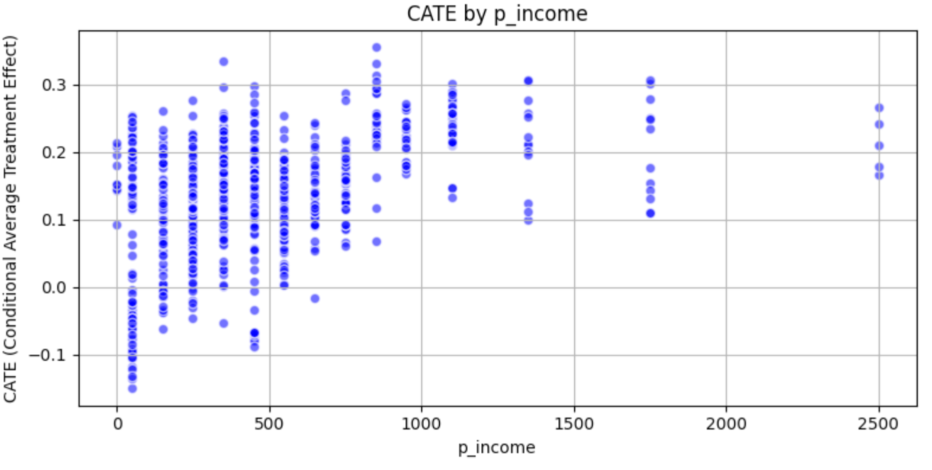


Figure 9

The scatter plot (Figure 9) indicates that personal income is also a significant factor in determining the effectiveness of psychological safety interventions. The CATE is 0.174 for those with a personal income of 500 or greater, with a 95% CI of (-0.009, 0.358). Higher-income individuals tend to benefit more consistently from the intervention, while lower-income individuals exhibit more variability in their responses.

4.2.2.3. Checking Overfit

Although this study does not aim to establish a predictive model, it followed the common practice in machine learning of dividing the dataset for training and testing to check for overfitting. In this study, 20% of the whole sample was separated for testing. MSE (Mean Squared Error) is a common measure used to evaluate the performance of a predictive model. Generally, a significant difference between training and test MSE can indicate overfitting if the test MSE is substantially higher than the training MSE. For the model generated in this study, the training MSE is 9.585, and the test MSE is 10.328, with a sample size of 912. The 7.8% increase in MSE from training to test is generally acceptable. This suggests that the model has a good balance between fitting the training data and generalizing to new data.

4.2.2.4. Feature Importance in Treatment Effects

As an approach extended from decision tree analysis, Causal Forest DML allows for the calculate of an indicator called "feature importance" for each covariate (feature). This metric evaluates how important each covariate is to the model, indicating the extent to which a particular covariate contributes to the splits in the decision trees. Thus, it is inherently different from the concept of direct/indirect effects of covariates used in regression analysis. For this reason, it is not common to calculate confidence intervals for feature importance values. The method used to calculate feature importance involves a probabilistic element, which means that the values themselves inherently include a certain level of uncertainty.

The importance of various features in determining the treatment effects is as follows: "Age" appears to be the most influential feature (importance: 0.39), followed by "personal income" (importance: 0.26). The remaining features are as follows: "work-related negative stress" (importance: 0.10), "years of education" (importance: 0.10), "gender" (importance: 0.04), "employment status" (importance: 0.04).

5. Discussion

This study investigated the relationship between psychological safety and health outcomes in the workplace, using a dual-analytical approach that combines OLR and Causal Forest DML. The results from both methods demonstrated an unignorable, although not as clear-cut as hoped, seemingly direct impact of psychological safety on physical health outcomes even after controlling for work-related stress and other basic ascriptive variables.

The OLR analysis demonstrated how “psychological safety at work” influences “physical health” outcomes, even after subtracting its indirect effects through covariates in the model. Additionally, the analysis identified “age,” “work-related negative stress,” and “personal income” as statistically significant predictors, alongside psychological safety. This indicates that in the workplace environment, increased psychological safety could directly benefit physical health.

Further, the Causal Forest DML analysis showed variability in the effects of psychological safety, notably emphasizing its significant benefits for employees aged 31 to 51. The heterogeneity of the treatment effect for those who are 30 years old or under makes sense considering the fact that most of them are physically fit in the first place, as indicated in the descriptive analysis (see Figure 1). Conversely, among those 52 years old and older, the treatment effect also varied.

This machine learning method generated the average decision tree, highlighting key variables such as “personal income” and “age,” along with node split criteria. The distinctions clarified through these results not only support the initial findings from the OLR but also deepen our understanding by pinpointing workforce segments that are particularly responsive to psychological safety interventions.

