Verbal Reports' Influence on Pilot Flight Performance and Mental Stress Under Spatial Disorientation

Yohan Kang; Myung Hwan Yun; Sungho Kim

BACKGROUND: Circumstances in flight can adversely affect a pilot's spatial abilities and lead to spatial disorientation (SD), increasing the potential for fatal accidents. To systematically understand the impact of SD, it is important to quantitatively and qualitatively analyze pilots' flight performance and mental stress, and to verify the effectiveness of verbal reports (VR) in enabling pilots to deal with an SD situation. This study investigated the effects of VR execution and type of SD on flight performance and mental stress.

METHODS: SD simulation experiments were conducted on 30 Air Force fighter pilots (15 in the VR group and 15 in the non-VR group) with electrocardiography (ECG) sensors attached. The pilots gave their VRs by immediately verbalizing their attention to instrument information and six potentially disorientating scenarios were implemented in each flight phase. Flight performance was analyzed using instructor evaluation and self-evaluation scores, and mental stress was measured using heart rate variability (HRV) and perceived distress score.

RESULTS: In maintaining flight performance, the VR group, respectively, had 8% and 10% higher scores for altitude and speed than the non-VR group. The self-evaluation scores were lowest for Coriolis, while Graveyard Spin was scored lowest in the instructor evaluations. Regarding mental stress, the VR group tended to have higher HFs and lower LF/HF ratios among HRV measures than the non-VR group, and an 11% lower perceived distress score. The highest perceived distress score was for Coriolis.

DISCUSSION: We suggest that pilots can be assisted to understand and overcome SD situations through VR.

KEYWORDS: verbal reports, vestibular and visual illusions, heart rate variability, instructor and self-evaluation, perceived distress.

Kang Y, Yun MH, Kim S. *Verbal reports' influence on pilot flight performance and mental stress under spatial disorientation*. Aerosp Med Hum Perform. 2020; 91(12):948–955.

Spatial disorientation (SD) is a major threat to pilots' flight safety. SD refers to pilots' inability to accurately recognize their position or aircraft's maneuver with respect to the direction of the Earth's gravity.[31] SD illusions include vestibular illusions such as Leans and Coriolis, and visual illusions such as the False Horizon and Black Hole Illusion.[4] In general, SD can be classified into three types according to whether a pilot consciously detects a disorientation situation.[2,25] In unrecognized SD (Type I), a pilot controls the aircraft without detecting any indications of SD, while in recognized SD (Type II), a pilot is consciously aware of the SD situation, but there is a discrepancy in the control of the aircraft between human senses and instrument information. Finally, in incapacitating SD (Type III), a pilot largely recognizes that he or she is disoriented but cannot interpret instrument information or control the aircraft under SD conditions. According to U.S. Air Force data analyzed between 1999 and 2009,[19] SD was an important factor in 11% of all aviation accidents and 42% of all mortalities. Because fighter aircraft account for 65% of aviation accidents caused by SD and the fatality rate of pilots is high, it is important to study how SD affects pilots and how to overcome it to ensure their flight safety.

Existing studies mostly focus on how SD affects pilots in terms of flight or cognitive performance; however, these studies do not comprehensively assess flight performance and mental stress using both quantitative and qualitative measures. For example, in some studies, changes in flight performance were

From Air Combat Command, Republic of Korea Air Force, and the Department of Industrial Engineering, Seoul National University, Seoul, Republic of Korea.

This manuscript was received for review in March 2020. It was accepted for publication in October 2020.

Address correspondence to: Sungho Kim, Department of Industrial Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea; ksh32035@snu.ac.kr.

DOI: https://doi.org/10.3357/AMHP.5620.2020

identified through deviations of certain flight specifications such as altitude, airspeed, attitude, or control reversal errors.[14,15,32] Others determined changes in cognitive performance through secondary tasks such as a duration discrimination task, a digit span task, and an addition task under SD conditions.[6,7,37] However, to systematically understand the impact of SD, the relationship between objective and subjective measures of flight performance and the mental stress of pilots under an SD condition must be analyzed.

In addition, previous studies of how to overcome SD focused on designing systems to support the functions of human sensory organs such as the visual, auditory, and tactile senses. However, research is needed to determine how pilots can overcome SD using conventional aircraft instruments. For instance, Malcolm[18] suggested a visual cue system that provides a virtual horizon to control spatial orientation in the pilot's peripheral vision. Lyons et al.[17] found that pilots overloaded with visual information could minimize workload and response time by gathering information through their auditory systems. In addition, Gilliland and Schlegel[5] and Rupert[28] proposed a method for preventing SD by enhancing situational awareness in flight through a head vibro-tactile display and mechanical tactile stimulators, respectively. However, a countermeasure strategy that enables pilots to overcome SD must be devised to help them proactively cope with SD situations.

