To link to this article: DOI: 10.1177/0018720813510735
URL: http://dx.doi.org/10.1177/0018720813510735

To cite this version: Dehais, Frédéric and Causse, Mickael and Vachon, François and Regis, Nicolas and Menant, Eric and Tremblay, Sébastien *Failure to Detect Critical Auditory Alerts in the Cockpit: Evidence for Inattentional Deafness.* (2014) Human Factors: The Journal of the Human Factors and Ergonomics Society, vol. 56 (n° 4). pp. 631-644. ISSN 0018-7208

Failure to Detect Critical Auditory Alerts in the Cockpit: Evidence for Inattentional Deafness

Frédéric Dehais[1], Mickaël Causse[1], François Vachon[2], Nicolas Régis[1], Eric Menant[1], Sébastien Tremblay[2]

[1]*DMIA, ISAE, Université de Toulouse, 10 av. E. Belin, 31055 Toulouse Cedex 4, France*

[2]*École de Psychologie, Université Laval, 2325, rue des Bibliothèques, Québec, Canada G1V 0A6*

Address all correspondence to Pr Frédéric Dehais, Institut Supérieur de l'Aéronautique et de l'Espace, Centre aéronautique et spatial, 10 avenue Édouard Belin BP 54032 - 31055 TOULOUSE Cedex 4

Précis: A flight-simulator experiment provides evidence of inattentional deafness to auditory landing-gear alarms in a bad weather scenario (windshear). The present study further reveals that the pre-exposure to an auditory alarm appears to protect against misperception of the same alarm.

ABSTRACT

Objective: The aim of this study was to test whether inattentional deafness to critical alarms would be observed in a simulated cockpit. **Background**: The inability of pilots to detect unexpected changes in their auditory environment (e.g., alarms) is a major safety problem in aeronautics. In aviation, the lack of response to alarms is usually not attributed to attentional limitations, but rather to pilots choosing to ignore such warnings due to decision biases, hearing issues, or conscious risk taking. **Method:** Twenty-eight general aviation pilots performed two landings in a flight simulator. In one scenario an auditory alert was triggered alone, whereas in the other the auditory alert occurred while the pilots dealt with a critical windshear. **Results:** In the windshear scenario, 11 pilots (39.3%) did not report nor react appropriately to the alarm whereas all the pilots perceived the auditory warning in the no-windshear scenario. Also, of those pilots who were first exposed to the no-windshear scenario and detected the alarm, only three suffered from inattentional deafness in the subsequent windshear scenario. **Conclusion:** These findings establish inattentional deafness as a cognitive phenomenon that is critical for air safety. Pre-exposure to a critical event triggering an auditory alarm can enhance alarm detection when a similar event is encountered subsequently. **Application:** Case-based learning is a solution to mitigate auditory alarm misperception.

Keywords: inattentional deafness; auditory alarms; warning misperception; aeronautics; eye tracking; psychophysiology

INTRODUCTION

Auditory alarms are known to present various advantages in emergency situations compared to visual alarms. In aeronautics, they provide information for pilots without requiring head/gaze movements (Edworthy, Loxley, & Dennis, 1991) and elicit faster reaction times (Wheale, 1981). Yet, the analysis of air safety reports reveals that a significant number of accidents are due to a lack of reaction to auditory alarms (Bliss, 2003). Three reasons are often raised to account for such a lack of response. One first explanation is that alerting systems, if perceived as unreliable, are likely to provoke the so-called 'Cry-Wolf Effect' (Breznitz, 1984; Wickens et al., 2009) and also lead to mistrust in alarms (Shapiro, 1994; Song & Kuchar, 2001; Sorkin, 1988) especially under high workload conditions (Bliss & Dunn, 2000). A second explanation is that the sometimes aggressive, distracting, and annoying nature of auditory alarms (Doll, Folds, & Leiker, 1984; Edworthy Loxley, & Dennis,1991) can increase the level of stress during warning events (Peryer, Noyes, Pleydell-Pearce, & Lieven, 2005). Indeed, for many pilots their initial response to alarms is to find a way to silence the noise, rather than to process the auditory stimulus for its meaning. Finally, a third explanation is related to frequent noise-exposure and aging issues, known to impair the pilots' ability to perceive auditory warnings (Beringer & Harris, 1999).

Nevertheless, these considerations are not sufficient to fully account for the lack of detection of critical auditory warnings as often reported in accident analyses (Bea, 1993, 2012) and observed in flight simulators (Dehais, Tessier, Christophe, & Reuzeau, 2010). An additional explanation is to consider the role of the sustained perceptual and attentional processes engaged in the cockpit. Evidence suggests that tasks involving high perceptual load consume most of attentional capacity, leaving little or none remaining for processing any task-irrelevant information (see Lavie, 1995). Consequently, high-load contexts tend to prevent the perceptual processing of task-irrelevant information and facilitate various forms of inattentional blindness (Mack & Rock, 1998; Simons & Chabris, 1999). Steelman, McCarley, and Wickens (2011) showed that salience alone does not guarantee visual attentional capture in complex dynamic workspaces such as cockpits; the detection of warning signals also depends upon attentional allocation over the flight instruments.