Both methods support the initial research question; nevertheless, several aspects require careful interpretation of these results. In OLR, the mild pseudo R-squared suggests that there may still be important variables not yet considered in the model. The presence of confounding variables can complicate the interpretation of the model and may distort the relationship between the explanatory variables and the outcome variable. At the time OLR was implemented, one possible confounding variable was employment status, which was excluded from the OLR model to avoid further increasing the number of zero frequency cells. However, as shown so far, the flexible Causal Forest DML model allowed the inclusion of this variable, and it accounts for only 4% of the total covariance comprised of six variables. Another factor worth considering, though not included in the dataset, is that corporate managements committed to providing a psychologically safer work environment might also encourage employees to enhance physical health by offering in-office fitness gyms or sport stretch services at a discount or even free of charge. However, even this senario would not explain the results from Causal Forest DML, which demonstrated the effect of the intervention across a certain age range, as clearly illustrated in Figure 8.

6. Conclusion

This analysis newly identified a direct pathway where psychological safety enhances physical health without reducing mental stress, in addition to the conventional pathway where high psychological safety enhances physical health by reducing mental stress.

While existing literature positions psychological safety as a critical lever for fostering organizational success, the findings of this paper suggest that psychological safety directly influences physical health, even when controlling for work-related stress and other basic ascriptive variables, particularly for those aged 31 to 51. This complements studies by Edmondson (1999) and others, extending the conversation around psychological safety to include its tangible benefits on physical health, beyond the well-documented mental and emotional well-being. Although the concept of psychological safety has been quite popular in the fields of business management and organizational behavior, this newly illuminated direction requires sociological inquiries.

Additionally, the use of machine learning methods among sociologists has been modest, except for a few cases where they collaborated with data scientists. However, as demonstrated in this paper, the application of such methods should be more frequent. Specifically, methods suitable for sociological causal inference, such as Causal Forest DML, rather than those focused on predictions, should be combined with traditional statistical methods to achieve fruitful findings.

For future research directions, it would be valuable to re-investigate these findings by including additional control variables, considering the possibility of other mediating pathways beyond those accounted for in the current analysis model. Also, given the study's limitations, including its cross-sectional design, future research should aim to employ longitudinal studies to further elucidate the causal pathways through which psychological safety impacts physical health.

Appendix

The distributions of the valid respondents' basic attributes are as follows.

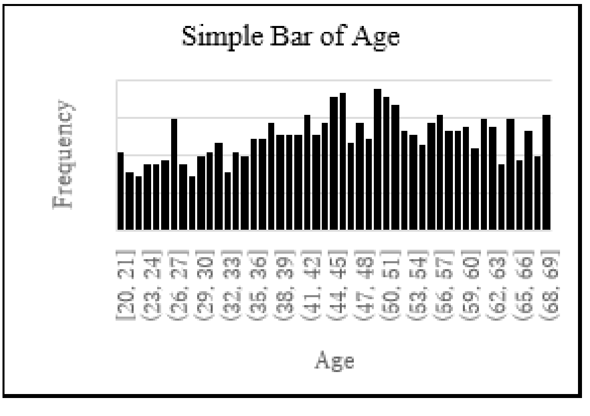


Figure 10

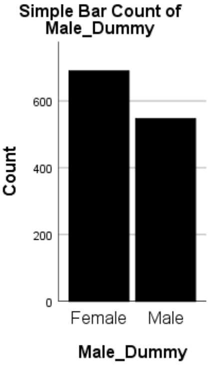


Figure 11

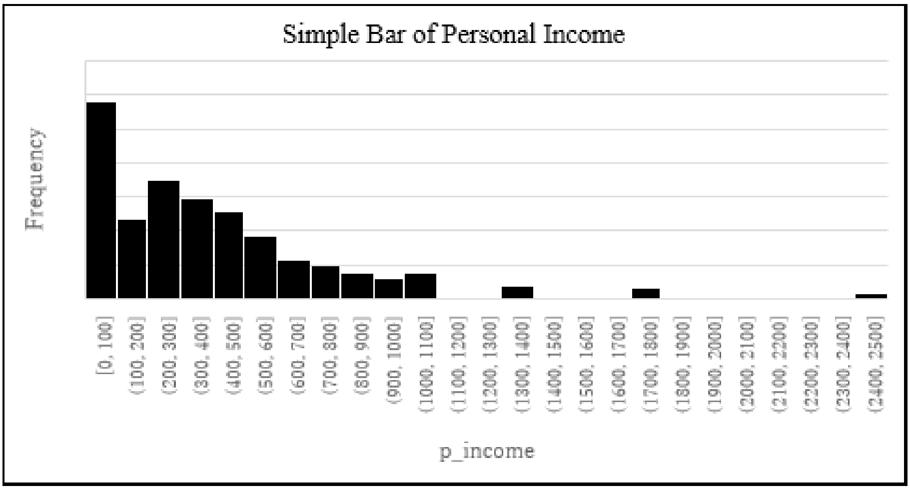
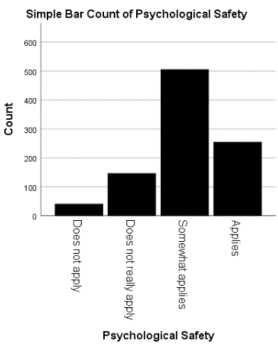
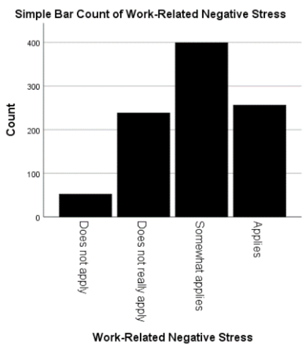