Methods for overcoming SD that pilots can perform during a flight are a transfer to instrument flight rules, use of the autopilot function, transfer of aircraft control authority, and the execution of verbal reports (VR). In particular, VR is an action taken by a pilot to verbally express procedures or instrument information. Therefore, it can be used as evidence of attention allocation and information perception during a flight.[3] VR is executed using a concurrent or retrospective method. In the concurrent method, the pilot immediately reports the process of allocating attention and perceiving information while performing the task, while in the retrospective method, the pilot recalls the process after completing the task.[13] VR execution can improve pilots' cognitive functions during flight training, help them understand their psychological state in each flight phase,[26] and strengthen their situational awareness in abnormal situations, including SD.[33]

On the other hand, a pilot's physiological status can be quantitatively measured using bio signals such as an electroencephalograph, electrooculograph, electrocardiograph (ECG), and electrodermal activity. For example, Horng et al.[9] observed changes in visual functions under a Coriolis condition based on electroencephalograph and electrooculograph data. Tropper et al.[36] evaluated a systematic motion-based training program to cope with SD phenomena using ECG data, and Tamura et al.[34] examined electrodermal activity data on SD situations to develop indices of pilots' physiological changes. Specifically, a pilot's mental anxiety or stress in a SD situation can cause a change in the autonomic nervous system.[39] One method to measure this is by frequency analysis of heart rate variability (HRV).

The purpose of this study was to analyze the impact of pilot's VR execution according to type of recognized SD. Flight performance was measured using an instructor evaluation score and a self-evaluation score. Mental stress was measured using the HRV scales and a perceived distress score. A statistical analysis was performed to determine the difference in flight performance and mental stress according to VR execution and type of SD.

METHODS

Subjects

The experiment was conducted on 30 male Air Force fighter pilots (30.6 ± 3.7 yr) who were randomly assigned into the VR group (15 pilots) and non-VR group (15 pilots). The random assignment was performed by generating and allocating random numbers to participants. This study recruited fighter pilots enrolled in the aerospace physiological training course. Of the enrolled fighter pilots, none were female; thus, the researchers could only recruit male fighter pilots. To avoid the effects of bias due to the flight cycle or SD training experience on the experiment, pilots currently operating in fighter squadrons who had undergone at least two training sessions using an SD simulator were selected to participate in the experiment. In addition, the selected participants did not have a cardiovascular disease and were instructed to have sufficient rest the day before the experiment. Smoking and the consumption of caffeine were banned on the day of the experiment. The study protocol was approved in advance by the Institutional Review Board of the Korean Air Force Aerospace Medical Center (Approval number: ASMC-19-IRB-003). Each subject provided written informed consent before participating.

Equipment

This study measured pilots' flight performance and mental stress during SD induced in an SD simulator. The SD simulator—GL-4000 (ETC Aircrew Training Systems, Southampton, PA, USA)—is designed to provide pilots with the ability to recognize and cope with the risk of an SD situation. In addition, this device can realize 24 types of SD with high fidelity using 4 axes of motion freedom. The ECG measurement equipment used, the MP-160 (BIOPAC Systems Inc., Goleta, CA, USA), measures bio signals through wireless receivers with two sensors attached to the upper left and right chest and one to the lower left chest.

Variables

In this study, VR execution and type of SD were employed as the independent variables. In addition, two flight performance scales (instructor evaluation score and self-evaluation score), two mental stress scales (HRV measures and perceived distress score), and the simulator sickness score were used as the dependent variables. VR execution was implemented using the concurrent method, which allows a pilot to promptly state the instrument information visually scanned during the simulated flight. Previous studies using the concurrent method processed the VR data by transcribing, segmenting, and coding to analyze the pilot's underlying cognitive process.[26,33] Since the purpose of this study was to confirm the effect of VR execution under SD scenarios, however, this study only monitored whether pilots performed VR as evidence of attention to the instrument information. Six types of SD often encountered by pilots were applied in each flight phase, namely four vestibular illusions and two visual illusions, as shown in **Table I**.

The instructor evaluation score, as an objective variable of flight performance, was derived based on the ability to manage flight specifications in each flight phase. One pilot qualified as an instructor was assigned to the evaluation per Kallus et al.[10] using the Air Force Flight Evaluation Criteria. Blind evaluation was conducted to prevent bias that might arise during the assessment. The instructor pilot was unaware of the VR execution groups while giving evaluation throughout the experiment and was asked to perform the same procedure as in an actual flight assessment. In addition, the instructor pilot did not interact at all with the subject during the SD scenarios in order to prevent experimenter cueing. The ability to manage flight specifications was measured with the elements of altitude, speed, and attitude. A 10-point scale (8–10 points: excellent; 6–7 points: satisfactory; 5 points or less: unsatisfactory) was used to rate each element depending on the degree of deviation from standard specifications. The self-evaluation score, as a subjective variable of flight performance, was measured using a questionnaire on pilots' perceived performance ability to manage their flight status under the SD situations in each flight phase that was scored on a 7-point scale (1 point: very difficult; 4 points: neutral; 7 points: very easy) based on pilots' subjective feelings.