This propensity to remain unaware of unexpected, though fully perceptible stimuli is not limited to vision, however. Back in the 50s, the seminal work of Cherry (1953) on dichotic listening revealed that unexpected changes (e.g., of language) in the message presented in an ignored auditory channel tended to remain unremarked by listeners. There is now contemporary evidence that unexpected salient sounds can remain unnoticed under attention-demanding conditions (e.g., Spence & Read, 2003; Fenn et al., 2011; Fuchs, Plack, Reese, & Palmer, 2010; Hughes, Hurlstone, Marsh, Vachon, & Jones, 2013), even in experts (e.g., Koreimann, Strauss, & Vitouch, 2009). Although less well-known than its visual counterpart, this inattentional deafness phenomenon (Dalton & Fraenkel, 2012; Koreimann et al., 2009; see also Wayand, Levin, & Varakin, 2005) could account for the pilots' inability to detect auditory alerts. Although there is still a predominant view according to which attention can be divided by modality (auditory vs. visual) – that is, many researchers assume there are two separate, and to some extent independent, pools of attentional resources available to perform cognitive tasks (see, e.g., Wickens, 1984) – a growing body of literature provides evidence that attention is shared between visual and auditory modalities at a more central level (Brand-D'Abrescia & Lavie, 2008; Santangelo, Olivetti Belardinelli, & Spence, 2007; Sinnett, Costa, & Soto-Faraco, 2006, Banbury, Macken, Tremblay, & Jones, 2001). In the case of separate pools, the tasks or cognitive activities from different modalities should not interfere with each other; however, a pool common to all modalities would lead to interference whenever attentional demand is high. For instance, the mere capacity of detecting an unexpected auditory stimulus has been shown to diminish when engaged in visual tasks of high perceptual load (Macdonald & Lavie, 2011). This attentional issue has also been demonstrated in the context of more ecological situations such as a radar-based monitoring and risk assessment task (Vachon, Nicholls, Jones, & Tremblay, 2011). Indeed, in Vachon et al.'s (2011) experiment, when participants had to monitor auditory channels for information critical to their assessment, in addition to monitoring the dynamic visual information and interacting with the visual interface, they missed up to 21% of unexpected but critical changes in the "urgent" auditory messages. In a similar operational context, some authors have shown electrophysiological evidence of this phenomenon as neural responses of the auditory system to unexpected sounds are attenuated when they conflict with visual information in the cockpit (Scannella, Causse, Chauveau, Pastor, & Dehais, in press), or when the visual primary task load increases (Giraudet, Saint-Louis, Causse, 2012; Kramer, Trejo, & Humphrey, 1995). Hence, since flying involves multitasking, induces high workload (Lee & Liu, 2003) and high engagement (Causse et al., 2013), it is more likely that auditory warnings could be missed.

Present Study

The objective of this study is to test whether inattentional deafness is likely to occur in the context of flying and, if so, to assess the potential impact of such a phenomenon on the pilot's behavior. An experiment was conducted in a flight simulator with pilots who had to perform landings in conditions that would induce either low or high cognitive load. At some point during landing, an audible alarm indicating a landing-gear failure was triggered. The detection of this alert should lead the pilot to abort landing and perform a go-around maneuver. In the high workload scenario, the alarm occurred concurrently with a buffeting-inducing windshear, yielding a sudden increase in cognitive load, whereas in the other (low-load) scenario, the alarm was triggered alone. Our prediction is that the level of cognitive load will affect the ability of pilots to detect the landing-gear failure auditory alarm. In the 'windshear' scenario, 1) we expected that the occurrence of the windshear would suddenly increase cognitive load and that pilots would become particularly susceptible to failing to notice the landing-gear auditory alarm. On the contrary, 2) it was predicted that the pilots would be much more likely to detect that alarm in the 'no-windshear' scenario. For half of the pilots the 'windshear' scenario was presented first, while the other half started with the 'no-windshear' scenario. We analyzed whether the order of exposure to conditions had an impact on vulnerability to inattentional deafness.

The assessment of inattentional deafness in a complex dynamic situation is particularly challenging as the non-detection of an auditory stimulus cannot directly translate into an observable response such as when an explicit change-detection task must be performed (cf. Vachon, Vallières, Jones, & Tremblay, 2012). Of course perceiving—and understanding—an auditory alarm (e.g., landing-gear failure) should lead to the application of the appropriate maneuver (e.g., go around); however the production of the expected action cannot guarantee the alarm was noticed as the behavior could have been motivated by another reason (e.g., the concurrent occurrence of a windshear). A more subjective way of measuring inattentional deafness is to ask participants after the experiment whether they faced special events during the scenario (cf. Dalton & Fraenkel, 2012; Macdonald & Lavie, 2011). No mention about an auditory alert can be taken as evidence for the non-detection of this alarm. However, the retrospective nature of such a post-hoc measure makes it susceptible to short-term memory (STM) limitations (Borrie, Ruggenbuck, & Hull, 1998) as a non-negligible amount of time can elapse between the instant the alarm is triggered and the moment the query is presented to the pilot. As a remedy to such pitfalls in measuring inattentional deafness, we advocated a multi-criteria approach that combined both subjective, post-experimental queries and objective, goal-related behaviors. We reasoned that a true instance of inattentional deafness should be reflected in a pilot who does not declare having heard the auditory warning regarding the landing-gear failure, and who at the same time fails to produce the expected reactions to this alarm such as a confirmatory glance at the visual landing-gear indicator and a go-around maneuver.

METHOD

Participants

Twenty-eight healthy male pilots (mean age = 38.22 years, SD = 16.3; flight experience = 2997.7 hours, range = 55–12000), all French defense staff from Institut Supérieur de l'Aéronautique et de l'Espace (ISAE) campus, were recruited by local advertisement and did not receive any payment for their participation. They all reported normal or corrected-to-normal vision and normal audition. Participants were randomly assigned into two independent groups. Those in the 'windshear first' group completed the windshear scenario first and then the no-windshear scenario, while this order was reversed for the other half of participants. Age and flight experience of each group are presented in Table 1.

Table 1
Characteristics of the sample of pilots of the present study.

	Total sample	'Windshear first' group	'No-windshear first' group
N	28[1]	14	14
Mean age (+*SD*) in years	29.52 (11.90)	30.54 (10.11)	28.33 (13.74)
Flight experience (+range) in hours	362 (30–3500)	338 (30–3500)	390 (32–1890)

[1] All pilots were males.

Flight Simulator

A 3-axis motion (roll, pitch and height) flight simulator built by the French flight test center was used to conduct the experiment (see Figure 1). It simulates a twin-engine aircraft flight model and reproduces aerodynamic effects such as buffeting (i.e., aircraft vibration during stall). Its user interface is composed of a Primary Flight Display, and a simplified Head-Up Display comprising a speed vector, a Navigation Display, and the upper Electronic Central Aircraft Monitoring Display page. The pilot has a stick to control the flight, a rudder, and two thrust levers.

Figure 1 about here

Two stereophonic speakers, located under the displays on each side of the cabin, were used to broadcast continuous radio communication and engine sound as background noise (77dB(SPL)), and to trigger four types of alarms (single chime, triple chime, repetitive chime, and pull up) presented at 86.3dB(SPL), that is, 8.5 times louder than the global ambient cockpit sound. Software was implemented to automatically manage the different events (e.g., failure, gusts of wind) that occurred during the landings.

