



# Effects of Temporal Characteristics on Pilots Perceiving Audiovisual Warning Signals Under Different Perceptual Loads

Xing Peng<sup>1\*</sup>, Hao Jiang<sup>1</sup>, Jiazhong Yang<sup>1</sup>, Rong Shi<sup>1</sup>, Junyi Feng<sup>2</sup> and Yaowei Liang<sup>1,3</sup>

<sup>1</sup> Institute of Aviation Human Factors and Cognitive Neuroscience, College of Flight Technology, Civil Aviation Flight University of China, Guanghan, China, <sup>2</sup> Technical Support Center, Operation Control Department, Beijing Capital Airlines, Beijing, China, <sup>3</sup> Flying Department of Southwest Branch, Air China Limited, Chengdu, China

Our research aimed to investigate the effectiveness of auditory, visual, and audiovisual warning signals for capturing the attention of the pilot, and how stimulus onset asynchronies (SOA) in audiovisual stimuli affect pilots perceiving the bimodal warning signals under different perceptual load conditions. In experiment 1 of the low perceptual load condition, participants discriminated the location (right vs. left) of visual targets preceded by five different types of warning signals. In experiment 2 of high perceptual load, participants completed the location task identical to a low load condition and a digit detection task in a rapid serial visual presentation (RSVP) stream. The main effect of warning signals in two experiments showed that visual and auditory cues presented simultaneously (AV) could effectively and efficiently arouse the attention of the pilots in high and low load conditions. Specifically, auditory (A), AV, and visual preceding auditory stimulus by 100 ms (VA100) increased the spatial orientation to a valid position in low load conditions. With the increase in visual perceptual load, auditory preceding the visual stimulus by 100 ms (AV100) and A warning signals had stronger spatial orientation. The results are expected to theoretically support the optimization design of the cockpit display interface, contributing to immediate flight crew awareness.

**Keywords:** temporal characteristics, pilots, warning signals, audiovisual integration, perceptual load

## OPEN ACCESS

### Edited by:

Giovanni Anobile,  
University of Florence, Italy

### Reviewed by:

Paolo Antonino Grasso,  
University of Florence, Italy  
Irene Togoli,  
International School for Advanced  
Studies (SISSA), Italy

### \*Correspondence:

Xing Peng  
fypx3688@163.com

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 03 November 2021

**Accepted:** 10 January 2022

**Published:** 10 February 2022

### Citation:

Peng X, Jiang H, Yang J, Shi R,  
Feng J and Liang Y (2022) Effects  
of Temporal Characteristics on Pilots  
Perceiving Audiovisual Warning  
Signals Under Different Perceptual  
Loads. *Front. Psychol.* 13:808150.  
doi: 10.3389/fpsyg.2022.808150

## INTRODUCTION

Pilots make use of information from the inner cockpit and outside the aircraft. About 80% of the information comes from the visual modality. Thus, pilots often suffer from visual signal overload, which may lead to flight accidents or dangerous situations. Audible information has been employed in the design of the cockpit to cope with visual overload, especially in the domain of warning signals. For example, Federal Aviation Regulation §25.1322 has been revised by Amendment No.131. It stipulates that warnings and caution alerts in the cockpit should provide prompt cues through no less than two distinct senses.

Existing studies have shown that the human brain can integrate visual and auditory signals into a unified, coherent, and meaningful perceptual content. This process can be called audiovisual integration (AVI). As indicated from the behavioral performances of the AVI, bimodal signals elicit faster and more accurate responses than unimodal (visual, auditory) signals alone

(Hershenson, 1962; Miller, 1982; Frassinetti et al., 2002). Recently, the advantages of AVI have achieved an extensive application for designing warning signals, e.g., ground proximity warning systems and take-over request systems in autopilot. These warnings consist of multiple sensory signals, resulting in immediate awareness of the driver (Fitch et al., 2011; Prewett et al., 2011; Haas and Erp, 2014). This has suggested that bimodal signals capture spatial attention more effectively than that of unimodal signals (Santangelo et al., 2008).

The spatial cueing task was widely adopted to determine the ability of bimodal stimuli to capture attention (Posner, 1980). This task used an abrupt peripheral onset as exogenous cues to be presented on the left or right peripheral position. The target showed up at the identical (valid condition) or at the opposed location (invalid condition) to the previous cue. The reaction time to the target under the identical position was significantly faster than that under the opposed position, which is termed as the “spatial cueing effect.” Using this task, studies showed a spatial cueing effect of unimodal and bimodal cues with a comparable magnitude under the no-load condition. Also, only bimodal cues trigger a significant spatial cueing performance in the high perceptual load condition of rapid serial visual presentation (RSVP) task (Santangelo and Spence, 2007; Santangelo et al., 2008). Similar results were achieved in other high load tasks, e.g., visual search task and temporal order judgment task (Matusz and Eimer, 2011; Barrett and Katrin, 2012). These studies demonstrated that bimodal cues could elicit a larger spatial cueing effect than unimodal cues in high-load conditions. Importantly, the perceptual load modulates the spatial cueing effect elicited by bimodal stimuli. During the different flight phases (departure, cruise, and arrival), the perceptual load faced by pilots was changing. Thus, the spatial cueing effect of auditory, visual, and audiovisual (bimodal) warning signals under different perceptual load conditions should further be examined.

A vital factor that facilitates AVI is close temporal proximity, known as the temporal rule. The temporal binding window (TBW) reflects the range of stimulus onset asynchronies (SOA) with audiovisual stimuli integrated for forming a single percept. Studies found that when audiovisual stimuli are presented at approximately the same time, the effect of AVI is pronounced (Stevenson et al., 2012). However, some found that audiovisual stimuli are presented asynchronously, for instance, when SOA was 40 ms (Lewald and Getzmann, 2011), 100 ms (Van de Par and Kohlrausch, 2000), and even 250 ms (McDonald et al., 2000), produced the significant AVI effect. For integrating of the simple stimuli (e.g., visual flashes and auditory beeps), the TBW is approximately 20–80 ms (Keetels and Vroomen, 2005). In terms of more complex and meaningful stimuli (e.g., audiovisual speech), the TBW increases to a few hundred milliseconds (Vroomen and Keetels, 2010). Several studies found the differential width of the TBW based on different groups. For instance, as compared with controls, people with schizophrenia (Stevenson et al., 2017) and autism spectrum disorders (Stevenson et al., 2014) have wider TBWs. Musicians relative to non-musicians showed significantly narrower TBWs for music and sine-wave speech (Lee and Noppeney, 2014). As

impacted by the differential life-related perceptual experiences, a simple sensory training can change the TBW and may impact multisensory processing (Donohue et al., 2010). It is hypothesized that whether the prior experiences and life history of the special group of long-term training pilots might affect their multisensory processing. Besides, the width of the TBW was also strongly dependent on task difficulty related to perceptual load and tasks with higher perceptual load produced wider TBWs (Stevenson and Wallace, 2013).

