Research Article *Open Access*

Journal of Scientific Engineering Advances

ISSN: 3071-0723

DOI: doi.org/10.63721/26JSEA0146

Comparative Studies of Supervised and Reinforcement Learning for Analyzing Thermal Environments and Airflow in Bus Cabins

Peizhi Rong[1], Chi Ho Wong[1], Ka Chun Li[2] and Wai Keung Anthony Loh[1*]

[1]Division of Science, Engineering and Health Studies, School of Professional Education and Executive Development, The Hong Kong Polytechnic University, Hong Kong, China
[2]Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Hong Kong, China

Citation: *Peizhi Rong, Chi Ho Wong, Ka Chun Li, Wai Keung Anthony Loh (2026) Comparative Studies of Supervised and Reinforcement Learning for Analyzing Thermal Environments and Airflow in Bus Cabins. J. of Sci Eng Advances 2(4) 1-7. WMJ/JSEA-146*

Abstract

Thermal comfort and indoor air quality (IAQ) in public transit play a critical role in protecting passenger health and enhancing travel satisfaction. While bus cabin's airflow and passenger movement continuously change in real-world operation, capturing these cabin dynamics and comfort is essential for timely and reliable customer feedback and environment monitoring. Most existing approaches rely on supervised learning, which is well suited for classification and prediction when abundant labeled data are available. However, collecting sufficiently large datasets for dynamic bus environments can be expensive, time-consuming, and difficult to scale. In contrast, reinforcement learning (RL) learns an adaptive decision policy from interaction logic rather than fixed supervision, which may allow effective performance even with smaller datasets. In this study, we investigate whether reinforcement learning is suitable for analyzing thermal pattern and IAQ in dynamic bus cabin situations.

[*]**Corresponding author**: Wai Keung Anthony Loh, Division of Science, Engineering and Health Studies, School of Professional Education and Executive Development, The Hong Kong Polytechnic University, Hong Kong, China.

Submitted: 22.06.2026 **Accepted**: 26.06.2026 **Published**: 06.07.2026

Keywords: Transportation, AI, Thermal Engineering

Introduction

The thermal comfort and indoor air quality (IAQ) of public transportation systems directly affect passenger health, safety, and travel satisfaction. In closed long-distance bus cabins, limited ventilation and high occupancy can cause carbon dioxide (CO_2) to accumulate, particularly in spring and autumn when air-conditioning systems may be switched off or when older vehicles operate with insufficient fresh-air supply in developing countries. Previous studies have reported that in-vehicle CO_2 concentrations can reach 2500 to 3722 ppm, and may become even higher under fully occupied conditions [1]. Elevated CO_2 levels are associated with deeper breathing, dizziness, reduced attention, and drowsiness [2-4]. Under severe exposure conditions, poor cabin air quality can contribute to confusion, fatigue and delayed reaction under high-CO_2[5-8]. Thermally uncomfortable conditions may also increase traffic accident risk for drivers. Therefore, monitoring thermal comfort and air-quality conditions inside bus cabins are needed for safer and more comfortable public transport.

While thermal cameras and flow sensors provide direct measurements, their use in public transport is complicated by privacy concerns, clothing interference, and the high cost of scaling and maintenance. To bypass these hurdles, Computational Fluid Dynamics (CFD) can be used as a transient tool to gain an initial understanding of cabin dynamics without the need for intrusive hardware. CFD provides a promising alternative for analyzing cabin airflow and thermal comfort. Regarding the time, cost, and controllability of physical experiments, simulation is often preferred for forecasting airflow in buses, as it can reveal local flow-field heterogeneity and its influence on thermal perception [9-11]. By controlling boundary conditions such as inlet velocity, air-supply direction, exhaust location, occupancy distribution, and heat sources, CFD allows different operating scenarios to be studied systematically without disturbing passengers or collecting identifiable thermal data.