Experimental Scenarios

Participants performed two scenarios that differed from each other in the level of cognitive demands required at the critical moment that the audio alarm occurred. Both windshear and no-windshear scenarios consisted of a manual landing on the 14R runway at Blagnac airport (Toulouse, France). The initial conditions were defined as follows: 2500 feet, heading 142 degrees, 130 knots, visibility 8100 m, slight rain, landing flaps configuration, the landing gear was in transit ("three red"). The landing-gear sequence was supposed to be complete ("three green") before the aircraft reached an altitude of 900 feet. At 900 feet, a failure of the undercarriage sequence occurred and participants were warned through the landing-gear indicator ("two green" and "one red", instead of "three green"; see Figure 2) and a triple-chime auditory alarm. This event should lead pilots to abort the landing and perform a go-around procedure. In the windshear scenario only, participants faced a windshear that critically dropped the speed of the aircraft simultaneously to the landing-gear failure. Such an addition to the basic (no-windshear) scenario was thought to induce a sudden increase in workload at the decisive moment of the flight: the occurrence of the critical landing-gear alarm. Both scenarios ended when participants reached the landing ground touchdown area, whatever their altitude, and the displays were switched off. No crashes were simulated. The scenario duration was about 2 minutes: the first segment lasted around 1min 30 sec (until 900ft) and the last segment (until touchdown) lasted around 30 sec.

Figure 2 about here

Procedure

Participants were told that the purpose of the experiment was to analyze cardiac responses and visual patterns during landings. A 20-min tutorial detailed the functioning of the simulator (user interface, important flight parameters). In particular, pilots were told that five different events were likely to occur during landings: an antiskid failure (simulated by an auditory "single chime"), an engine failure (simulated by an auditory "repetitive chime" and a red warning on the corresponding engine indicator), a decision height issue (poor external visibility); a ground proximity issue (simulated by an auditory “Pull Up” alarm); and finally a landing-gear failure (simulated by an auditory “triple chime” warning and “one red and two green” on the landing-gear indicator if the undercarriage sequence was not completed at 900 feet). All the auditory alarms were well known transportation airplane alerts. The associated procedures were explained (respectively: antiskid failure: “do not exceed 130 knots at touchdown”; engine failure: “set the corresponding throttle lever on idle and use the rudder”; decision height issue: “perform a go-around if the runway is not visible at 200 feet”; ground proximity issue: “perform an immediate go-around”; and a landing-gear failure: “proceed to an immediate go-around to further recycle the landing gear”). In fact, pilots only encountered the landing-gear failure during the two experimental scenarios.

Participants sat in the flight simulator and the sensors (electrocardiogram, eye tracker) were set before starting a 5-min resting period without any stimulation. They then completed a 1-hour training session in which they performed manual landings, in particular supervising the automatic undercarriage sequence that was supposed to end before 900 feet (every 250 feet, a wheel was locked one after the other). Training was performed with no simulator motion and no auditory radio communications in order to ensure that pilots were capable of performing the task appropriately before placing them within a proper immersive scenario. Each new landing, commencing midair, was progressively more difficult due to slight changes in landing conditions (i.e., stronger crosswind, lower visibility, etc.). Participants were told that they were free to perform a go-around if necessary and that there was no traffic in the landing pattern. During training, all the different alarms were presented before the beginning of the third, fifth, seventh, and ninth landings and participants were asked to identify the events and recall the associated procedures. Participants were also trained to fill out a questionnaire and a self-report after each landing (see next section). After the ninth practice landing, the simulator motion was engaged to reproduce realistic flight sensations, and a continuous radio communication was also broadcasted to reproduce more ecological flight conditions. Introducing the motion and the radio communication created an immersive environment in which participants were then more likely to be surprised by a sudden falling sensation induced by the simulator motion, and in turn, to miss the auditory alert. Participants then performed the two experimental scenarios with no break, in the order predetermined by group.

Metrics

A set of measures, ranging from subjective, self-reported metrics to objective, behavioral and psychophysiological measurements, was extracted in order to determine whether the introduction of the windshear was successful in increasing cognitive workload/psychological stress, and to assess how such an increase in load/stress affected the detection of the landing-gear audio alarm.

Subjective measurements. Participants were asked to fill out a 4-item questionnaire directly after the end of each scenario. The questions were: 1) “Describe the weather and wind conditions”; 2) “Describe the status of the aircraft”; 3) “Describe the particular events you have faced”; and 4) “Describe your actions and decisions”. From this debriefing questionnaire, we extracted information to verify perception of the auditory alert and of the failure via the landing-gear indicator, and to analyze the decision that led to a go-around when one had been performed. Moreover, a self-report of mental workload, psychological stress, perceived difficulty, and self-estimated performance level was collected using a visual analog scale (1 for very low, 7 for very high).

Heart rate measurement. Heart rate (HR) was taken as an objective measure of the level of mental workload and psychological stress (Causse, Sénard, Démonet, & Pastor, 2010). An electrocardiogram was used to collect participants' cardiac activity at a sampling rate of 2048Hz with the Biopac® system. Three electrodes connected to an extender cable were applied to participants' chests using Uni-Gel to enhance the quality of the signal. The Biopac Acqknowledge software was used to export and filter the HR derived from the inter-beat-interval. A continuous measurement of HR was recorded during both experimental scenarios. A time domain analysis was not performed as the duration of all scenarios was shorter than 4 minutes (i.e., the minimum period of time to calculate heart rate variability; Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996). In order to test the validity of our load and psychological stress manipulation, the cardiovascular response was contrasted across two segments. Hence, the first segment (S1) included the interval between the beginning of the scenario and the moment the plane reached an altitude of 900 feet (i.e., when the alarm was triggered) whereas the second segment (S2) began when the plane arrived at 900 feet until the end of the scenario (i.e., with an erroneous landing or a go-around maneuver). HR was averaged across the whole duration of each segment. The impact of the windshear occurrence on cognitive workload/stress level was assessed by comparing the change in HR from S1 to S2 between the two scenarios.

Ocular measurement. A Pertech® head-mounted eye tracker was used to analyze participants' ocular behavior. This 80-g non-intrusive device has 0.25° of accuracy and a 50-Hz sampling rate. The EyeTechLab software provided data such as timestamps and the x,y coordinates of the participants' eye gaze on the visual scene. Eye-tracking data were used to check whether participants glanced at the landing gear indicator during the scenario and the associated latency.

RESULTS

Validation of the mental workload/psychological stress manipulation

The impact of the windshear on mental workload/psychological stress was assessed through subjective and objective measures. Table 2 displays the mean score obtained in each scenario for each of the four self-rated metrics. These means were computed from the data of 26 subjects as 2 pilots did not complete the self-rated scales. The results of the dependent-samples *t* tests were compelling (see Table 2): pilots judged that the introduction of the windshear into the scenario had a significant negative impact on their work. Indeed, they found the windshear scenario more demanding in terms of mental load, more stressful, and more difficult than the no-windshear scenario. Moreover, subjects felt less confident about their performance in the presence of the windshear.