In summary, perceptual load and SOA in audiovisual stimuli affected the AVI of bimodal warning signals. The present study examines how temporal characteristics affect pilots perceiving bimodal warning signals under different perceptual loads. We used the spatial cueing task and manipulated the types of warning signal and cue validities to investigate the spatial cueing effect of auditory, visual, and audiovisual (bimodal) warning signals, and how temporal characteristics affect the pilots who are perceiving a bimodal warning under different perceptual loads. Participants were required to conduct the discrimination of the position (right vs. left) of visual targets preceded by these different warning signals in experiment 1 (low perceptual load). In experiment 2 (high perceptual load), subjects were to perform an additional demanding central RSVP task, in which they were required to identify digit numbers in a series of presented letters. We hypothesized that: (1) bimodal warning signals were likely to have significantly greater effectiveness than unimodal warning signals to capture the attention of a pilot and (2) the perceptual load modulates the effects of temporal characteristics for bimodal warning signals.

## EXPERIMENT 1

### Method

#### Participants

Twenty-four pilots (average age:  $22.08 \pm 1.06$  years; age scope: 20–24 years) were recruited from the Civil Aviation Flight University of China. All participants had normal or corrected-to-normal vision, were all right-handed, and did not have any neurological or psychiatric disorders history. They got aviation commercial licenses from the Civil Aviation Administration of China (CAAC) and have logged an average of more than 230 h of flight in simulators and real aircraft. This study was approved by the Ethics Committee of Civil Aviation Flight University of China and was performed in accordance with the approved guidelines and the Declaration of Helsinki. The written informed consent was offered by each subject engaged.

#### Apparatus and Materials

There existed five warning signals serving as the cue, presented at the left or right side of the screen to capture the attention of the participants. The visual (V) warning signal was a black rectangle (size:  $0.2^\circ \times 0.1^\circ$ ;  $4.5^\circ$  away from the center of the screen). The auditory (A) warning signal was pure tones of 1,000 Hz (65 dB, 100 ms, 10 ms rising and falling time) played through earphones. Unimodal V and A warning signals were presented for 100 ms. Moreover, visual and auditory signals were combined to generate

bimodal warning cues. The visual signal in the bimodal cues could be presented simultaneously with, 100 ms prior to, or 100 ms after the auditory signal. These three temporal conditions were labeled as AV, VA100, or AV100, respectively. Thus, five warning signal cue types were V, A, AV, VA100, and AV100, respectively. In the target screen, a visual target was presented, i.e., a red (RGB: 255, 0, 0; 27.5 cd/m<sup>2</sup>) solid circle (size: 1° × 1°) presented at the identical (valid cue) or opposite (invalid cue) location to the previous cue.

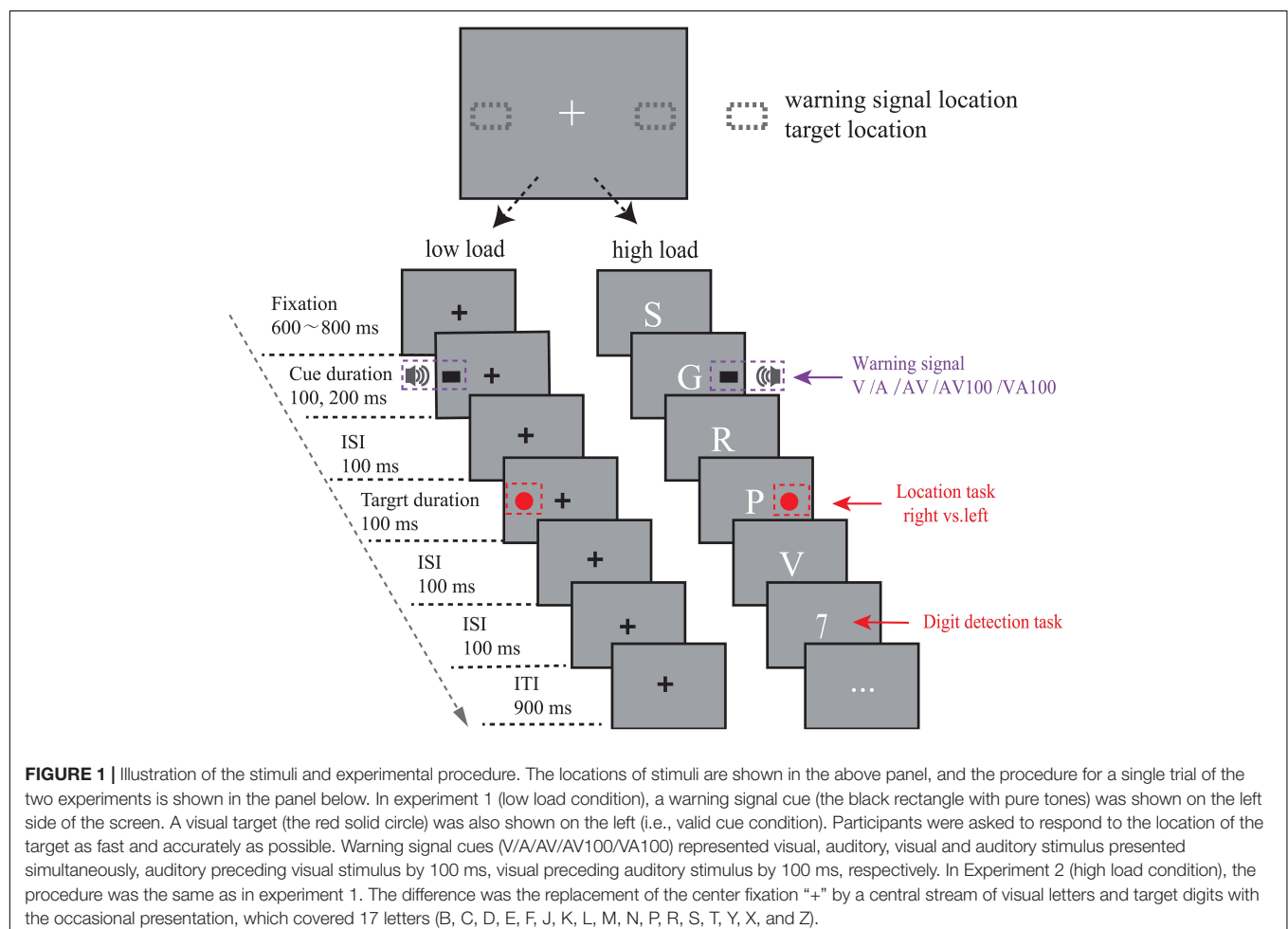
The experimental design was a 5 (warning signal cue type: V, A, AV, VA100, and AV100) × 2 (cue validity: 50% valid cue and 50% invalid cue) factorial design. As shown in **Figure 1**, the fixation stimulus had the presentation inside the display center and lasted for 600–800 ms. Then, one type of warning cues (V, A, AV, VA100, AV100) was randomly presented to left or right of the central fixation, lasting for 100 ms (V, A, AV) or 200 ms (VA100, AV100). After an interval time of 100 ms, a target stimulus (red solid circle) was randomly shown on the identical (valid cue location) or opposite (invalid cue location) sides of the warning signal, lasting for 100 ms duration. The next trial started 900 ms later.