To bridge the gap before collecting identifiable thermal data of passengers, CFD-generated simulations serve as a vital "synthetic" testing ground. By training models on simulated airflow and thermal patterns, AI could recognize cabin dynamics to quickly identify the most effective thermal comfort and air-quality conditions. While supervised learning efficiently maps these inputs for near-real-time monitoring, it is often limited by the scarcity of labeled, real-world data. Reinforcement learning (RL) addresses this by focusing on sequential decision-making; it optimizes control policies to maximize long-term rewards, making it better suited for navigating evolving cabin states even when large identifiable datasets are unavailable [12-13]. Hence, this work investigates how close reinforcement learning can match, or even exceed, the accuracy of supervised learning when trained on a limited thermal dataset [14-18]. By comparing the two, we aim to determine if RL's reward-based feedback offers a more computationally efficient and privacy-conscious alternative for real-time bus-cabin management.

Methodologies

Generate CFD Images for Training

We introduce a three-level architecture of CFD Visualization. First, it tackles multi-parameter coupling analysis by comprehensively considering the interaction effects of air conditioning outlet velocity (0, 0.5, 1.5, and 3 m/s), direction (across five orthogonal directions), and exhaust port position (front, middle, and rear). Second, it focuses on the visualization of flow field features by transforming PMV cloud images into interpretable visual representations, thereby overcoming the dimensional limitations inherent in traditional numerical analysis.

The first step in the CFD simulation design involves creating a detailed geometric model of a single-decker bus, constructed at a 1:1 scale with dimensions of 13.225 m in length, 2.645 m in width, and 2.568 m in height. This model includes 37 seats and a designated driver's area, providing an accurate representation of the interior layout in bus. For the meshing strategy, a mixed grid approach is utilized. The grid is designed to maintain a Y+ value of less than 5, with localized refinement in critical areas, particularly around the human mouth and nose, where a mesh size of 1.5 mm was employed. The total grid count reached 34,939,793, which was verified to ensure grid independence, confirming that the results are not significantly affected by mesh resolution.

In establishing the boundary conditions, the RNG k-ε turbulence model was selected based on the work of Tominaga and Stathopoulos [19]. This model is well-suited for capturing the complexities of turbulent flows within the bus environment [20,21]. The simulation parameters included a systematic exploration of variable combinations, resulting in a total of 45 groups derived from three different wind speeds, five distinct directions, and three exhaust port positions. From these combinations, ten typical working conditions were selected for analysis, as outlined in Table 1. To ensure the accuracy of the simulation, specific convergence criteria were established. The residual of the continuity equation was required to be less than 1×10^{-6}, while the residual of the energy equation was also set to remain below 1×10^{-6}. These criteria guarantee that the simulation converges with a stable and reliable solution, providing a solid foundation for subsequent analysis. On the other hand, we calculated the PMV index based on the formula in literature work (ASHRAE standard (2021) [22]).

$$PMV = \left(0.303e^{-0.036M} + 0.028\right) \cdot \left\{ \begin{aligned} &(M-W) - 3.05 \times 10^{-3}\left[5733 - 6.99\,(M-W) - p_a\right] \\ &- 0.42\left[(M-W) - 58.15\right] - 1.7 \times 10^{-5} M\left(5867 - p_a\right) - 0.0014M\left(34 - t_a\right) \\ &- 3.96 \times 10^{-8} f_{cl}\left[\left(t_{cl} + 273\right)^4 - \left(\bar{t}_r + 273\right)^4 - f_{cl} h_c\left(t_{cl} - t_a\right)\right] \end{aligned} \right\}$$

Table 1: Typical Simulation working Condition combinations [23], where D, F, B, L, and R are abbreviations for downward, thirty degrees forward, thirty degrees backward, thirty degrees leftward, thirty degrees rightward, respectively.

Experimental variables	Inlet velocity	Inlet direction	Exhaust location
Inlet velocity	0m/s	D mode	Rear
	0.5m/s	D mode	Rear
	1.5m/s	D mode	Rear
	3m/s	D mode	Rear
Inlet direction	1.5m/s	F mode	Rear
	1.5m/s	B mode	Rear
	1.5m/s	L mode	Rear
	1.5m/s	R mode	Rear
Exhaust location	1.5m/s	D mode	Front
	1.5m/s	D mode	Middle