Table 2

Mean scores (+SE) on each of four self-rated scales obtained after each scenario.

Self-rated scale	Scenario		Observed *t* value
	No windshear	Windshear	(*df* = 25)
Mental workload	3.83 (0.23)	4.90 (0.25)	4.10*
Psychological stress	3.50 (0.19)	4.67 (0.22)	5.76*
Perceived difficulty	3.48 (0.22)	4.92 (0.20)	7.56*
Self-estimated performance	4.37 (0.23)	2.92 (0.19)	-5.04*

* $p < .001$, power < .99

With regards to the objective measurement of workload/stress level, we computed the cardiovascular response to the alarm/windshear event through HR change—here, increase—from S1 to S2. Two participants were excluded from this analysis due to missing data. The mean HR change was larger in the windshear scenario (*M*=6.56 bpm, *SE*=1.14) than in the no-windshear scenario (*M*=4.90 bpm, *SE*=1.05). This difference was significant, *t*(25)=1.74, *p*=.048 (one-tailed, dependent samples), power=.517, suggesting that introducing a windshear intensified the level of objective workload/stress. Although we cannot preclude the possibility that this increased HR reflected instead some sort of arousal, we nonetheless conclude from both subjective and objective stress indicators that the windshear manipulation contribute to increasing mental workload/psychological stress.

Assessment of Inattentional Deafness

Given that participants were asked about whether they had noticed a special event in the scenario up to 30 s after the audio alarm was triggered, STM rather than attentional limitations could be responsible for any inability to recall the occurrence of that alarm. Accordingly, we employed a multi-criteria approach to evaluate inattentional deafness in which we crosschecked pilots' alarm perception with gaze behavior towards the landing-gear indicator and pilots' decision regarding the correct maneuver to perform. We reasoned that, firstly, a pilot that truly missed the audio alarm is much less likely to glance at the landing-gear indicator immediately following the alarm. Secondly, the non-detection of the auditory and visual alerts should not lead to the appropriate maneuver, that is, a go-around performed due to the failure. Therefore, we considered as a true instance of 'deafness' when a pilot 1) did not report having heard the triple-chime warning during the scenario, 2) did not glance at the landing-gear indicator after the alarm, and 3) did not perform the expected maneuver, either by landing the plane or by justifying the go-around through the need to stabilize the plane, and not in reaction to the alarm. The meeting of these three criteria constitutes a clear indication that the pilot was unaware of the landing-gear failure due to the non-perception of the critical alarm.

As expected, in the no-windshear scenario, all pilots reported having detected the auditory and visual alarms and performed the go-around for the appropriate reason (i.e. the failure); we thus focused our analysis of inattentional deafness on the data from the windshear scenario. Table 3 presents the data relative to the three above-mentioned criteria for every pilot in the windshear scenario. Overall, 12 out of the 28 pilots (42.9%) did not report at the end of the scenario having perceived the auditory alarm during the windshear. However, one of these 12 participants (Subject 3) did not declare having heard the audio alert in the post-experimental questionnaire but nevertheless reported having seen the visual alarm—he indeed looked at the landing-gear indicator 4.05 s after the alarm was triggered—and performed the go-around maneuver because of the landing-gear failure. He was thus not considered a 'deaf' pilot. Accordingly, we concluded that 11 pilots (39.3%) suffered from inattentional deafness in the windshear scenario based on our multi-criteria approach. Among these 'deaf' pilots, 45.5% of them landed the plane despite the landing-gear failure, while the remaining pilots (54.5%) declared they performed the go-around because the windshear destabilized the plane. Besides, there was a strong relationship between the detection of the auditory alarm and the further go-around action in the windshear scenario, $\chi^2(1, N=28)=8.12, p=.004$, power=.813. In fact, all pilots who consciously noticed the audio alert during the windshear did perform the correct maneuver (i.e. the go-around).

Table 3

Pilots' behavioral performance in the windshear scenario according to whether that scenario was encountered first or second. Bold characters highlight pilots who suffered inattentional deafness.

Subject	Audio alarm detection	Visual alarm detection	Timing to glance at landing-gear indicator (s)	Maneuver following the alarm	Reported origin of the go-around
Windshear = 1st scenario					
1	Yes	Yes	0.49	Go-around	Failure
2	Yes	Yes	8.24	Go-around	Failure
3	No	Yes	4.05	Go-around	Failure
4	Yes	Yes	14.2	Go-around	Failure
5	No	No	--	Go-around	Unstabilized
6	No	No	--	Go-around	Unstabilized
7	Yes	Yes	11.50	Go-around	Failure
8	No	No	--	Landing	--
9	No	No	--	Landing	--
10	Yes	Yes	0.50	Go-around	Failure
11	No	No	--	Landing	--
12	No	No	--	Go-around	Unstabilized
13	No	No	--	Landing	--
14	No	No	--	Go-around	Unstabilized
Windshear = 2nd scenario					
15	No	No	--	Go-around	Unstabilized
16	Yes	Yes	1.00	Go-around	Failure
17	Yes	Yes	0.1	Go-around	Failure
18	Yes	No	--	Go-around	Failure
19	Yes	Yes	0.22	Go-around	Failure
20	Yes	Yes	4.40	Go-around	Failure
21	Yes	No	--	Go-around	Failure
22	No	No	--	Landing	--
23	Yes	Yes	0.1	Go-around	Failure
24	No	No	--	Go-around	Unstabilized
25	Yes	Yes	2.42	Go-around	Failure
26	Yes	Yes	24.1	Go-around	Failure
27	Yes	Yes	1.00	Go-around	Failure
28	Yes	Yes	0.27	Go-around	Failure

Although inattentional deafness is considered a cognitive phenomenon that can affect anyone, one may argue that the cockpit flight experience may influence the ability to detect an audible alarm. In order to rule out the hypothesis that the deafness observed in the current study could be attributable to a lack of flight experience, we contrasted the hours of flight time of the 'deaf' pilots with those of the 'non-deaf' pilots. The flight experience of one 'non-deaf' pilot was not available. The analysis revealed no significant difference between the number of hours of flight time between 'deaf' (M=413.16 hours; SE=312.71) and 'non-deaf' pilots (M=307.13 hours; SE=146.61), $t(26)<1$ (independent samples), power=.061. Such a result confirms that flight experience cannot account for the non-detection of the auditory alarm.