Figure 12





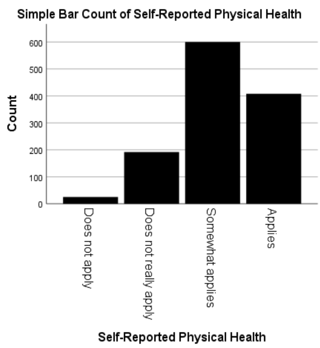


Figure 13 **Figure 14** **Figure 15**

Acknowledgements

This research was conducted with the support of the JSPS (Japan Society for the Promotion of Science) Grant-in-Aid for Scientific Research, Grant Number 19H01574. We are deeply grateful for the financial assistance provided, which made this study possible.

I would like to extend my deepest gratitude to my esteemed colleagues and the dedicated members of our research group, the 'Global Consumer Culture Research Group.' Their expertise and support have been invaluable in successfully conducting the social survey.

NOTES

1. Though there are several versions to scale "psychological safety" (for instance, Edmondson (1999) used 7 items, also sometimes only four out of the seven items are used), in this particular paper, due to the limitation of the questionnaire space, only one question was used.
2. While this method offers a sophisticated tool for exploring these dynamics, it is crucial to note that the cross-sectional nature of the dataset limits the ability to definitively establish causal relationships. Instead, Causal Forest DML aids in identifying areas where psychological safety might significantly influence health outcomes, suggesting the need for further investigation with longitudinal data.
3. Some of the early research on this method was conducted starting since 2016 by Susan Athey and Guido Imbens, along with many others such as Victor Chernozhukov et al. (2017). After further refinements, Causal Forest DML (v0.1) eventually became publicly available in the econML library by Microsoft Research in 2019. The library has since seen numerous updates and has evolved into what we know today.
4. For more detailed explanations, refer to Jennie E. Brand et al. (2021).
5. In this study, Causal Forest DML analysis was implemented in Python version 3.10.12, utilizing the econML library. Python was chosen for its flexibility and the powerful machine learning libraries it supports, which are well-suited for this advanced analytical approach.

REFERENCES

ATHEY, S. & IMBENS, G., 2016, Recursive Partitioning for Heterogeneous Causal Effects. *Proceedings of the National Academy of Sciences*, vol. 113, no. 27, pp. 7353 – 7360.

BRAND, J.E., XU, J., KOCH, B., GERALDO, P., 2021. Uncovering Sociological Effect Heterogeneity Using Tree-Based Machine Learning. *Socio Methodology*. vol. 51, no. 2, pp. 189 – 222.

CHERNOZHUKOV, V., CHETYERIKOV, D., DEMIRER, M., DUFLO, E.,
HANSEN, C., & NEWEY, W., 2017. Double/Debiased/Neyman Machine Learning of Treatment Effects. *American Economic Review*, vol. 107, no. 5, pp. 261 – 265.

EDMONDSON, A. C., 1999. Psychological Safety and Learning Behavior in Work Teams. *Administrative Science Quarterly*, vol. 44, no. 2, pp. 350 – 383.

EDMONDSON, A. C., 2018. *The Fearless Organization: Creating Psychological Safety in the Workplace for Learning, Innovation, and Growth.*: Wiley.

KAHN, W. A., 1990. Psychological Conditions of Personal Engagement and Disengagement at Work. *Academy of Management Journal*, vol. 33, no. 4, pp. 692 – 724.

NEWMAN, A., DONOHUE, R., & EVA, N., 2017. Psychological safety: A systematic review of the literature. *Human Resource Management Review*, vol. 27, no. 3, pp. 521 – 535.
SCHEIN, E. H., & BENNIS, W. G., 1965. *Personal and Organizational Change through Group Methods: The laboratory Approach*.: Wiley.

✉ **Dr. Satoshi Maeda**
Researcher number (e-Rad #): 80966325
J-GLOBAL ID: 202201001911673190
Researchmap Member ID: R000041102
Lecturer (PT) at Chuo University
Tokyo, Japan
E-mail: msatoshi003n@g.chuo-u.ac.jp