The HRV measures, variables that objectively measure mental stress, were low frequency (LF), high frequency (HF), and LF/HF ratio in the frequency domain.[35] LF refers to the spectral components of HRV in the range 0.04–0.15 Hz and HF those from 0.15–0.4 Hz. LF represents the activities of both the sympathetic nervous systems (SNS) and parasympathetic nervous systems (PNS), and HF is used as an indicator of PNS activity. LF/HF is the ratio between LF and HF, which indicates the overall balance of the autonomic nervous system.[8,22,27,29] Meanwhile, stress can simultaneously activate two axes: the hypothalamus-pituitary-adrenal and sympathetic-adrenal medullary axes. The activation of the hypothalamus-pituitary-adrenal system results in cortisol changes in body fluids, and the activation of the sympathetic-adrenal medullary system changes heart rate. When the autonomic nervous system is activated by stress, the heart rate may increase or decrease depending on the shift direction of the sympatho-vagal response.[20,40] Generally, it is known that as stress level increases, HF tends to decrease, while LF and the LF/HF ratio are more likely to increase. This is because the PNS withdraws the inhibitory effect, promoting the SNS's reactions to stress.[12] The perceived distress score, a subjective variable of mental stress, was evaluated using a questionnaire on pilots' perceived level of mental stress under the SD situations in each flight phase. A 7-point scale (1 point: very low; 4 points: neutral; 7 points: very high) was used to rate the items based on pilots' subjective opinions. **Table II** shows the experimental dependent variables used in this experiment. Finally, simulator sickness score was assessed using the simulator sickness questionnaire (SSQ).[11] The SSQ can be used to calculate nausea, oculomotor, disorientation, and total severity scores, reflecting pilots' perceived severity of simulator sickness. A 4-option scale (none, slight, moderate, and severe) was used to evaluate 16 symptoms.

Table I. General Description of the Six Types of Spatial Disorientation.

TYPES OF SPATIAL DISORIENTATION	DESCRIPTION
Vestibular Illusions	
Somatogravic Illusion	False climbing sensation caused by the increasing thrust of an aircraft.
Coriolis	False abrupt sensation of rotation when pilots move their head while the aircraft is turning.
Leans	False roll attitude sensation due to changes in the stimulus threshold of the semicircular canal.
Graveyard Spin	False rotation sensation when trying to level the aircraft following a prolonged descending turn.
Visual Illusions	
False Horizon	False horizontal sensation on virtual horizons due to the predominance of the peripheral visual field.
Black Hole Illusion	False shape or size sensation on the runway due to limited visibility in night-time landing situations.

Procedures

The experiment consisted of four stages: experimental preparation, preliminary experiment, main experiment, and postsurvey questionnaire. In the experimental preparation stage, participants were provided with a description of the experimental process, including flight procedures and specifications in each flight phase, and provided their informed consent in writing. Additionally, the participants of the VR execution group were asked to sequentially give a verbal report on the information they obtained from the aircraft instruments based on procedures specified in the Air Force manual. In the preliminary experiment stage, the participants attached ECG sensors to their bodies to measure ECG baseline data, and then conducted a practice flight for 5 min under normal flight conditions. In the main experiment stage, the ECG data were obtained as the participants flew for 15 min under specific types of SD situations in each flight phase. At the same time, the instructor evaluation was completed using the same cockpit instrument panel data displayed on the screen in the simulator control station by a pilot qualified as an instructor. The instructor pilot was asked to evaluate flight performance as objectively as possible based on the Air Force Flight Evaluation Criteria, and was blinded to the participants' group status to avoid unconscious bias. **Fig. 1** shows the flight profiles and types of SD employed in the experiment. In the postsurvey questionnaire stage, participants assessed their self-evaluation score, perceived distress score, and simulator sickness score, and then provided feedback and other opinions related to the experiment.

Table II. General Description of Experimental Dependent Variables.

EXPERIMENTAL VARIABLES	UNITS	DESCRIPTION
Flight Performance		
Instructor evaluation	Point	A 10-point scale measured by an instructor-qualified pilot using the Air Force flight evaluation criteria.
Altitude		
Airspeed		
Attitude		
Self-evaluation	Point	A 7-point scale measured by participants based on subjective flight evaluation judgments.
Mental Stress		
HRV measures	ms^2	
LF		Power in frequency range from 0.04 Hz to 0.15 Hz indicating sympathetic and parasympathetic nervous system activity.
HF		Power in frequency range from 0.15 Hz to 0.4 Hz indicating parasympathetic nervous system activity.
LF/HF ratio	--	Ratio between LF and HF indicating the balance of the autonomic nervous system activity.
Perceived stress	Point	A 7-point scale measured by participants based on subjective mental stress judgments.

HRV: heart rate variability; LF: low frequency; HF: high frequency.

ECG Data Processing

ECG data were processed in four steps: data extraction, data preprocessing, data normalization, and data integration. In the data extraction step, ECG raw data were collected using a sampling frequency of 2000 Hz for 5 min in the baseline and 15 min in the main experiment. The R-peak values were extracted from the ECG raw data using the Pan-Tompkins algorithm[24] based on the MATLAB program (ver. R2019a; MathWorks, Natick, MA, USA). In the data preprocessing step, a low-level artifact correction method was performed to remove any outliers of the R-peak values and an autoregressive method was employed to calculate LF, HF, and the LF/HF ratio in the frequency domain using the HRV analysis program Kubios HRV (ver. 3.3.0; Kubios, Kupio, Finland). The autoregressive method is known to be more effective in analyzing short-term HRV data than the fast Fourier transform method.[16] In the data normalization step, the preprocessed ECG data were normalized through the ratio of individual data to the baseline data to offset differences in the size of the bio signal data for each participant. In the data integration step, 2-min analysis intervals during the six types of SD phenomena in each flight phase were determined, and the data of each analysis interval for each individual were combined into one data set.