An interesting finding arose when comparing the rate of inattentional deafness according to whether pilots encountered the windshear in the first or in the second scenario they performed. Indeed, whereas 57.1% of the pilots who first completed the windshear scenario failed to perceive the auditory alarm, such a rate dropped to 21.4% for pilots who started with the no-windshear scenario. In fact, a pilot who initially performed the task with no windshear was 4.89 times more likely to detect the audio alarm in the subsequent windshear scenario than a pilot who experienced the windshear first. This relationship between the order of the scenarios and the tendency to deafness was significant, $\chi^2(1, N=28)=5.25, p=.022$, power=.630. These results suggest that pre-exposure to the auditory landing-gear failure alarm primed pilots to subsequently detect the same alarm in a more complex situation.

DISCUSSION

The objective of this study was to show that inattentional deafness could be one cause of aircraft pilots' inability to react to auditory alarms. A particular issue was to demonstrate that the inability to recall the presence of the alarm did not ensue from STM or “inattentional amnesia” (Wolfe, 1999). Indeed, often in inattentional deafness paradigms, the assessment of auditory stimulus detection is based solely upon questions immediately after the occurrence of the stimuli (see Macdonald & Lavie, 2011). In our study, for the sake of ecological validity and to ensure that participants were not interrupted in their piloting task, the debriefing question was presented 30 s after the occurrence of the auditory alarm, after the scenario had ended. Also we used a multi-criteria approach based on objective and subjective measurements. Our results seem to support the hypothesis that a salient and relevant auditory alert could remain unintentionally unnoticed. Indeed, in the windshear scenario, 39.3% (i.e. 11) of the pilots reported neither the auditory warning nor the landing-gear failure and continued to land or perform a go-around maneuver due to a stabilization—and not a landing-gear—issue. In this latter case, participants declared they performed the go around because the energy or the trajectory of the aircraft could not guarantee a safe landing. On the other hand, all the pilots who reported having heard the alarm demonstrated a subsequent eye fixation towards the landing-gear indicator and performance of the expected maneuver to avoid a gear-up landing.

It is noteworthy that inattentional deafness occurred only in the windshear condition. Subjective results tend to confirm that this scenario elicited the highest subjective workload, psychological stress and increased task difficulty as the introduction of the windshear led pilots to perform a series of corrective actions to “restabilize” the aircraft. Psychophysiological results also revealed faster HR following the windshear, suggesting that the sudden sensation of falling induced by the windshear-like motion intensified pilots’ mental workload and psychological stress (Dehais, Causse, & Tremblay, 2011; Dehais, Causse,Vachon, Tremblay, 2012). It is true, however, that based solely on HR measurements, it is difficult to exclude an explanation in terms of increased arousal (Causse et al., 2010). Nevertheless, the consistency between our subjective and physiological measures of mental workload/stress suggests that the windshear induced a mobilization of mental resources to the detriment of processing the failure. This is in line with the growing evidence that high cognitive load (Macdonald & Lavie, 2011) and high task difficulty (Eramudugolla, Irvine, McAnally, Martin, & Mattingley, 2005) promote the failure to detect unexpected auditory stimuli.

To account for the failure to notice unexpected visual objects, Most (2010) proposed two distinct loci of inattentional deafness. First, an unexpected stimulus can remain undetected when covert spatial attention is focused away from that stimulus. Given the evidence that looking directly at an unexpected object does not guarantee its detection (e.g., Most, Simons, Scholl, & Chabris, 2000), it has also been proposed that inattentional blindness can originate from a central bottleneck independent of the locus of spatial attention, whereby objects are missed due to a failure of visual awareness. Evidence for 'central' inattentional blindness comes, among others, from studies showing that the phenomenon can be simply induced by increasing cognitive load (e.g., Todd, Fougnie, & Marois, 2005). A similar dual-mechanism approach has been recently applied to change blindness, a related phenomenon, by Vachon et al. (2012) where the failure to notice a change in a visual scene can ensue from either the misallocation of spatial attention—away from the changing object—or an attentional breakdown—an overload in attentional processes leaving the change with insufficient resources to reach consciousness. With regards to the present results, the non-detection of the auditory alarm is more likely to reflect the 'central' source of inattentional deafness than the 'spatial' source. Indeed, the alert was missed only when co-occurring with a sudden increase in cognitive load, suggesting that the temporarily high demand in attentional resources induced by the windshear temporarily reduces cognitive access (cf. Block, 2007) to the alarm's perceptual representation. Moreover, the fact that the very same alert was invariably detected under a condition with no such load variation (i.e. the no-windshear condition) indicates that pilots were able to appropriately allocate their attention toward the alarm. Although we established that central inattentional deafness can take place in the simulated cockpit, it is noteworthy that this conclusion does not preclude an alarm being missed because attention was focused elsewhere.

One could argue that our results ensued from a mistrust in alarms rather than the phenomenon of inattentional deafness. Indeed, the main explanation for alarm misperception, based on accident statistics (Bliss, 2003) and research (e.g., Breznitz, 1984; Wickens et al., 2009), is related to issues regarding a lack of alarm reliability. Besides, Bliss and Dunn (2000) showed that increasing task workload can magnify alarm mistrust and, in turn, degrade alarm response performance. However, alarm mistrust is unlikely to be responsible for the missing of the auditory alert in the present study. First, there was no false alarm implemented in the current experimental design whereas false alarms are a necessary condition for the cry-wolf phenomenon to take place (Breznitz, 1984). Second, pilots encountered an auditory alarm only twice while performing a scenario (i.e. during the two experimental scenarios), leaving very few alarm instances for mistrust to build up. In addition, if mistrust in alarms was at play in our study, we should have observed some undetected alarms in all scenarios, not only in the presence of the windshear.

The present study indicated that inattentional deafness is a robust phenomenon, as the propensity for pilots to miss the auditory alarm was not related to their cockpit flight experience. This result is in line with the empirical work of Drew et al. (in Press) and Koreimann et al. (2009) with experts in other domains (with cardiologists and musicians, respectively). Their studies showed that expertise cannot fully protect individuals from the attentional failures potentially responsible for inattentional blindness and inattentional deafness.

The analysis also revealed a scenario order effect. The participants that were submitted to the no-windshear scenario first were about five times more likely to perceive the auditory alarm in the subsequent windshear scenario than those who started the experiment with the windshear scenario. The pre-exposure to a (detected) alarm in the first scenario increased the likelihood of noticing the same alarm even if presented concurrent to a windshear, as if pilots were expecting the alarm to ring in the second scenario. This result parallels findings from the inattentional blindness literature whereby expectation of the occurrence of the ‘unexpected’ object can promote its detection (e.g., Mack & Rock, 1998; see Levin, 2002, for a discussion). This pre-exposure effect is also consistent with the demonstration that the attentional-capture power of an irrelevant deviant sound faded away when participants were expecting this sound to occur (Hughes et al., 2013).