Participants were tested in a quiet and dimly lit room. The stimuli were presented on a 24-in computer monitor with a

refresh rate of 60 Hz. The distance between the eyes of the participants and the computer monitor was about 60 cm. The subjects participating in the experiment were instructed to respond as soon and as accurately as possible. When the target showed up on the left side of the screen, the “←” key on the keyboard should be pressed. Otherwise, the “→” key should be pressed. The experiment covered a total of 1,080 trials, arranged in 6 blocks of 180 trials within each block. The experiment lasted for approximately 40 min.

## Data Analysis

Incorrect trials and trials with the reaction time below 100 ms or those over 1,000 ms were deleted and not calculated, because they were assumed to be the result of anticipation or not paying attention to the task, respectively. In addition to the reaction time and accuracy, we also calculated the inverse efficiency score (IES). This single index was calculated by combining RT and accuracy: RT/percentage of correct responses (Townsend and Ashby, 1983). Please note that the smaller the IES, the better the performance. The 5 × 2 repeated measures analysis of variance was used to analyze the accuracy, reaction time, and IES data with the factor warning signals (V, A, AV, VA100, and AV100) and cue validity (valid cue and invalid cue). The Bonferroni correction was used



for multiple comparisons correction. The effect size of the partial eta-squared ( $\eta_p^2$ ) was calculated for the ANOVA. All statistical levels were set to .05.

## Results

### Accuracy

**Figure 2A** illustrated the results: the main effect of warning signals is significant,  $F(4, 92) = 22.55, p = 0.001, \eta_p^2 = 0.50$ . The accuracy of A (97.4%) was significantly higher than that of AV (95.6%,  $p = 0.007$ ), AV100 (92.7%,  $p < 0.001$ ), and VA100 (94.8%,  $p = 0.009$ ); The accuracy of AV (95.6%) was significantly higher than AV100 (92.7%,  $p < 0.001$ ); The accuracy of V (97.2%) was significantly higher than AV (95.6%,  $p = 0.013$ ), AV100 (92.7%,  $p < 0.001$ ), and VA100 (94.8%,  $p = 0.019$ ). The accuracy of VA100 (94.8%) was significantly higher than AV100 (92.7%,  $p = 0.023$ ). The main effect of cue validity was significant,  $F(1, 23) = 55.24, p < 0.001, \eta_p^2 = 0.71$ . The accuracy of valid cues (96.6%) was significantly higher than the invalid cues (94.5%),  $p < 0.001$ . A significant interaction was identified between warning signals and cue validity,  $F(4, 92) = 15.77, p < 0.001, \eta_p^2 = 0.40$ . To be specific, accuracy to validly cued AV targets ( $M = 96.4\%$ ,  $SD = 4.95$ ) was significantly higher than to invalidly cued AV targets ( $M = 94.8\%$ ,  $SD = 6.12$ ),  $t(23) = 2.67, p = 0.014$ , and  $d = 0.28$ ; Accuracy to validly cued AV100 targets ( $M = 96\%$ ,  $SD = 5.32$ ) was significantly higher than to invalidly cued AV100 targets ( $M = 89.4\%$ ,  $SD = 7.48$ ),  $t(23) = 6.97$ , and  $p < 0.001$ , and  $d = 0.99$ ; Moreover, compared to invalidly cued VA100 targets ( $M = 93.9\%$ ,  $SD = 5.94$ ), accuracy to validly cued VA100 targets ( $M = 95.7\%$ ,  $SD = 5.79$ ) was significantly higher,  $t(23) = 6.97, p = 0.009$ , and  $d = 0.30$ . No difference in accuracy was identified between validly and invalidly cued A and V targets.

Three types of warning signals (AV, AV100, and VA100) can promote the detection of targets on the valid cue condition. Because the accuracy rates of the warning signals are above 95%, we further analyze the reaction time.

### Reaction Time

**Figure 2B** presents the results: the main effect of warning signals is significant,  $F(4, 92) = 30.90, p < 0.001, \eta_p^2 = 0.573$ . The reaction time of A (224 ms) was significantly faster than that of V (253 ms,  $p < 0.001$ ); The reaction time of AV (217 ms) was

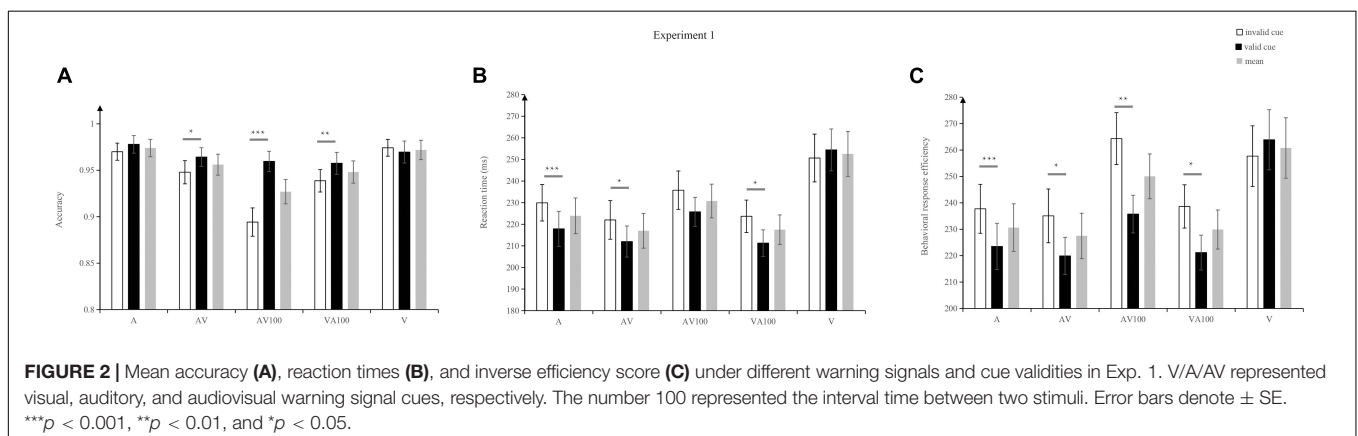
significantly faster than that of A (224 ms,  $p = 0.003$ ), AV100 (231 ms,  $p < 0.001$ ), and V (253 ms,  $p < 0.001$ ); The reaction time of AV100 (231 ms) was significantly faster than that of V (253 ms,  $p = 0.004$ ); The reaction time of VA100 (217 ms) was significantly faster than that of AV100 (231 ms,  $p = 0.001$ ) and V (253 ms,  $p < 0.001$ ). The main effect of cue validity was significant,  $F(1, 23) = 5.24, p = 0.032, \eta_p^2 = 0.19$ , the reaction time of valid cue (224 ms) was significantly faster than that of invalid cue (232 ms).