Statistical Analysis

In this study, a two-way mixed analysis of variance (ANOVA) was conducted at a significance level of 0.05 to investigate the effects of VR execution and type of SD on flight performance and mental stress level. VR execution was analyzed as a between-subject variable and type of SD as a within-subject variable. The results of the ANOVA for each dependent variable were applied to the modified *P*-value by performing Greenhouse-Geisser adjustments when Mauchly's test of sphericity was violated. Multiple comparisons were implemented with a Bonferroni correction to the significance level for the effects of VR execution and type of SD. For the simulator sickness score, a 2-sample *t*-test was performed at a significance level of 0.05 to evaluate the effect of VR execution using the statistics program R (ver. 3.6.1; The R Foundation, Vienna, Austria).

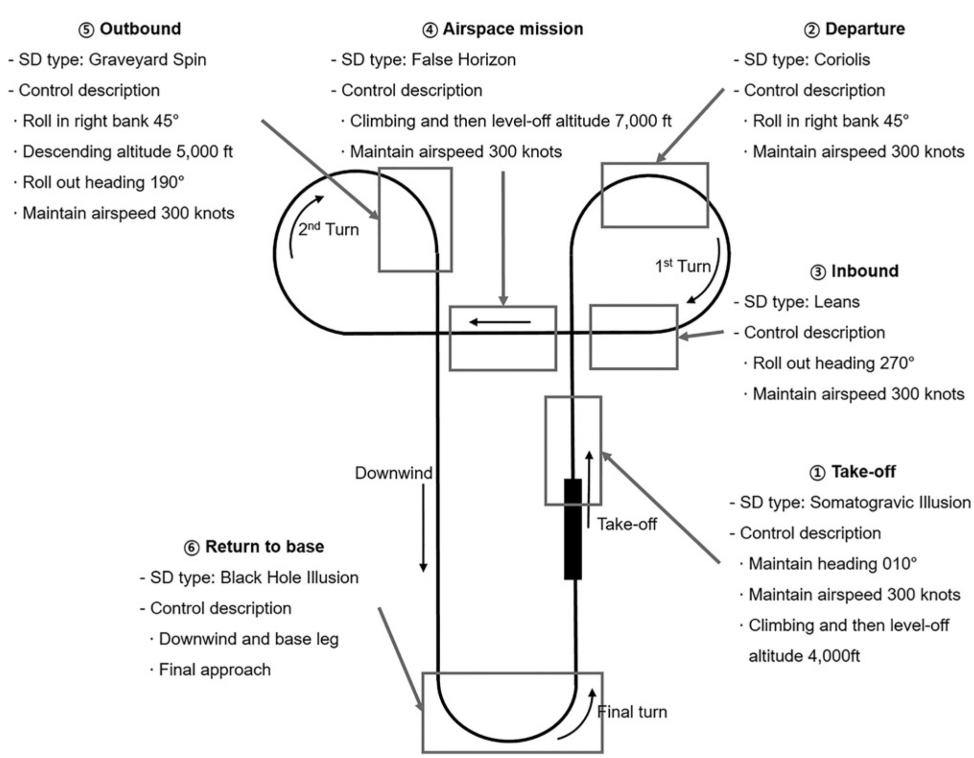

Fig. 1. General description of flight profiles. SD: Spatial disorientation.

RESULTS

For flight performance, the results showed a significant difference in instructor evaluation scores of altitude and airspeed for the main effect of VR execution, and in all instructor evaluation scores and self-evaluation scores for the main effect of type of SD. Regarding mental stress, significant differences were evident in only the perceived distress scores for both main effect of VR execution and type of SD. The interaction effects between VR execution and type of SD did not significantly affect flight performance and mental stress. **Table III** and **Table IV** summarize the results for the descriptive statistics and ANOVA, respectively.