The AI models

Supervised Learning in combination of transfer learning is used. Instead of training a model from scratch, the tool downloads a MobileNet model pre-trained on millions of images (like the ImageNet dataset) via the Google teachable Machine [18]. It uses the existing layers to extract "features" (shapes, textures, colors) and only trains the final layer to map those features to your specific categories. The simulation involves extracting 100 sub-planes at the height of the passenger's head ($Z = 1.5$ m) to generate a dataset of 1024×768 RGB cloud images [23]. This dataset comprises 10 categories, totaling 1,000 images, which are divided into a training set of 700 images and a test set of 300 images at a 7:3 ratio. To enhance the prediction accuracy, data augmentation techniques such as rotation (±5°) and brightness adjustment (±10%) are applied [24-26]. Additionally, standardization is performed by normalizing the pixel values to a range of [1]. The hyperparameters such as Batch size, epoch, learning rate will be analyzed [27-30].

Comparative analysis of reinforcement learning methods such as Q-learning in MATLAB will be conducted against the supervised learning technique [31]. To apply Q-learning to image classification, the task is reframed as a reinforcement learning problem where the standard components are mapped to specific RL elements: the state represents the input image or features extracted, the action corresponds to the available class labels, and the reward provides positive feedback for correct classifications versus negative or zero feedback for errors [31]. Through this structure, the agent learns a policy to select the specific action—or class label—that maximizes the cumulative reward over time, effectively treating the classification decision as an optimal strategy.

Results and Discussion

First, we conduct training in AI using a small sample dataset, with 10 categories × 40 pieces = 400 pieces. The training set (280) and the test set (120) will be divided into a 7:3 ratio. As shown in Table 2, the predictive accuracy for SL is 100% under the categories of 0m/s, 0.5m/s, 3m/s, and L mode with significant features. All the tests are plotted in Table 2. The predictive results for SL of the B mode, F mode, Front, and R mode are also close to 100% in the supervised learning model. However, it is worth noting that the limited dataset causing SL fails to distinguish between the two categories of the exit at the back (1.5m/s) and the exit in the Middle (Middle) since the features of the two may be similar as seen in Figure 1a and 1c. The poor accuracy for SL in these categories could be due to overfitting because those 40 pictures were captured only about 200mm from the head, and each thermal comfort picture doesn't differ much.

Table 2: The Average Recognition rate of Each Category in the Small Sample Dataset. Supervised learning and reinforcement learning are abbreviated as SL and RL, respectively.

	0m/s	0.5m/s	1.5m/s	3m/s	B mode	F mode	L mode	Front	R mode	Middle
SL	100%	100%	53%	100%	98%	99%	100%	98%	99%	31%
RL	72%	73%	46%	88%	79%	76%	85%	77%	61%	28%

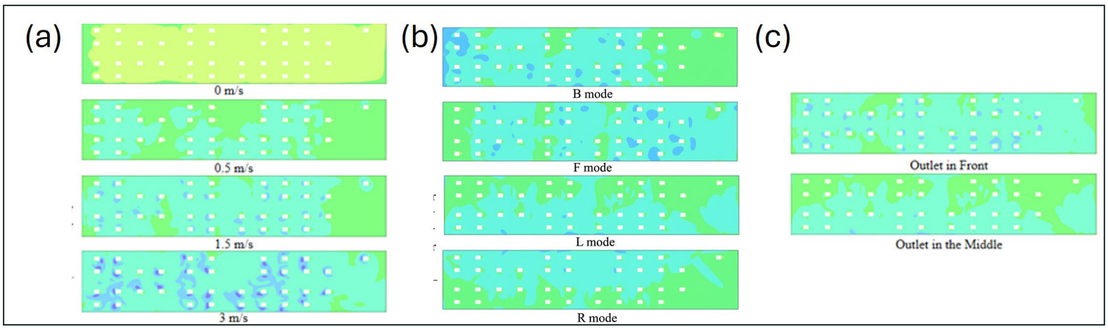

Figure 1: Sample Images for each of the Ten Categories. (a) Wind Speed is the Only variable; (b) The Wind Direction is the Only Variable; (c) The Position of the Exhaust is the Only Variable.