There was a significant interaction between warning signals and cue validity,  $F(4, 92) = 6.65, p < 0.001, \eta_p^2 = 0.22$ . To be specific, reaction time to validly cued AV targets ( $M = 212$  ms,  $SD = 35$  ms) was significantly faster than to invalidly cued AV targets ( $M = 222$  ms,  $SD = 44$  ms),  $t(23) = 2.28, p = 0.032$ , and  $d = 0.25$ ; Reaction time to validly cued VA100 targets ( $M = 211$  ms,  $SD = 30$  ms) was significantly faster than to invalidly cued VA100 targets ( $M = 224$  ms,  $SD = 37$  ms),  $t(23) = 2.78, p = 0.011$ , and  $d = 0.38$ ; Also, reaction time to validly cued A targets ( $M = 218$  ms,  $SD = 40$  ms) was significantly faster than to invalidly cued A targets ( $M = 230$  ms,  $SD = 41$  ms),  $t(23) = 4.51, p < 0.001$ , and  $d = 0.30$ ; No difference was identified in reaction time between validly and invalidly cued V and AV100 targets. From the results of the reaction time, A, AV and VA100 warning signals produced a significant exogenous cueing effect, demonstrating that the mentioned three types of warning signals are capable of capturing the pilot's attention effectively.

### Inverse Efficiency Score

**Figure 2C** presents the results: the main effect of warning signals is significant,  $F(4, 92) = 23.189, p < 0.001, \eta_p^2 = 0.502$ . The IES of AV (227) was significantly smaller than that of AV100 (250,  $p < 0.001$ ) and V (261,  $p < 0.001$ ); The IES of VA100 (230) was significantly smaller than that of AV100 (250,  $p < 0.001$ ) and V (261,  $p < 0.001$ ); The IES of A (231) was significantly smaller than that of AV100 (250,  $p < 0.001$ ) and V (261,  $p < 0.001$ ). The main effect of cue validity was significant,  $F(1, 23) = 11.23, p = 0.003, \eta_p^2 = 0.328$ . The IES of valid cue (232) was significantly smaller than that of invalid cue (247).

There was a significant interaction between warning signals and cue validity,  $F(4, 92) = 10.028, p < 0.001, \eta_p^2 = 0.304$ . To be specific, IES to validly cued AV targets ( $M = 220, SD = 34$ ) was





significantly smaller than to invalidly cued AV targets ( $M = 235$ ,  $SD = 50$ ),  $t(23) = 2.51$ ,  $p = 0.019$ , and  $d = 0.34$ ; The IES to validly cued VA100 targets ( $M = 221$ ,  $SD = 32$ ) was significantly smaller than to invalidly cued VA100 targets ( $M = 239$ ,  $SD = 40$ ),  $t(23) = 3.14$ ,  $p = 0.005$ , and  $d = 0.5$ ; Also, IES to validly cued A targets ( $M = 223$ ,  $SD = 43$ ) was significantly smaller than to invalidly cued A targets ( $M = 238$ ,  $SD = 45$ ),  $t(23) = 4.52$ ,  $p < 0.001$ , and  $d = 0.34$ ; IES to validly cued AV100 targets ( $M = 236$ ,  $SD = 35$ ) were significantly smaller than to invalidly cued AV100 targets ( $M = 264$ ,  $SD = 48$ ),  $t(23) = 3.67$ ,  $p = 0.001$ , and  $d = 0.65$ ; No difference was identified in IES between validly and invalidly cued V targets.

The IES results showed AV, VA100, and A warning signals had better IES regardless of cue validity. Considering the cue validity condition, A, AV, AV100, and VA100 produced better IES to valid target stimuli.

## EXPERIMENT 2

The visual load faced by pilots will change in different flight stages. For example, during the taxiing and cruise phase, there involve low visual loads. During the take-off, climb and final approach phase, the pilot needs to pay more attention to monitor the flight parameters of the cockpit, which involves high visual loads. With the visual load increased, will the TBW between the V and A signals affect the AVI? For the mentioned reason, experiment 2 employed the RSVP stream paradigm to increase the visual load, examining which warning signal cues are more effective to capture the spatial attention and how temporal characteristics affect the bimodal warning signals under a high perceptual load.

## Methods

### Participants

Twenty-four pilots (age range: 20–24 years; mean age:  $22.09 \pm 1.06$  years) were recruited from the Civil Aviation Flight University of China. Two participants were excluded from further analyses because their accuracy rates were below 90% in the digit detection task of experiment 2. The written informed consent was provided by all subjects participating in the experiment. All pilot participants had normal or corrected-to-normal vision, were all right-handed, and did not have any neurological or psychiatric disorders history. They got aviation commercial licenses from the CAAC and have logged an average of more than 230 h of flight in simulators and real aircraft.

### Apparatus and Materials

In the high-load condition, the center fixation in experiment 1 was replaced by a central stream of visual letters for target digits with occasional presentation, which is named RSVP stream. The set of distractors within the RSVP covered 17 letters (B, C, D, E, F, J, K, L, M, N, P, R, S, T, Y, X, and Z), which does not require a response, and the set of targets covered six digits (2, 3, 4, 5, 6, and 9;  $1^\circ \times 1^\circ$ ). The five distractor letters in the stream were selected randomly before each trial with the

sole restriction that no distractor was repeated within a given stream. Each letter was presented for 100 ms. The warning signal was presented as a cue in the third-stream position, and equiprobably on either side of fixation, while the target digit appeared equiprobably in the sixth or seventh stream position. When presented, the red circle target of the spatial cueing task showed up in the fifth position after the warning cue. The target digit had the presentation on 20% of the trials, whereas the red circle peripheral target had the presentation on the remaining 80% of the trials. Subjects participating in the experiment should execute the location task (requiring a discrimination response), which was identical to experiment 1, and the digit detection task (press the “↓” key after seeing the target digit in the screen center). The experiment lasted for approximately 40 min. There were 1,080 trials in total, which consisted of 864 trials of the location task and 216 trials of the digit detection task.

### Data Analysis

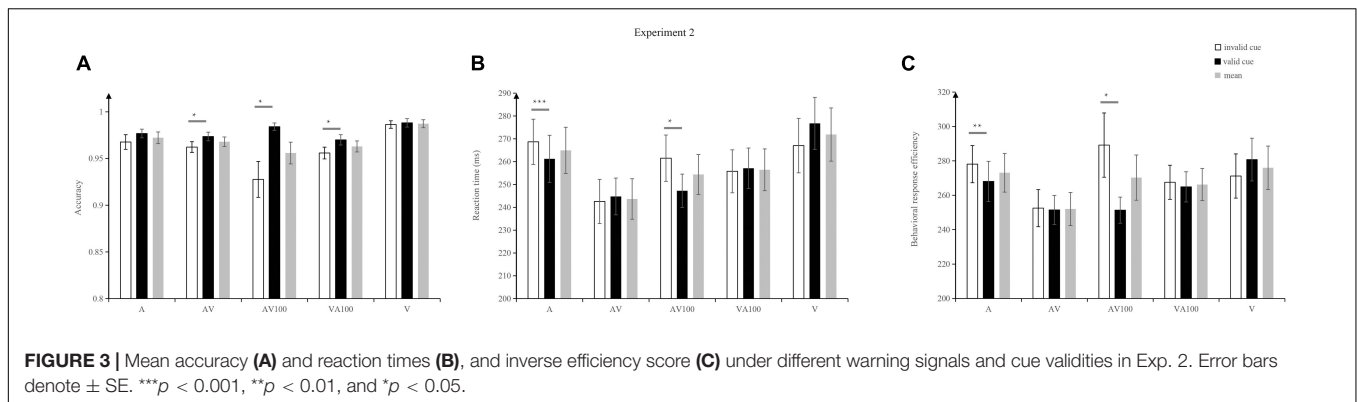
The standard of data elimination and data analysis method of accuracy and reaction time were the same as experiment 1. In experiment 2, pilots participating in the experiment should complete both the location and the digit detection tasks. Two participants were excluded from further analyses because their accuracy rates were below 90% in the digit detection task. The purpose of the digit task was to increase the visual load, so we did not analyze the data from this task. The following analysis was only for the location task. The Bonferroni correction was used for multiple comparisons correction. The effect size of the partial eta-squared ( $\eta_p^2$ ) was calculated for the ANOVA. All statistical levels were set to 0.05.