The scenario order effect may reflect some sort of priming (or learning) effect from the previous encounter with a significant flight event. Indeed, given that pilots invariably perceived the audio alarm—and thus consciously experienced the associated landing-gear failure—in the absence of windshear, one could consider these pilots who began with the no-windshear condition as having been pre-exposed to this specific critical event. Such a pre-exposure—or experience—is likely to have primed the pilots, on the basis of their attentional set (Most, Scholl, Clifford, & Simons, 2005), to respond appropriately when facing the same critical event in the subsequent scenario, despite the increase in cognitive load and psychological stress induced by the windshear. In fact, this could be an instance of case-based reasoning (Kolodner, 1993) whereby pilots used their memory of their first encounter with the landing-gear failure situation to respond to the same event in the windshear scenario. Whereas O’Hare and Wiggins (2004) demonstrated that case-based learning and reminding can actually improve pilot decision making, the present results provided indirect evidence that using a previous case or experience in responding to a critical flight event can also enhance pilot perception. This suggests that a flight training system that incorporates case-based learning of audio alarm events may be a potentially useful means of improving auditory perception and, hence, counteracting inattentional deafness.

To conclude, despite some limitations, the present study supports the existence of the inattentional deafness phenomenon in a simulated cockpit. Our pattern of results suggests that such a robust cognitive limitation may lead to inappropriate decision making even with experts, which in turn may have dramatic consequences. Such a conclusion in the auditory domain is in line with research demonstrating failures of visual awareness in safety-critical situations that may be disastrous (see Varakin, Levin, & Filder, 2004) and reports of failure to detect auditory alarms in other safety-critical domains of work (e.g., emergency medicine; see Edworthy, 2013). Of course, further research is required to extend our findings to a larger sample of airline pilots and also to integrate neurophysiological measurements (e.g. EEG) so as to pinpoint a neural signature of attentional failures.

KEY POINTS

- An experiment was conducted in a motion flight simulator to test the vulnerability of pilots to inattentional deafness – that is, their inability to detect a critical auditory alarm – under weather conditions that may promote attentional tunneling effects (windshear vs. no windshear at landing approach).
- A multi-criteria approach based on self-reported as well as objective behavioral data, including eye movements, provides evidence of inattentional deafness in a simulated cockpit.
- Participants were better to detect the auditory alarm when they had been previously exposed to the alarm in a no-windshear condition. Pre-exposure seems to reduce the vulnerability to inattentional deafness.

REFERENCES

Banbury, S. P., Macken, W. J., Tremblay, S., & Jones, D. M. (2001). Auditory distraction and short-term memory: Phenomena and practical implications. *Human Factors, 43*(1), 12-29.

BEA. (1993) Bureau Enquête et d'Analyse - *Accident investigation report GF072*. Technical report a40-ek000823a.

BEA. (2012) Bureau Enquête et d'Analyse – *On the accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France flight AF447 Rio de Janeiro-Paris*. Technical report

Beringer, D. B., & Harris, Jr., H. C. (1999). Automation in general aviation: Two studies of pilot responses to autopilot malfunctions. *The International Journal of Aviation Psychology, 9*, 155-174.

Bliss, J. P. (2003). Investigation of alarm-related accidents and incidents in aviation. *The International Journal of Aviation Psychology, 13*, 249-268.

Bliss, J. P., & Dunn, M. C. (2000). Behavioural implications of alarm mistrust as a function of task workload. *Ergonomics, 43*, 1283-1300.

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences, 30*, 481–548.

Borrie, W. T., Roggenbuck, J. W., & Hull, R. B. (1998). The problem of verbal reports in recreation research: Review, recommendations, and new directions. *Tourism Analysis, 2*, 175-183.

Brand-D'Abrescia, M., & Lavie, N. (2008). Task coordination between and within sensory modalities: Effects on distraction. *Perception & psychophysics, 70*(3), 508-515.

Breznitz, S. (1984). *Cry wolf: The psychology of false alarms*. Lawrence Erlbaum Associates.

Causse, M., Péran, P., Dehais, F., Caravasso, C. F., Zeffiro, T., Sabatini, U., & Pastor, J. (2013). Affective decision making under uncertainty during a plausible aviation task: An fMRI study. *NeuroImage, 71*, 19-29.

Causse, M., Sénard, J. M., Démonet, J. F., & Pastor, J. (2010). Monitoring cognitive and emotional processes through pupil and cardiac response during dynamic versus logical task. *Applied Psychophysiology and Biofeedback, 35*, 115-123.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, *25*, 975.

Dalton, P., & Fraenkel, N. (2012). Gorillas we have missed: Sustained inattentional deafness for dynamic events. *Cognition, 124*, 367-372.

Dehais, F., Causse, M., Tremblay, S. (2011). Mitigation of conflicts with automation. *Human Factors, 53*, 448-460.

Dehais, F., Causse, M., Vachon, F., & Tremblay, S. (2012). Cognitive conflict in human-automation interactions: A psychophysiological study. *Applied Ergonomics, 43*, 588-595.

Dehais, F., Tessier, C., Christophe, L., & Reuzeau, F. (2010). The perseveration syndrome in the pilot's activity: Guidelines and cognitive countermeasures. *Human Error, Safety and Systems Development, 5962*, 68-80.

Doll, T. J., Folds, D., & Leiker, L. A. (1984). *Auditory information systems in military aircraft: Current configurations versus the state of the art.* Final Report, 1 May-30 Sep. 1983 Georgia Inst. of Tech., Atlanta. Systems Engineering Lab.

Drew, T., Vo, M. L. H., Wolfe, J. M. (in press). The invisible gorilla strikes again: Sustained inattentional blindness in expert observers. *Psychological Science*.

Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors, 33*, 205-231.

Eramudugolla, R., Irvine, D. R. F., McAnally, K. I., Martin, R. L., & Mattingley, J. B. (2005). Directed attention eliminates ‘change deafness’ in complex auditory scenes. *Current Biology, 15*, 1108-1113.

Fenn, K. M., Shintel, H., Atkins, A. S., Skipper, J. I., Bond, V. C., & Nusbaum, H. C. (2011). When less is heard than meets the ear: Change deafness in a telephone conversation. *The Quarterly Journal of Experimental Psychology, 64*, 1442-1456.