Despite the potential of reinforcement learning (RL) to handle small datasets and generalize well, its performance in this bus cabin study was inferior to supervised learning, trailing by approximately 20% in accuracy, except for the comparable accuracy for the categories of “1.5m/s” and “Middle” [32,33]. The performance gap is primarily due to reinforcement learning’s (RL) lack of explicit guidance compared to supervised learning (SL). The definition of a "large" or "small" dataset is context-dependent and should be determined on a problem-by-problem basis. The near-perfect accuracy of the SL model in most of the categories suggests that the dataset could be large enough for SL to distinguish features. In contrast, the RL agent faces a disadvantage as it tries to match the performance of a SL baseline [33]. Since the visual differences between categories are minimal (such as the thermal comfort images captured only 200mm from the head), the RL agent may struggle to correlate specific actions with sparse rewards. The complexity of the reward signal and the high-dimensional state space of the images likely led to instability during convergence, causing the RL model to falter where the supervised model could still find a reasonable mapping [32,33]. However, in the “1.5m/s” and “Middle” categories, RL performs comparably to SL. This party is likely because the datasets for these specific categories are not large enough for SL to establish a dominant predictive advantage. In such cases, RL’s reliance on trial-and-error and strategic decision-making allows it to remain competitive where SL lacks sufficient data to generalize effectively.

We further optimized various training parameters, including epochs, batch size, and learning rate [29,30]. As shown in Figure 2, while adjusting these training parameters results in slight changes in recognition accuracy, the recognition rate for the 1.5 m/s category remains low. Since the visual features of the 1.5 m/s exit and the middle exit are similar, the neural network faces a "feature overlap" problem where the mathematical distance between the two classes is negligible. In this scenario, the model is likely converging to a local minimum where it may be unable to distinguish between these two specific patterns. Simply increasing the number of training iterations or adjusting the learning speed does not provide the model with the new discriminative information it needs to separate these categories. This suggests that the performance bottleneck may stem from low image resolution, insufficient dataset size, or a lack of unique spatial markers within the CFD-generated visuals. These factors prove that even a supervised model cannot overcome high levels of similarity in the raw input data.

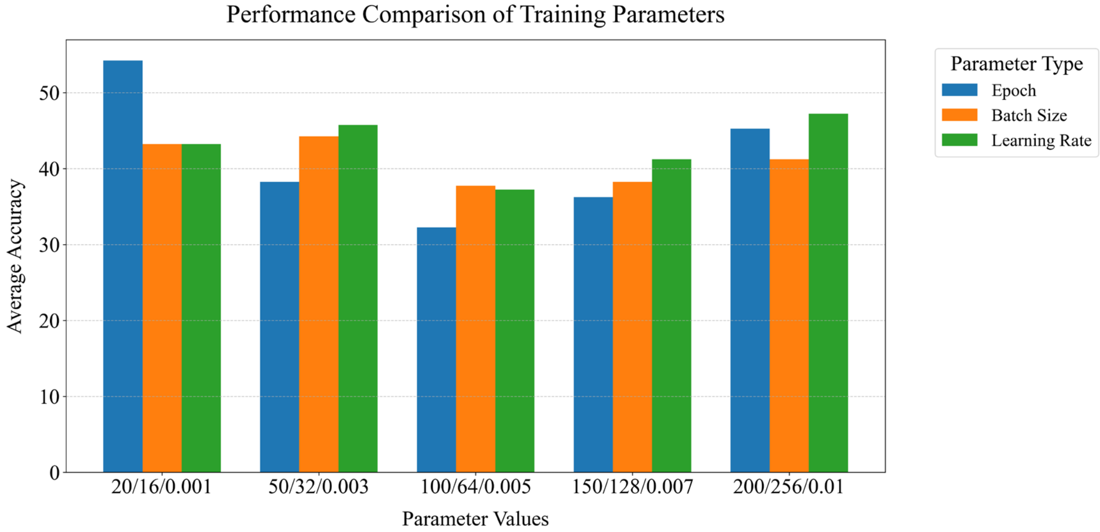

Figure 2: The Average Recognition Accuracy of Different Training Parameters (SL) in the Category of 1.5m/s.

Considering this, increasing the dataset from 40 to 100 images per category provides the supervised learning (SL) model with enough statistical variance to more successfully differentiate between the previously indistinguishable 1.5 m/s and middle exit categories. This surge in accuracy to near-perfection (99.87%–100%) confirms that the SL model's primary bottleneck is data scarcity rather than architectural capability and image resolution. However, the reinforcement learning (RL) model failed to exhibit a similar performance leap despite the increased sample size [33].