## Results

### Accuracy

The results are shown in **Figure 3A**: the main effect of warning signals was observed,  $F(4, 84) = 4.58$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.18$ . The accuracy of V (98.7%) was significantly higher than that of AV (96.8%,  $p = 0.02$ ) and VA100 (96.3%,  $p = 0.004$ ). The main effect of cue validity was significant,  $F(1, 21) = 13.19$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.39$ . The accuracy of valid cues (97.9%) was significantly higher than that of invalid cues (96%),  $p = 0.002$ .

There was a significant interaction between cue validity and warning signals,  $F(4, 84) = 4.12$ ,  $p = 0.004$ ,  $\eta_p^2 = 0.16$ . To be specific, the accuracy to validly cued AV targets ( $M = 97.4\%$ ,  $SD = 2.17$ ) were significantly higher than to invalidly cued AV targets ( $M = 96.2\%$ ,  $SD = 2.69$ ),  $t(21) = 2.15$ ,  $p = 0.04$ , and  $d = 0.49$ ; The accuracy to validly cued AV100 targets ( $M = 98.4\%$ ,  $SD = 1.9$ ) were significantly higher than to invalidly cued AV100 ( $M = 92.8\%$ ,  $SD = 10.23$ ) targets,  $t(21) = 2.64$ ,  $p = 0.015$ , and  $d = 0.59$ ; the accuracy to validly cued VA100 targets ( $M = 97\%$ ,  $SD = 2.6$ ) were significantly higher than to invalidly cued VA100 ( $M = 95.6\%$ ,  $SD = 2.96$ ) targets,  $t(21) = 2.29$ ,  $p = 0.014$ , and  $d = 0.5$ ; There was no difference in accuracy between different validly conditions on A and V targets. Three types of warning signals (AV, AV100, and VA100) can promote the detection of targets on the valid cue condition.



## Reaction Time

The results are presented in **Figure 3B**: a main effect of warning signals was observed,  $F(4, 84) = 21.38, p < 0.001, \eta_p^2 = 0.50$ . The reaction time of AV (244 ms) was significantly faster than that of V (272 ms,  $p < 0.001$ ), A (265 ms,  $p < 0.001$ ), AV100 (254 ms,  $p = 0.001$ ), and VA100 (256 ms,  $p < 0.001$ ). The reaction time of AV100 (254 ms) was significantly faster than that of V (272 ms,  $p = 0.01$ ). The reaction time of VA100 (256 ms) was significantly faster than that of A (265 ms,  $p = 0.035$ ) and V (272 ms,  $p = 0.004$ ). The main effect of cue validity was not significant,  $F(1, 21) = 0.44, p = 0.51$ .

There was a significant interaction between cue type and cue validity,  $F(4, 84) = 6.81, p < 0.001, \eta_p^2 = 0.25$ . To be specific, reaction time to validly cued A targets ( $M = 261$  ms,  $SD = 48$  ms) were significantly faster than to invalidly cued A targets ( $M = 269$  ms,  $SD = 46$  ms),  $t(21) = 4.34, p < 0.001$ , and  $d = 0.17$ ; Reaction time to validly cued AV100 targets ( $M = 247$  ms,  $SD = 34$  ms) significantly faster than to invalidly cued AV100 targets ( $M = 262$  ms,  $SD = 48$  ms),  $t(21) = 2.32, p = 0.03$ , and  $d = 0.35$ . There was no difference in reaction time between different validly conditions on AV, V, and VA100. From the results of the reaction time, A and AV100 warning signal exerted an significant spatial cueing effect.

## Inverse Efficiency Score

**Figure 3C** presents the results: the main effect of warning signals was significant,  $F(4, 84) = 4.543, p = 0.002, \eta_p^2 = 0.178$ . The IES of AV (252) was significantly smaller than that of V (266,  $p < 0.001$ ), A (273,  $p < 0.001$ ), and VA100 (276,  $p < 0.001$ ).

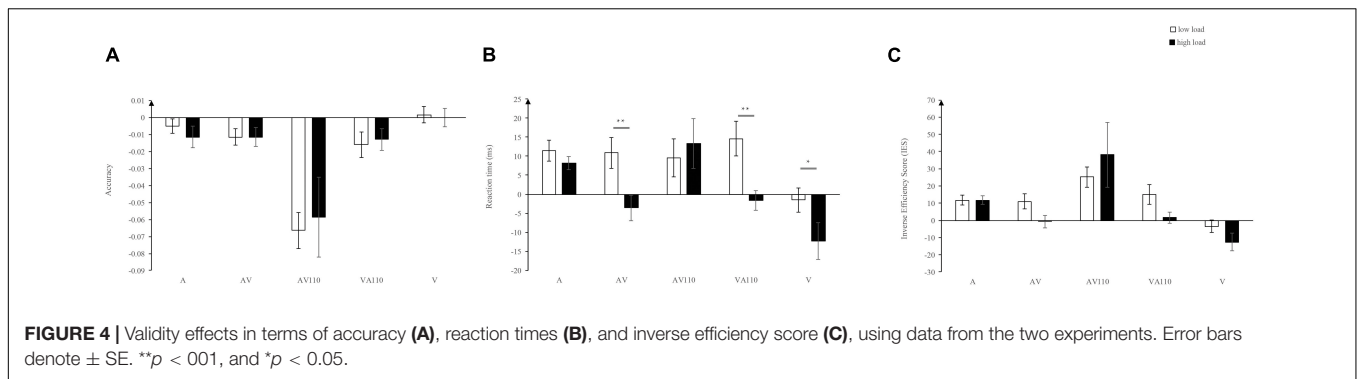
The main effect of cue validity was significant,  $F(1, 21) = 5.219, p = 0.033, \eta_p^2 = 0.199$ . The IES of valid cue (263) was significantly smaller than that of invalid cue (271). There was a significant interaction between warning signals and cue validity,  $F(4, 84) = 4.406, p = 0.003, \eta_p^2 = 0.173$ . To be specific, IES to validly cued A targets ( $M = 268, SD = 55$ ) were significantly smaller than to invalidly cued A targets ( $M = 278, SD = 51$ ),  $t(21) = 3.86, p = 0.001$ , and  $d = 0.19$ ; Also, IES to validly cued AV100 targets ( $M = 251, SD = 36$ ) were significantly smaller than to invalidly cued AV100 targets ( $M = 289, SD = 88$ ),  $t(21) = 2.21, p = 0.039$ , and  $d = 0.5$ ; No difference was identified in IES between validly and invalidly cued V, AV, and VA100 targets.