Flight Performance

The main effect of VR execution significantly affected instructor evaluation scores of altitude and airspeed. However, no significant impact on the self-evaluation score was evident. The altitude score in the VR execution group (7.62 ± 0.10) was 8% higher than in the non-VR execution group (7.07 ± 0.10), and the airspeed score in the VR execution group (7.63 ± 0.10) was 10% higher than in the non-VR execution group (6.96 ± 0.10). Furthermore, the analysis indicated that the main effect of type of SD significantly influenced all the instructor evaluation scores: altitude, airspeed, and attitude. In addition, the self-evaluation scores were significantly affected. The altitude score for Graveyard Spin (6.63 ± 0.18) was significantly lower than for Somatogravic Illusion (7.63 ± 0.16), Coriolis (7.57 ± 0.15), Leans (7.53 ± 0.17), and False Horizon (7.73 ± 0.10). The airspeed score for Graveyard Spin (6.63 ± 0.19) was significantly lower than for Leans (7.70 ± 0.15) and False Horizon (7.57 ± 0.14). Finally, the attitude score for Graveyard Spin (5.90 ± 0.15) was significantly lower than for Somatogravic Illusion (7.00 ± 0.17), Coriolis (6.83 ± 0.17), Leans (6.90 ± 0.15), False Horizon (6.73 ± 0.16), and the Black Hole Illusion (7.37 ± 0.19). On the other hand, the self-evaluation scores were significantly lower for Coriolis (2.37 ± 0.25) than Somatogravic Illusion (4.10 ± 0.24), Leans (3.50 ± 0.21), and False Horizon (3.73 ± 0.25). The interaction of VR execution and type of SD did not have a significant effect on altitude, airspeed, or attitude in the instructor evaluation and self-evaluation.

Mental Stress

A significant main effect of VR execution and type of SD was not found in any HRV scale. However, this study identified that the VR execution group tended to have a higher HF and lower LF/HF ratio than the non-VR execution group, but the difference between the two groups in LF was relatively small. On the other hand, the perceived distress score had a significant effect on both VR execution and type of SD. The perceived distress scores were significantly lower in the VR execution group (4.83 ± 0.16) than in the non-VR execution group (5.42 ± 0.13) by 11%. The perceived distress scores for Coriolis (6.53 ± 0.12) were significantly higher than for Somatogravic Illusion (4.00 ± 0.27), Leans (4.70 ± 0.22), False Horizon (4.60 ± 0.16), and Graveyard Spin (4.90 ± 0.20). The interaction between VR execution and type of SD did not significantly impact the LF, HF, and LF/HF ratio in the HRV measures and perceived distress score.

Simulator Sickness

The main effect of VR execution was not significant for nausea [$t(28) = 0.418$, $P = 0.679$], oculomotor [$t(28) = -0.725$, $P = 0.475$], disorientation [$t(28) = 0.367$, $P = 0.717$], or the total severity score [$t(28) = -0.205$, $P = 0.839$] of the SSQ subscales. The SSQ scores of the VR execution group were as follows: nausea, 1.91 ± 1.02; oculomotor, 6.06 ± 1.52; disorientation, 4.64 ± 1.75; and total severity score, 4.99 ± 0.87. The SSQ scores for the non-VR execution group were as follows: nausea, 2.54 ± 1.13; oculomotor, 4.55 ± 1.44; disorientation, 5.57 ± 1.82; and total severity score, 4.74 ± 0.85.

DISCUSSION

This study systematically investigated the characteristics of flight performance and mental stress of pilots based on VR execution and type of SD. The participants were divided into two groups according to VR execution, and six types of SD that pilots often experience during an actual flight were employed in a simulator environment. The effects of VR execution and type of SD were verified through the instructor and self-evaluation scores in terms of flight performance, and mental stress with the HRV scales and perceived distress scores. The results of this study objectively confirm that pilots' VR execution in SD situations can help to improve their flight safety.

Table III. Summary of the Analysis of Variance (ANOVA).

DEPENDENT VARIABLE	VR			SD			VR×SD		
	DF	F	P	DF	F	P	DF	F	P
Flight performance									
Altitude	(1, 28)	9.87	0.004	(5, 140)	10.58	< 0.001	(5, 140)	1.26	0.372
Airspeed	(1, 28)	11.37	0.004	(5, 140)	8.04	< 0.001	(5, 140)	1.39	0.287
Attitude	(1, 28)	3.06	0.091	(3.8, 105.5)	13.82	< 0.001	(3.8, 105.5)	1.29	0.239
Self-evaluation	(1, 28)	0.49	0.490	(5, 140)	14.64	< 0.001	(5, 140)	0.61	0.270
Mental stress									
Normalized LF	(1, 28)	0.26	0.613	(1.9, 54.4)	1.15	0.324	(1.9, 54.4)	0.55	0.577
Normalized HF	(1, 28)	3.77	0.062	(1.9, 52.0)	0.54	0.575	(1.9, 52.0)	0.69	0.498
LF/HF ratio	(1, 28)	1.45	0.239	(2.3, 65.3)	0.56	0.598	(2.3, 65.3)	1.15	0.329
Perceived stress	(1, 28)	9.72	0.004	(5, 140)	28.13	< 0.001	(5, 140)	1.08	0.372

DF: degree of freedom; VR: verbal reports; SD: spatial disorientation; LF: low frequency; HF: high frequency.

Table IV. Summary of Descriptive Statistics.