This disparity highlights a noticeable inefficiency in RL for this specific application, in which RL struggles to match SL with large image datasets but can achieve comparable performance when data is limited. This indicates that RL may be more resilient in scenarios where there is insufficient data for SL to establish a dominant predictive advantage.

Our work demonstrates that in the context of high-similarity fluid dynamics and thermal images, the trial-and-error mechanism of the RL model introduces a "sample inefficiency" where the agent may spend too much computational effort exploring incorrect labels before identifying the subtle visual cues that distinguish different airflow velocities. Consequently, the RL agent could not bridge the gap to the SL model's performance. These findings suggest that for monitoring bus cabin environments, where visual features are dense and airflow nuances are subtle, supervised learning (SL) is the more robust and efficient choice if a large dataset can be collected. However, in scenarios with smaller image datasets, the performance gap narrows, and neither SL nor RL may show a significant advantage over the other. The choice between SL and RL in this application could depend on the volume of passenger thermal images available for training.

Conclusion

This study investigated the efficacy of different AI methodologies for predicting thermal comfort and indoor airflow in dynamic bus cabin environments. The reinforcement learning model reached a performance ceiling below that of the supervised learning approach for a reasonable sample size. These results highlight a profound sample inefficiency in the reinforcement learning model when applied to bus cabin imagery. The exploratory trial-and-error nature of reinforcement learning may lack the direct gradient guidance found in supervised learning, making it prone to instability when faced with subtle thermal gradients and complex airflow patterns [34].

Author Contributions

Conceptualization: WKL; methodology: PR, CHW, WKL; Computation: PR; validation: PR; formal analysis: PR; data curation: PR; writing: PR, CHW, K.CL; manuscript editing: CH Wong, K.C.L, W.K.L; Visualization: PR, CHW, KCL WKL.

Data Availability Statement

The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments

The work described in this paper were fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project Ref No. UGC/FDS24/E03/22).

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Huang HL, & Hsu DJ (2009) Exposure levels of particulate matter in long-distance buses in Taiwan. Indoor Air 19: 234-242.
2. GD Mathur (2008) Field tests to monitor build-up of carbon dioxide in vehicle cabin with AC system operating in recirculation mode for improving cabin IAQ and safety. SAE International Journal of Passenger Cars – Mechanical Systems 1: 757-767.
3. Park SK, Kim KW, & Kwon HY (2009) Experimental study on the air quality of vehicle's cabin by evaluating CO_2 concentration and fine dust on the actual road. SAE Technical Paper
4. Čorňák Š, Braun P (2010) The evaluation of interior car's air quality and safety of traffic. Science & Military Journal 5: 36-39.
5. Čorňák Š, Horák V, Chládek Z, Ulman J (2012) The evaluation of air quality in military vehicles. Science & Military Journal 7: 50.
6. Seo J, Choi Y (2012) Estimation of the air quality of a vehicle interior: The effect of the ratio of fresh air to recirculated air from a heating, ventilation and air-conditioning system. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering 227: 1162-1172.
7. Chiu C F, Chen M H, Chang F H (2015) Carbon dioxide concentrations and temperatures within tour buses under real-time traffic conditions. PLOS ONE 10: e0125117.
8. Jung HS, Grady ML, Victoroff T, Miller AL (2017) Simultaneously reducing CO_2 and particulate exposures via fractional recirculation of vehicle cabin air. Atmospheric Environment 160: 77-88.
9. Chan A (2003) Commuter exposure and indoor–outdoor relationships of carbon oxides in buses in Hong Kong. Atmospheric Environment 37: 3809-3815.
10. Mansour MK, Musa MN, Hassan MNW, Saqr KM (2008) Development of novel control strategy for multiple circuit, roof top bus air conditioning system in hot humid countries. Energy Conversion and Management, 49: 1455-1468.
11. Constantin D, Mazilescu CA, Nagi M, Drăghici A, Mihărtescu AA (2016) Perception of cabin air quality among drivers and passengers. Sustainability 8.
12. Heo S, Nam K, Loy-Benitez J, Li Q, Lee S, et al. (2019) A deep reinforcement learning-based autonomous ventilation control system for smart indoor air quality management in a subway station. Buildings 202: 109440.
13. Valladares W, Galindo M, Gutiérrez J, Wu WC, Liao KK (2019) Energy optimization associated with thermal comfort and indoor air control via a deep reinforcement learning algorithm. Building and Environment 155: 105-117.
14. Liu L, Chen RC (2017) A novel passenger flow prediction model using deep learning methods. Transportation Research Part C: Emerging