## Comparison Between Low and High Perceptual Loads

Finally, data from the two experiments were combined for further analysis. We focused on the validity effect, which refers to the difference of accuracy, reaction time, or IES between invalid-cue and valid-cue conditions. Then, we would like to investigate how the validity effect would be influenced by perceptual loads and types of warning signals. To this end, three 2 (perceptual load level) by 5 (types of warning signals) within-subjects ANOVAs were performed, for accuracy, reaction time, and IES, respectively. The results are shown in **Figures 4A–C**, respectively. When analyzing, data from the twenty shared subjects from the two experiments were used.

Results showed that, for accuracy, the main effect of warning signal cue type was significant,  $F(4, 76) = 11.05, p < 0.001, \eta_p^2 = 0.368$ . Validity effect (in terms of accuracy) for AV100 ( $-0.062$ ) was smaller than A ( $-0.008, p = 0.016$ ), AV ( $-0.012, p = 0.032$ ) and V ( $0.001, p = 0.005$ ). The main effect of perceptual load level was not significant,  $F(1, 19) = 0.1, p = 0.92$ , nor was the interaction between the two factors,  $F(4, 76) = 0.211, p = 0.931$ .

For reaction times, the main effect of warning signal cue type was significant,  $F(4, 76) = 8.366, p < 0.001, \eta_p^2 = 0.306$ . Validity effect of V ( $-6.88$  ms) was significantly smaller than A (9.84 ms,  $p = 0.001$ ), AV (3.65 ms,  $p = 0.006$ ), AV100 (11.44 ms,  $p = 0.03$ ), and VA100 (6.5 ms,  $p = 0.003$ ). The main effect of perceptual load level was significant,  $F(1, 19) = 8.03, p = 0.011, \eta_p^2 = 0.297$ . Validity effect (in terms of reaction times) in low load condition (9 ms) was significantly larger than that in high load condition (0.82 ms,  $p = 0.011$ ). Importantly, there was a significant interaction between cue type of warning signal and perceptual load level,  $F(4, 76) = 4.422, p = 0.003, \eta_p^2 = 0.189$ . Simple effect analysis revealed a significantly smaller validity effect in high load condition compared with low load condition for AV ( $p = 0.003$ ), VA100 ( $p = 0.007$ ), and V ( $p = 0.025$ ). In low load condition, validity effect of V ( $-1.48$  ms) was significantly smaller than A (11.43 ms,  $p = 0.03$ ), AV (10.89 ms,  $p = 0.015$ ), and AV100 (9.56 ms,  $p = 0.034$ ). In high load condition, validity effect of A (8.26 ms) was significantly larger than AV ( $-3.59$  ms,  $p = 0.014$ ), VA100 ( $-1.60$  ms,  $p = 0.34$ ), and V ( $-12.29$  ms,  $p = 0.003$ ). AV100 (13.31 ms) was significantly larger than V ( $-12.29$  ms,  $p = 0.037$ ).



For IES, it was similar with the accuracy. The main effect of warning signal cue type was significant,  $F(4, 76) = 7.722$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.289$ . Validity effect (in terms of IES) of V ( $-8.09$ ) was significantly smaller than A ( $11.76$ ,  $p = 0.001$ ), AV ( $5.12$ ,  $p = 0.002$ ), AV100 ( $31.64$ ,  $p = 0.03$ ), and VA100 ( $8.43$ ,  $p = 0.001$ ). The main effect of perceptual load level was not significant,  $F(1, 19) = 0.982$ ,  $p = 0.334$ , nor was the interaction between the two factors,  $F(4, 76) = 1.379$ ,  $p = 0.249$ .

## DISCUSSION

### The Effectiveness of Different Types of Warning Signals

This study tested several types of unimodal and multimodal warning signals to explore their effects in capturing the attention of the pilot. In doing so, reaction times, accuracy, and IES were analyzed. For the sake of clarity, we discussed the IES results as it provides an integration of the other two. The main effect of warning signals for IES in two experiments showed that AV, VA100, and A warning signals produced better processing advantages in low load conditions, AV warning cues produced a better processing advantage in high load conditions.

In low load condition, the IES of AV, VA100, and A was significantly smaller than that of V and AV100. In other words, AV, VA100, and A warning signals showed advantages in arousal. Firstly, AV warnings had a better task performance. This result is consistent with previous studies: when audiovisual stimuli are presented at approximately the same time, the effect of AVI is pronounced (Stevenson et al., 2012). Secondly, unimodal A cues can be significantly more effective than V alone in enhancing the performance of the pilot, even better than bimodal cues (AV100). Auditory displays are more effective than visual displays alone in improving user performance and effectively attracting the attention of the drivers (Edworthy and Hellier, 2006; Haas and Edworthy, 2006). Santangelo et al. (2006, 2008) investigated whether multisensory cues might be more effective in capturing the spatial attention of a person than unimodal cues. Their results showed the reaction time of A cues was the fastest, significantly faster than AV and V, which is partly consistent with ours. When pilots were required to accomplish a visual-spatial information processing, their visual attention resources were occupied, so the auditory-led warning signals

showed the dominance effect. Thirdly, some studies found that AVI is often the most pronounced when the stimuli from different modalities are presented asynchronously. Specifically, VA100 produced better IES than AV100 in the low load condition of our study. Previous studies showed bimodal signals in which visual preceding auditory stimulus by about 100 ms can produce the maximal response enhancement (Rowland et al., 2007; Perrault et al., 2011). This can be due to the differences in sensory conduction and processing time, and sensory input time may not be precisely aligned in time (King and Palmer, 1985). Auditory signals arrive in the brain about 30–50 ms earlier than visual signals, so the integration can probably occur when visual stimulation precedes auditory stimulation.

In the high load condition, the IES of AV was significantly smaller than that of V, A, and VA100. An AV warning had a better task performance than V or A warning alone regardless of other AV conditions (VA100). This result is consistent with our hypothesis and previous studies. Multimodal warnings can potentially enhance risk communication (van Erp et al., 2015). The bimodal stimulus presented synchronously elicited faster and more accurate responses than unimodal visual or auditory stimuli (Raab, 1962; Stein and Meredith, 1993). According to the temporal principle of audiovisual processing, the synchronous occurrence of the A and V stimuli of a multisensory stimulus can be a percept into a single and coherent perceptual representation (Meredith et al., 1987; Stein and Meredith, 1993; Senkowski et al., 2007). When V and A signals are presented simultaneously, like the AV warning signals in this study, they can elicit more rapid and more precise responses in contrast to the auditory or visual modality alone. This advantage can be termed as a “redundant signals effect” (Hershenson, 1962).