	VR EXECUTION		SD TYPES					
DEPENDENT VARIABLE	VR	NON-VR	SD1	SD2	SD3	SD4	SD5	SD6
Flight performance								
Altitude (Point)	7.62^{A} (0.10)	7.07^{B} (0.10)	7.63^{A} (0.16)	7.57^{A} (0.15)	7.53^{A} (0.17)	7.73^{A} (0.10)	6.63^{B} (0.18)	6.97^{AB} (0.19)
Airspeed (Point)	7.63^{A} (0.10)	6.96^{B} (0.10)	7.00^{AB} (0.19)	7.43^{AB} (0.19)	7.70^{A} (0.15)	7.57^{A} (0.14)	6.63^{B} (0.19)	7.43^{AB} (0.18)
Attitude (Point)	6.99^{A} (0.11)	6.59^{A} (0.09)	7.00^{A} (0.17)	6.83^{A} (0.17)	6.90^{A} (0.15)	6.73^{A} (0.16)	5.90^{B} (0.15)	7.37^{A} (0.19)
Self-evaluation (Point)	3.13^{A} (0.14)	3.39^{A} (0.17)	4.10^{A} (0.24)	2.37^{B} (0.25)	3.50^{AC} (0.21)	3.73^{A} (0.25)	3.33^{AB} (0.28)	2.53^{BC} (0.22)
Mental stress								
Normalized LF (-)	2.46^{A} (0.30)	2.10^{A} (0.28)	2.93^{A} (0.87)	2.43^{A} (0.40)	1.87^{A} (0.25)	1.80^{A} (0.33)	2.54^{A} (0.56)	2.11^{A} (0.36)
Normalized HF (-)	3.48^{A} (0.60)	1.09^{A} (0.20)	2.15^{A} (0.75)	2.41^{A} (0.64)	1.92^{A} (0.70)	2.94^{A} (1.24)	2.40^{A} (0.71)	1.88^{A} (0.62)
LF/HF ratio (-)	2.52^{A} (0.43)	4.08^{A} (0.28)	3.08^{A} (0.97)	3.64^{A} (0.74)	2.85^{A} (0.69)	2.50^{A} (0.68)	4.13^{A} (1.52)	3.62^{A} (1.00)
Perceived stress (Point)	4.83^{A} (0.16)	5.42^{B} (0.13)	4.00^{A} (0.27)	6.53^{B} (0.12)	4.70^{A} (0.22)	4.60^{A} (0.16)	4.90^{A} (0.20)	6.03^{B} (0.18)

VR: verbal reports; SD: spatial disorientation; LF: low frequency; HF: high frequency; SD1: Somatogravic Illusion; SD2: Coriolis; SD3: Leans; SD4: False Horizon; SD5: Graveyard Spin; SD6: Black Hole Illusion. Values indicate mean (standard error). Letters represent statistical significance.

For the main effect of VR execution, flight performance differed significantly for altitude and airspeed on the instructor evaluation scales, while mental stress differed slightly for HF and LF/HF ratio of the HRV measures and significantly for perceived distress scores. The VR execution group had increased accuracy in maintaining altitude and airspeed during recovery from a disorientating scenario. This might be because pilots performing VR tend to focus on managing flight specifications, thereby compensating for the spatial ability reduced by SD. In the case of self-evaluation, there was no difference in the effectiveness of VR execution, which needs to be studied in the future. In addition, the VR execution group was likely to have higher HF and lower LF/HF ratio than the non-VR execution group. The results suggest that the task of maintaining flight specifications in the SD scenarios was likely to cause significant mental stress and changes in HRV to pilots and that VR execution tended to reduce mental stress in an SD situation by activating the PNS.[1] This is consistent with the results of previous studies that acute stress increases LF/HF and decreases HF, suggesting activation of the SNS as well as withdrawal of PNS activity under stress.[23] The perceived distress scores of the VR execution group were lower than those of the non-VR execution group, which may be attributed to the characteristics of verbal working memory. Echoic memory is known to disappear more slowly than iconic memory and to be more easily remembered when provided for a short time.[21] For example, Wickens et al.[38] found that pilots could learn better when navigational information was presented in auditory rather than visual paths in high workload situations. Therefore, when pilots efficiently perform VR in an SD situation, they can benefit from the effect of redundancy gain by checking the information verbally and visually. Thus, redundancy gain can enhance attention allocation and the working memory of a pilot's instrument information, and eventually reduce perceived mental stress.

For the main effect of SD types, flight performance differed significantly for all instructor evaluation scores and self-evaluation scores, while mental stress differed significantly only for perceived distress scores. Coriolis scored the lowest in the self-evaluations and Graveyard Spin the lowest in all instructor evaluations. The reason for the discrepancy between the instructor and self-evaluation results is that Coriolis is likely to have a greater impact on self-evaluation, as it occurs when a pilot moves his or her head directly to check for friendly/hostile aircraft or the surrounding terrain during a turn. However, since this type of SD can allow a pilot to quickly correct flight specifications, the impact on the instructor evaluation would be relatively low. Furthermore, Graveyard Spin tends to have a weaker effect on self-evaluation because it occurs when a pilot is not aware of the adaptation of his or her vestibular senses by an aircraft's continuous turning. Therefore, this type of SD could negatively affect the instructor evaluation because a pilot cannot properly manipulate flight specifications. These findings can help in the development of effective training plans and response procedures for pilots based on the characteristics of each SD. In the case of HRV measures, there was no difference in the effect of SD types, which means that the difference in mental stress by SD types was likely not to be significant. On the other hand, the perceived distress score according to type of SD was similar to that for the self-evaluation, meaning a high correlation between flight performance and pilots' judged mental stress.