Fuchs, P. A., Plack, C. J., Rees, A., & Palmer, A. R. (2010). *The Oxford Handbook of Auditory Science: Hearing* (Vol. 3). Oxford University Press, USA.

Giraudet, L., Saint-Louis, M-E., Causse, M. (2012). Electrophysiological correlates of inattentional deafness: No hearing without listening. *Proceedings of the Human Factors and Ergonomics Society Europe Chapter Conference*. Toulouse, France.

Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Vachon, F., & Jones, D. M. (2013). Cognitive control of auditory distraction: Impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 539-553.

Kolodner, J. (1993). *Case-based reasoning*. San Mateo, CA: Morgan Kaufmann.

Koreimann, S., Strauss, S., & Vitouch, O. (2009). Inattentional deafness under dynamic musical conditions. *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music* (pp. 246-249).

Kramer, A. F., Trejo, L. J., & Humphrey, D. (1995). Assessment of mental workload with task-irrelevant auditory probes. *Biological Psychology,* 40(1), 83-100.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 451-468.

Lee, Y. H., & Liu, B. S. (2003). Inflight workload assessment: comparison of subjective and physiological measurements. *Aviation, Space, and Environmental Medicine, 74*, 1078-1084.

Levin, D. T. (2002). Change blindness blindness as visual metacognition. *Journal of Consciousness Studies, 9,* 111-130.

Macdonald, J. S. P., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics, 73*, 1780-1789.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge: The MIT Press.

Most, S. B. (2010). What's "inattentional" about inattentional blindness? *Consciousness and Cognition, 19*, 1102–1104.

Most, S. B., Simons, D. J., Scholl, B. J., & Chabris, C. F. (2000). Sustained inattentional blindness: The role of location in the detection of unexpected dynamic events. *Psyche, 6*(14).

Most, S. B., Scholl, B. J., Clifford, E. R., & Simons, D. J. (2005). What you see is what you set: Sustained inattentional blindness and the capture of awareness. *Psychological Review, 112*, 217–242.

O'Hare, D., & Wiggins, M. (2004). Remembrance of cases past: Who remembers what, when confronting critical flight events? *Human Factors, 46*, 277-287.

Peryer, G., Noyes, J., Pleydell-Pearce, K., & Lieven, N. (2005). Auditory alert characteristics: A survey of pilot views. *International Journal of Aviation Psychology, 15*, 233-250.

Santangelo, V., Olivetti Belardinelli, M., & Spence, C. (2007). The suppression of reflexive visual and auditory orienting when attention is otherwise engaged. *Journal of Experimental Psychology: Human Perception and Performance, 33(1),* 137.

Scannella, S., Causse, M., Chauveau, N., Pastor, J., & Dehais, F. (in press). Effects of the audiovisual conflict on auditory early processes. *International Journal of Psychophysiology*.

Shapiro, N. (1994). *NBC Dateline*. New York: National Broadcasting Corporation.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception-London*, *28*, 1059-1074.

Sinnett, S., Costa, A., & Soto-Faraco, S. (2006). Manipulating inattentional blindness within and across sensory modalities. *The Quarterly Journal of Experimental Psychology, 59*(8), 1425-1442.

Song, L., & Kuchar, J. K. (2001). Describing, predicting, and mitigating dissonance between alerting systems. *Proceedings of the 4th International Workshop on Human Error, Safety, and System Development*. Linkoping, Sweden.

Sorkin, R. D. (1988). Why are people turning off our alarms. *Journal of the Acoustical Society of America, 84*, 1107-1108.

Spence, C., & Read, L. (2003). Speech shadowing while driving: On the difficulty of splitting attentions between eye and ear. *Psychological Science, 14,* 251-256.

Steelman, K. S., McCarley, J. S., & Wickens, C. D. (2011). Modeling the control of attention in visual workspaces. *Human Factors, 53*, 142-153.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability: Standards of measurement, physiological interpretation and clinical use. *Circulation, 93*(5), 1043-1065.

Todd, J. J., Fougnie, D., & Marois, R. (2005). Visual short-term memory load suppresses temporo-parietal junction activity and induces inattentional blindness. *Psychological Science, 16*, 965–972.

Vachon, F., Tremblay, S., Nicholls, A. P., & Jones, D. M. (2011). Exploiting the auditory modality in decision support: Beneficial "warning" effects and unavoidable costs. *Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting* (pp. 1402-1406). Santa Monica, CA: Human Factors and Ergonomics Society.

Vachon, F., Vallières, B. R., Jones, D. M., & Tremblay, S. (2012). Nonexplicit change detection in complex dynamic settings: What eye movements reveal. *Human Factors, 54*, 996-1007.

Varakin, D., Levin, D. T., & Fidler, R. (2004). Unseen and unaware: Implications of recent research on failures of visual awareness for human-computer interface design. *Human-Computer Interaction, 19*, 389-422.

Wayand, J. F., Levin, D. T., & Varakin, D. A. (2005). Inattentional blindness for a noxious multimodal stimulus. *The American Journal of Psychology, 118*, 339-352.

Weltman, G., & Egstrom, G. (1966). Perceptual narrowing in novice divers. *Human Factors, 8*, 499-506.

Wheale, J. L. (1981). The speed of response to synthesized voice messages. *British Journal of Audiology, 15*, 205-212.

Wickens. C. D., Rice. S., Keller. D., Hutchins. S., Hughes. J., & Clayton. K. (2009). False alerts in air traffic control conflict alerting system: Is there a “cry wolf” effect? *Human Factors, 51,* 446-462.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds) *Varieties of Attention* (p. 63-101). New York: Academic Press.

Wolfe, J. M. (1999). Inattentional Amnesia. In V. Coltheart (Ed.), *Fleeting memories: Cognition of brief visual stimuli* (pp. 71-94). Cambridge, MA: MIT Press.

ACKNOWLEDGEMENT

This work was funded by Réseau thématique de recherche avancée Sciences et Technologies pour l'Aéronautique et l'Espace (RTRA STAE). We would like to express our sincere gratitude to D. Le Quéau and O. Jankowiak for their support on this project; thank Helen Hodgetts for her critical reading of a draft of this article; P. Labedan and G. Garrouste for their work on the experimental set up and the participants who volunteered their time to take part in this research. Part of this research was presented at the 56th annual meeting of the Human Factors and Ergonomics Society held in Boston, MA, in October 2012.