### Comparison Between Low and High Perceptual Loads

We compared the validity effect under different perceptual loads. There was a significant interaction between cue type of warning signal and perceptual load level in reaction time, but not in accuracy and IES. Specifically, there was a significant validity effect of reaction time (spatial cueing effect) for AV and VA100 in low load conditions, but no such effect in high load conditions. The perceptual load modulates the effects of temporal characteristics for bimodal warning signals. Previous studies

showed the spatial cueing effect was found in bimodal (audio-tactile and audiovisual) cues both in high and low perceptual load, but no spatial cueing effect was found in unimodal auditory, visual, or tactile cues under conditions of high load (Santangelo and Spence, 2007; Santangelo et al., 2008; Ho et al., 2009). In other words, only bimodal cues still captured attention under conditions of high perceptual load. However, in our study, AV cues showed no processing advantage under high load condition. On the contrary, AV, VA100, and A elicited faster responses to the valid target stimuli under low load conditions. The AV100 and A produced faster response time to the valid target stimuli under high load conditions. The additional auditory signal can effectively attract the attention of the pilot under the visual overload situation, especially in A and AV100 conditions. There are some reasons for this result: as visual load increases, auditory stimulus plays an important role in warning. The task does not consume all available perceptual resources in low-load conditions, and pilots can process additional stimuli. However, high-load tasks consume considerable attention resources, causing additional task-related information to be less processed. As indicated from perceptual load theory (Lavie, 2005), attention is a limited capacity system that proceeds automatically until such a capacity is reached. Reduced attention capacity for processing multisensory stimuli can impair multisensory temporal processing. This implies that the perceptual operations under high load conditions take more time to reach completion, suggesting additional perceptual resources can be used to process a multisensory temporal in low load conditions. Flying requires pilots to continuously monitor various instrument information in the cockpit and the environmental information outside the cockpit *via* predominantly visual input. Around 80% of the information necessary for the pilots is acquired through the visual system (Dehais et al., 2017). We speculate with the visual overload in the RSVP task; it can be argued that presenting warning signals *via* the auditory modality causes less workload of pilots compared to presenting the same information *via* the visual modality. Therefore, pilots benefited most from auditory-led warning signals (A and AV100).

It should be noted that there was no replication of Posner's classic facilitation effect under visual conditions. This is inconsistent with most previous studies, but there are a few studies having similar results with ours. For example, Van der Stoep et al. (2015) used an exogenous cue-target paradigm and modulated the modality of the target stimulus (A, V, and AV) to investigate the relationship between exogenous spatial attention and audiovisual integration. They found the validity effect (facilitating effect) appeared in the V and AV targets, but not for A targets in the location task (experiment 2). This result was like ours, when the cue and the target were from the same modality (our study: V cue\_ V target; Van der Stoep et al: A cue\_ A target), no validity effect was produced in this modality. Some researchers modulated the modality of the cue stimulus to investigate whether multisensory cues might be more effective in capturing the spatial attention of a person than unimodal cues. Results showed that the validity effect was found in AV and A cues, but no validity effect in unimodal V cues (Santangelo

et al., 2006). There are other possibilities for the lack of Posner's effect: we used headphones to present auditory signals, and this may affect the facilitation effect (Spence and Driver, 1997); Task requirements were different: we used the azimuth task (left vs. right), while other studies used the elevation task (up vs. down).

The limitation of the present study is that the experiments were performed in a laboratory environment, and the real flying experiment, which is necessary to avoid risks for the participants, was not performed. Also, the warning signals of the visual and auditory senses were mainly discussed, whereas other warning modalities (e.g., tactile sense) were not examined. In addition, the effect sizes in our study are relatively small. Combined with the results, although the bimodal warning signal has obvious advantages, the cockpit warning signal in actual flight should adopt a more rigorous grading system, and the integrated warning signal should be used in more dangerous situations. In some common signal designs, the use of unimodal warnings, especially audible warnings, can sufficiently elicit responses of the pilots. Moreover, according to "inattention blindness," people are likely to have insufficient visual/auditory stimuli awareness if stressing a distinct stimulus (Macdonald and Lavie, 2011; Murphy and Greene, 2017). Since flying commonly covers a significant visual perceptual load, the ability of the pilots to perceive the auditory warning signal may have an adverse effect. The combination of different senses should be used cautiously.

## CONCLUSION

In summary, temporal parameters were deployed to build integrated audiovisual warning signals. No matter in the case of low or high visual load, AV warning cues can effectively and efficiently arouse the attention of pilots. Specifically, A, AV, and VA100 warning cues increased spatial orientation to valid target stimuli in low load condition. With the increase in visual perceptual load, AV100, and A warning cues had a stronger spatial orientation to valid target stimuli. The results are expected to theoretically support the optimization design of the cockpit display interface, contributing to immediate flight crew awareness.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee at the Civil Aviation Flight University of China. The patients/participants provided their written informed consent to participate in this study.



## AUTHOR CONTRIBUTIONS

XP conceived the study, designed the experiments, and wrote the main manuscript text. HJ and RS performed the testing and data collection. JF and YL performed the statistical analysis and interpretation under the supervision of XP. JY reviewed the manuscript. All authors contributed to the manuscript revision, read, and approved the final version of the manuscript for submission.

## REFERENCES

- Barrett, D. J. K., and Katrin, K. (2012). Evidence for multisensory integration in the elicitation of prior entry by bimodal cues. *Exp. Brain Res.* 222, 11–20. doi: 10.1007/s00221-012-3191-8
- Dehais, F., Roy, R. N., Durantin, G., Gateau, T., and Callan, D. (2017). “EEG-engagement index and auditory alarm misperception: an inattentive deafness study in actual flight condition,” in *Paper presented at the International Conference on Applied Human Factors and Ergonomics* (New York, NY: AHFE). doi: 10.1007/978-3-319-60642-2\_21
- Donohue, S. E., Woldorff, M. G., and Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. *Attent. Percept. Psychophys.* 72:1120. doi: 10.3758/APP.72.4.1120
- Edworthy, J., and Hellier, E. (2006). “Complex nonverbal auditory signals and speech warnings,” in *Handbook of Warnings*, ed. M. S. Wogalter (Mahwah, NJ: Lawrence Erlbaum Associates), 199–220.
- Fitch, G. M., Hankey, J. M., Kleiner, B. M., and Dingus, T. A. (2011). Driver comprehension of multiple haptic seat alerts intended for use in an integrated collision avoidance system. *Transport. Res. Part F Traffic Psychol. Behav.* 14, 278–290.
- Frassinetti, F., Bolognini, N., and Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Exp. Brain Res.* 147, 332–343. doi: 10.1007/s00221-002-1262-y
- Haas, E., and Edworthy, J. (2006). “An introduction to auditory warnings and alarms,” in *Handbook of Warnings*, ed. M. S. Wogalter (Mahwah, NJ: Lawrence Erlbaum Associates), 189–198.
- Haas, E. C., and Erp, J. B. F. V. (2014). Multimodal warnings to enhance risk communication and safety. *Safety Sci.* 61, 29–35.
- Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *J. Exp. Psychol.* 63:289. doi: 10.1037/h0039516
- Ho, C., Santangelo, V., and Spence, C. (2009). Multisensory warning signals: when spatial correspondence matters. *Exp. Brain Res.* 195, 261–272. doi: 10.1007/s00221-009-1778-5
- Keetels, M., and Vroomen, J. (2005). The role of spatial disparity and hemifields in audio-visual temporal order judgments. *Exp. Brain Res.* 167, 635–640. doi: 10.1007/s00221-005-0067-1
- King, A. J., and Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurons in the guinea-pig superior colliculus. *Exp. Brain Res.* 60, 492–500. doi: 10.1007/BF00236934
- Lavie, N. (2005). Distracted and confused?: selective attention under load. *Trends Cognit. Sci.* 9:75.
- Lee, H., and Noppeney, U. (2014). Music expertise shapes audiovisual temporal integration windows for speech, sinewave speech, and music. *Front. Psychol.* 5:868. doi: 10.3389/fpsyg.2014.00868
- Lewald, J., and Getzmann, S. (2011). When and where of auditory spatial processing in cortex: a novel approach using electrotopography. *PLoS One* 6:e25146. doi: 10.1371/journal.pone.0025146
- Macdonald, J. S. P., and Lavie, N. (2011). Visual perceptual load induces inattentive deafness. *Springer Open Choice* 73, 1780–1789. doi: 10.3758/s13414-011-0144-4
- Matusz, P. J., and Eimer, M. (2011). Multisensory enhancement of attentional capture in visual search. *Psychonomic Bull. Rev.* 2011, 1–6. doi: 10.3758/s13423-011-0131-8