The simulator sickness scores measured using the SSQ were not significant in terms of VR execution and the total severity scores of the VR execution and non-VR execution group differed by less than 5 points. Simulator sickness is a form of motion sickness that results from a mismatch between the simulated visual motion and sense of movement detected by the vestibular organ. Therefore, in this study, the effects of simulator sickness might have confounded the impact of VR execution and type of SD on flight performance and mental stress. However, the simulator sickness scores of the VR execution group (total severity score: 4.99 $\pm$ 0.87) and non-VR execution group (total severity score: 4.74 $\pm$ 0.85) were both negligible according to the SSQ scoring category criteria.[30] These results indicate that the effects of simulator sickness on the measurement of flight performance and mental stress under an SD situation were insignificant.

This study has limitations in terms of pilot recruitment, the experimental environment, and flight performance measure. The study was conducted with 30 male Air Force fighter pilots in their 20s and 30s. However, additional analyses, including factors such as gender, age, type, and flight qualification, are required to validate the results for flight performance and mental stress according to VR execution and type of SD. In addition, although the experiment was conducted using an SD simulator, it was difficult to accurately analyze physiological and psychological responses in the experimental environment because of differences from the actual flight environment. Lastly, though the evaluation score given by one instructor pilot was used as an objective variable of flight performance, it is necessary to extract and analyze flight specification data from the simulator in order to more objectively measure flight performance of pilots. Therefore, consideration of specific pilot groups, an actual SD situation, and simulator data analysis in future studies will further contribute to developing strategies to effectively understand and overcome SD through VR.

ACKNOWLEDGMENTS

The authors wish to thank the Aerospace Medical Training Centre of the Republic of Korea Air Force for supporting our research by allowing the use of their spatial disorientation training facilities.

Financial Disclosure Statement: The authors have no competing interests to declare.

Authors and affiliations: Yohan Kang, B.S., M.S., Air Combat Command, Republic of Korea Air Force, and Myung Hwan Yun, M.S., Ph.D., and Sungho Kim, B.S., M.S., Department of Industrial Engineering, Seoul National University, Seoul, Republic of Korea.