Technologies 84: 74-91.
15. Song G, Ai Z, Zhang G, Peng Y, Wang W (2022) Using machine learning algorithms to multidimensional analysis of subjective thermal comfort in a library. Building and Environment 212: 108-790.
16. Fukami K, Fukagata K, Taira K (2019) Super-resolution reconstruction of turbulent flows with machine learning. Journal of Fluid Mechanics 870: 106-120.
17. Wang Y, Li Z, Yuan Z, Peng W, Liu T (2024) Prediction of turbulent channel flow using Fourier neural operator-based machine-learning strategy. Physical Review Fluids 9: 084-604.
18. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC (2018) MobileNetV2: Inverted residuals and linear bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 4510-4520.
19. Tominaga Y, Stathopoulos T (2007) Turbulent Schmidt numbers for CFD analysis with various types of flowfield. Atmospheric Environment 41: 8091-8099.
20. Luo Q, Ou C, Hang J, Luo Z, Yang H (2022) Role of pathogen-laden expiratory droplet dispersion and natural ventilation explaining a COVID-19 outbreak in a coach bus. Building and Environment 220: 109-160.
21. Zhu S, Lin T, Laurent JGC, Spengler JD, Srebric J (2022) Tradeoffs between ventilation, air mixing, and passenger density for the airborne transmission risk in airport transportation vehicles. Building and Environment 219: 109-186.
22. Olesen BW, Brager GS (2004) A better way to predict comfort: The new ASHRAE standard 55.
23. Rong P, Zhou W, Wang J, Zhang H (2024) Evaluating thermal comfort and air quality in buses with different ventilation strategies in Hong Kong. Renewable and Sustainable Energy Reviews 24: 103-449.
24. Shijie J, Ping W, Peiyi J, Siping H (2017) Research on data augmentation for image classification based on convolutional neural networks. Proceedings of the 2017 Chinese Automation Congress (CAC) 4165-4170.
25. Mikołajczyk A, Grochowski M (2018) Data augmentation for improving deep learning in image classification problem. 2018 International Interdisciplinary PhD Workshop (IIPhDW) 117-122.
26. Alomar K, Aysel HI, Cai X (2023) Data augmentation in classification and segmentation: A survey and new strategies. Journal of Imaging 9: 46.
27. Smith SL, Kindermans PJ, Ying C, Le QV (2018) Don't decay the learning rate, increase the batch size. arXiv preprint 11.
28. Kandel I, Castelli M (2020) The effect of batch size on the generalizability of convolutional neural networks on a histopathology dataset. IEEE Access 6: 312-315.
29. Afaq S, Rao S (2020) Significance of epochs on training a neural network. International Journal of Scientific & Technology Research 9: 485-488.
30. Jacobs RA (1988) Increased rates of convergence through learning rate adaptation. Neural Networks 1: 295-307.
31. Jang B, Kim M, Harerimana G, Kim J W (2019) Q-learning algorithms: A comprehensive classification and applications. IEEE Access 7: 133653-133667.
32. Xiang Li, Jinghuan Shang, Srijan Das, Michael S Ryoo (2022) Does self-supervised learning really improve reinforcement learning from pixels?. Advances in Neural Information Processing Systems (NeurIPS) 36.
33. Gabriel Dulac-Arnold, Nir Levine, Daniel J Mankowitz, Jerry Li, Cosmin Paduraru (2021) Challenges of real-world reinforcement learning: definitions, benchmarks and analysis. Machine Learning 110: 2419-2468.
34. Girolamo Macaluso, Alessandro Sestini, Andrew D Bagdanov (2024) Small dataset, big gains: enhancing reinforcement learning by offline pre-training with model-based augmentation. Computer Science and Mathematics Forum 9: 4.