## FUNDING

This study was supported by the Youth Project of Humanities and Social Sciences Financed by Ministry of Education of China (Grant No. 21YJC190012), the General Project of Civil Aviation Flight University of China (Grant No. J2020-004), and the National Natural Science Foundation of China (Award number(s): 31871025) and the Key Laboratory of Flight Techniques and Flight Safety, CAAC (FZ2021ZZ02).

- Mcdonald, J. J., Teder-SaLejarvi, W. A., and Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature* 407, 906–908.
- Meredith, M. A., Nemitz, J. W., and Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J. Neurosci.* 7, 3215–3229. doi: 10.1523/JNEUROSCI.07-10-03215.1987
- Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognit. Psychol.* 14, 247–279. doi: 10.1016/0010-0285(82)90010-x
- Murphy, G., and Greene, C. M. (2017). The Elephant in the Road: Auditory Perceptual Load Affects Driver Perception and Awareness. *Appl. Cognit. Psychol.* 31, 258–263. doi: 10.1002/acp.3311
- Perrault, T. J., Stein, B. E., and Rowland, B. A. (2011). Non-stationarity in multisensory neurons in the superior colliculus. *Front. Psychol.* 2:144. doi: 10.3389/fpsyg.2011.00144
- Posner, M. I. (1980). Orienting of attention. *Quart. J. Exp. Psychol.* 32, 3–25.
- Prewett, M. S., Elliott, L. R., Walvoord, A. G., and Coovert, M. D. (2011). A Meta-Analysis of Vibrotactile and Visual Information Displays for Improving Task Performance. *IEEE Transact. Syst. Man Cybernet. Part C* 42, 123–132.
- Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Trans. N. Y. Acad. Sci.* 24, 574–590.
- Rowland, B. A., Quessy, S., Stanford, T. R., and Stein, B. E. (2007). Multisensory integration shortens physiological response latencies. *J. Neurosci.* 27, 5879–5884. doi: 10.1523/JNEUROSCI.4986-06.2007
- Santangelo, V., and Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *J. Exp. Psychol. Hum. Percept. Performance* 33:1311. doi: 10.1037/0096-1523.33.6.1311
- Santangelo, V., Van der Lubbe, R. H., Belardinelli, M. O., and Postma, A. (2006). Spatial attention triggered by unimodal, crossmodal, and bimodal exogenous cues: a comparison of reflexive orienting mechanisms. *Exp. Brain Res.* 173, 40–48. doi: 10.1007/s00221-006-0361-6
- Santangelo, V., Van der Lubbe, R. H., Belardinelli, M. O., and Postma, A. (2008). Multisensory integration affects ERP components elicited by exogenous cues. *Exp. Brain Res.* 185, 269–277. doi: 10.1007/s00221-007-1151-5
- Senkowski, D., Talsma, D., Grigutsch, M., Herrmann, C. S., and Woldorff, M. G. (2007). Good times for multisensory integration: effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsychologia* 45, 561–571. doi: 10.1016/j.neuropsychologia.2006.01.013
- Stein, B. E., and Meredith, M. A. (1993). *The Merging of the Senses*. Cambridge, MA: The MIT Press.
- Spence, C., and Driver, J. (1997). On measuring selective attention to an expected sensory modality. *Percept. Psychophys.* 59, 389–403. doi: 10.3758/bf03211906
- Stevenson, R. A., and Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Exp. Brain Res. Exp. Hirnforschung Experimentation Cérébrale* 227, 249–261. doi: 10.1007/s00221-013-3507-3
- Stevenson, R. A., Fister, J. K., Barnett, Z. P., Nidiffer, A. R., and Wallace, M. T. (2012). Interactions between the spatial and temporal stimulus factors that influence multisensory integration in human performance. *Exp. Brain Res.* 219, 121–137. doi: 10.1007/s00221-012-3072-1
- Stevenson, R. A., Park, S., Cochran, C., Mcintosh, L. G., Noel, J. P., Barense, M. D., et al. (2017). The associations between multisensory temporal processing on symptoms of schizophrenia. *Schizophr. Res.* 179, 97–103. doi: 10.1016/j.schres.2016.09.035

- Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., et al. (2014). Multisensory Temporal Integration in Autism Spectrum Disorders. *J. Neurosci.* 34, 691–697. doi: 10.1523/JNEUROSCI.3615-13.2014
- Townsend, J. T., and Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.
- Van de Par, S., and Kohlrausch, A. (2000). Sensitivity to auditory-visual asynchrony and to jitter in auditory-visual timing. *Proc. SPIE Int. Soc. Optical Engine.* 3959, 234–242. doi: 10.1117/12.387160
- Van der Stoep, N., Van der Stigchel, S., and Nijboer, T. C. (2015). Exogenous spatial attention decreases audiovisual integration. *Attent. Percept. Psychophys.* 77, 464–482.
- van Erp, J. B., Toet, A., and Janssen, J. B. (2015). Uni-, bi- and tri-modal warning signals: effects of temporal parameters and sensory modality on perceived urgency. *Saf. Sci.* 72, 1–8. doi: 10.1016/j.ssci.2014.07.022
- Vroomen, J., and Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attent. Percept. Psychophys.* 72, 871–884. doi: 10.3758/APP.72.4.871

**Conflict of Interest:** YL was employed by the company Air China Limited. JF was employed by the company Beijing Capital Airlines.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Peng, Jiang, Yang, Shi, Feng and Liang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.