REFERENCES

1. Castaldo R, Melillo P, Bracale U, Caserta M, Triassi M, Pecchia L. Acute mental stress assessment via short term HRV analysis in healthy adults: a systematic review with meta-analysis. Biomed Signal Process Control. 2015; 18:370–377.
2. Davis JR, Johnson R, Stepanek J, Fogarty JA. Fundamentals of aerospace medicine, 4th ed. Philadelphia (PA, USA): Lippincott Williams & Wilkins; 2008.
3. Ericsson KA, Simon HA. Verbal reports as data. Psychol Rev. 1980; 87(3):215–251.
4. Gibb R, Ercoline B, Scharff L. Spatial disorientation: decades of pilot fatalities. Aviat Space Environ Med. 2011; 82(7):717–724.
5. Gilliland K, Schlegel RE. Tactile stimulation of the human head for information display. Hum Factors. 1994; 36(4):700–717.
6. Gresty MA, Golding JF, Le H, Nightingale K. Cognitive impairment by spatial disorientation. Aviat Space Environ Med. 2008; 79(2):105–111.
7. Gresty MA, Waters S, Bray A, Bunday K, Golding JF. Impairment of spatial cognitive function with preservation of verbal performance during spatial disorientation. Curr Biol. 2003; 13(21):R829–R830.
8. Heathers JA. Everything Hertz: methodological issues in short-term frequency-domain HRV. Front Physiol. 2014; 5:177.
9. Horng CT, Liu CC, Kuo DI, Shieh PC, Wu YC, et al. Changes in visual function during the Coriolis illusion. Aviat Space Environ Med. 2009; 80(4):360–363.
10. Kallus KW, Tropper K, Boucsein W. The importance of motion cues in spatial disorientation training for VFR-pilots. Int J Aviat Psychol. 2011; 21(2):135–152.
11. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int J Aviat Psychol. 1993; 3(3):203–220.
12. Kim HG, Cheon EJ, Bai DS, Lee YH, Koo BH. Stress and heart rate variability: a meta-analysis and review of the literature. Psychiatry Investig. 2018; 15(3):235–245.
13. Kuusela H, Paul P. A comparison of concurrent and retrospective verbal protocol analysis. Am J Psychol. 2000; 113(3):387–404.
14. Lewkowicz R, Fudali-Czyż A, Bałaj B, Francuz P. Change detection flicker task effects on simulator-induced spatial disorientation events. Aerosp Med Hum Perform. 2018; 89(10):863–872.
15. Lewkowicz R, Stróżak P, Bałaj B, Francuz P, Augustynowicz P. Selective auditory attention and spatial disorientation cues effect on flight performance. Aerosp Med Hum Perform. 2018; 89(11):976–984.
16. Li K, Rüdiger H, Ziemssen T. Spectral analysis of heart rate variability: time window matters. Front Neurol. 2019; 10:545.
17. Lyons TJ, Gillingham KK, Teas DC, Ercoline WR, Oakley C. The effects of acoustic orientation cues on instrument flight performance in a flight simulator. Aviat Space Environ Med. 1990; 61(8):699–706.
18. Malcolm R. Pilot disorientation and the use of a peripheral vision display. Aviat Space Environ Med. 1984; 55(3):231–238.
19. Mathews RSJ, Previc F, Bunting A. USAF spatial disorientation survey. Paper presented at the Research and Technology Organization and Human Factors and Medicine Symposium on Spatial Disorientation in Military Vehicles: Causes, Consequences and Cures; 2002 April; La Coruna, Spain. Neuilly-sur-Seine (France): NATO RTO; 2002. Report No.: HFM-085, RTO-MP-086.
20. Mohammadi A, Emamgoli A, Shirinkalam M, Meftahi GH, Yagoobi K, Hatef B. The persistent effect of acute psychosocial stress on heart rate variability. Egypt Heart J. 2019; 71(1):18.
21. Nilsson LG, Ohlsson K, Rönnberg J. Capacity differences in processing and storage of auditory and visual input. In: Dornick S, editor. Attention and performance VI. Hillsdale (NJ, USA): Lawrence Erlbaum Associates; 1977.
22. Pagani M, Lombardi F, Guzzetti S, Rimoldi O, Furlan R, et al. Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and conscious dog. Circ Res. 1986; 59(2):178–193.
23. Pagani M, Montano N, Porta A, Malliani A, Abboud FM, et al. Relationship between spectral components of cardiovascular variabilities and direct measures of muscle sympathetic nerve activity in humans. Circulation. 1997; 95(6):1441–1448.
24. Pan J, Tompkins WJ. A real-time QRS detection algorithm. IEEE Trans Biomed Eng. 1985; 32(3):230–236.
25. Previc FH, Ercoline WR, editors. Spatial disorientation in aviation, 1st ed. Reston (VA): AIAA; 2004.
26. Purtee MD, Gluck KA, Krusmark MA, Kotte SA, Lefebvre AT. Verbal protocol analysis for validation of UAV operator model. Proceedings of the 25th Interservice/Industry Training, Simulation, and Education Conference. Orlando (FL, USA): National Defense Industrial Association; 2003:1741–1750.
27. Quintana DS, Heathers JA. Considerations in the assessment of heart rate variability in biobehavioral research. Front Psychol. 2014; 5:805.
28. Rupert AH. Tactile situation awareness system: proprioceptive prostheses for sensory deficiencies. Aviat Space Environ Med. 2000; 71(9, Suppl.):A92–A99.
29. Schubert C, Lambertz M, Nelesen RA, Bardwell W, Choi JB, Dimsdale JE. Effects of stress on heart rate complexity—a comparison between short-term and chronic stress. Biol Psychol. 2009; 80(3):325–332.
30. Stanney KM, Kennedy RS, Drexler JM. Cybersickness is not simulator sickness. Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting; Albuquerque, NM. Santa Monica (CA, USA): HFES; 1997; 41(2):1138–1142.
31. Stott JR, Benson AJ. Spatial orientation and disorientation in flight. In: Gradwell DP, Rainford DJ, editors. Ernsting's aviation and space medicine, 5th ed. Boca Raton (FL, USA): CRC Press; 2016.
32. Stróżak P, Francuz P, Lewkowicz R, Augustynowicz P, Fudali-Czyż A, et al. Selective attention and working memory under spatial disorientation in a flight simulator. The International Journal of Aerospace Psychology. 2018; 28(1–2):31–45.
33. Sullivan C, Blackman HS. Insights into pilot situation awareness using verbal protocol analysis. Proceedings of the Human Factors and Ergonomics Society 35th Annual Meeting; San Francisco, CA, USA. Santa Monica (CA, USA): HFES; 1991; 35(2):57–61.

34. Tamura A, Iwamoto T, Ozaki H, Kimura M, Tsujimoto Y, Wada Y. Wrist-worn electrodermal activity as a novel neurophysiological biomarker of autonomic symptoms in spatial disorientation. Front Neurol. 2018; 9:1056.
35. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Circulation. 1996; 93(5):1043–1065.
36. Tropper K, Kallus KW, Boucsein W. Psychophysiological evaluation of an antidisorientation training for visual flight rules pilots in a moving base simulator. Int J Aviat Psychol. 2009; 19(3):270–286.
37. Webb CM, Estrada A III, Kelley AM. The effects of spatial disorientation on cognitive processing. Int J Aviat Psychol. 2012; 22(3):224–241.
38. Wickens CD, Sandry DL, Vidulich M. Compatibility and resource competition between modalities of input, central processing, and output. Hum Factors. 1983; 25(2):227–248.
39. Wilson GF. An analysis of mental workload in pilots during flight using multiple psychophysiological measures. Int J Aviat Psychol. 2002; 12(1):3–18.
40. Yaribeygi H, Panahi Y, Sahraei H, Johnston TP, Sahebkar A. The impact of stress on body function: a review. EXCLI J. 2017; 16:1057–1072.