Dear Dr. William Horrey

RE: Failure to Detect Critical Auditory Alerts in the Cockpit: Evidence for Inattentional Deafness (Dehais, Causse, Vachon, Régis, Menant, & Tremblay)

The authors are very grateful for the editor's interest in our paper and the reviewers' suggestions and comments. In the following pages, we address your and the reviewers' minor comments in greater detail. Hopefully, this revision will meet your concerns.

Best regards

Frédéric Dehais

1) Editor - p.14, for the second criterion, if relevant please indicate the time window for which the pilot could glance at the landing-gear indicator (i.e., within x-seconds of the onset of the alarm)

According to aircraft designers, pilots are generally expected to react within 5 s to a critical event especially when collision with the terrain is imminent. However, in our study we did not specify a time window within which the pilot should glance at the indicator. In any case, on average participants took 4.84 s to look at the indicator, which includes one pilot who took 24 s. Only 4 out of the 28 pilots took more than 5 s to glance at the landing-gear indicator.

2) Editor - p.17, missing word: "4.89 TIMES more likely to…"

The word has been added.

3) Reviewer 1 - Generally, the writing still needs some tightening (need to be more efficient; say more with fewer words).

A thorough proofreading has been performed by an English native speaker to tighten the writing.

4) Reviewer 1 - Lack of a power analysis is still troubling, and can not be justified simply by significant results. Recommend a post-hoc power analysis be done and reported.

The post hoc power level is now provided for every statistical test.

5) Reviewer 1- Differences between training and testing conditions should be acknowledged as they are likely a threat to internal validity.

We now provide an explicit justification in the Method section for why simulator motion and radio communications were introduced only during testing scenarios.

6) **Reviewer 1** - ***Though it would increase length a little, I would still like to see more mention of signal trust/false alarms in relation to your findings.***

We now discussed how our results relate to the issue of alarm mistrust in the Discussion.

7) Reviewer 2 - "indicating that introducing a windshear intensified the level of objective workload/stress" --> As noted in my first review, this claim is based on the assumption that heart rate changes measure only workload or stress. An increase in heart rate might not reflect either of those two. It could just reflect an increase in arousal because something interesting and unexpected happened. The reviewers should not present their speculative interpretation of the meaning of a finding as if it were the finding itself. The phrasing on this point is better in the discussion section.

Following Reviewer 2's suggestion, we toned down our conclusion regarding HR in both the Results and Discussion sections.

8) **Reviewer 2** - ***The Dattel et al (2012) does not contradict the results in this paper -- it addressed a completely different issue (as did work by Benn). Those papers showed that experiencing one inattentional blindness situation does not affect noticing rates for a different inattentional blindness situation. In this study, the pilots experience the same event in the same context (landing a plane). In that context, the second trial is just a divided attention case of the same thing. That makes it effectively the same as all prior inattentional blindness trials that included a divided attention trial after the critical trial. result is similar to other work by Benn and others. It is not contradictory to the present finding which used the same context and task. I think that entire discussion could be cut.***

And, I think the authors make far to much of their order effect -- it's exactly what you would expect. Once asked about an unexpected event, the next trial is a "divided attention" trial in which most subjects are expected to detect the now somewhat-expected event.

We thank R2 for this very useful comment. We thus deleted the discussion about Dattel et al.'s study. However, we kept our discussion about the scenario order effect because it is the first demonstration of such an effect in aviation and we are confident that it has valuable training implications for the domain.

Figure 1: Photos of the ISAE flight simulator. Cockpit view (left panel) and view from outside the simulator (right panel). Participants flew the aircraft from the left seat.

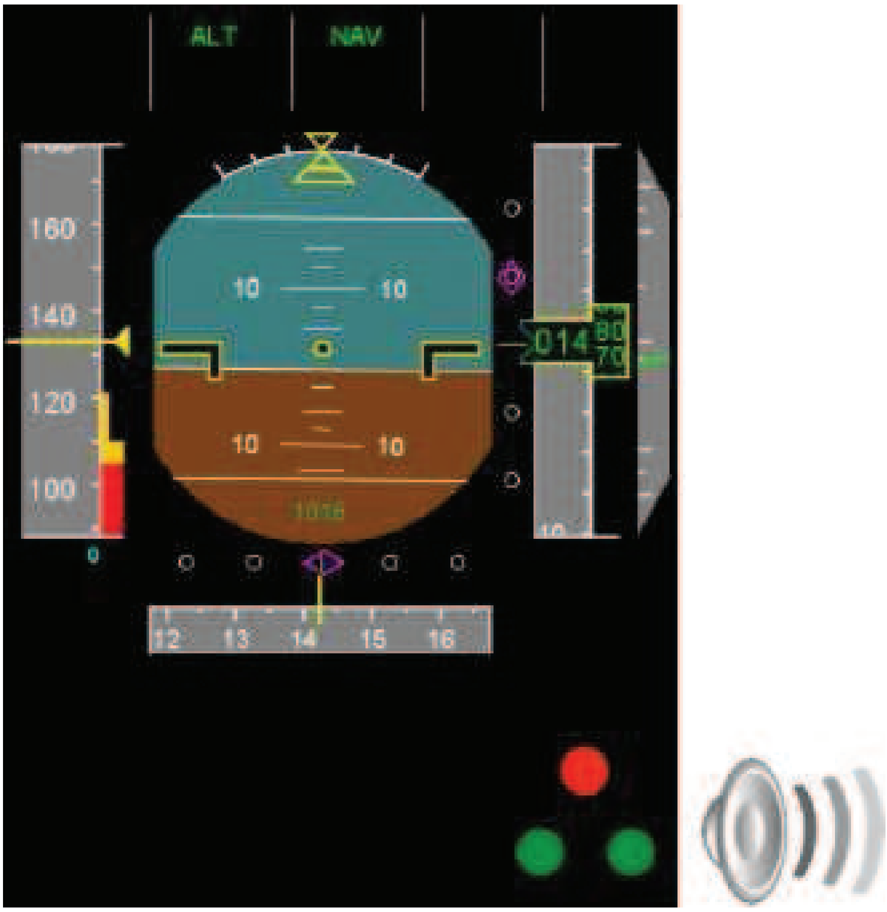

Figure 2. The landing gear indicator was displayed in the right lower part of the PFD along the scenarios. In the scenarios, the undercarriage sequence failed as the “nose” wheel was still not locked at 900 feet (“one red” and “two green” instead of “three green”). A triple chime auditory warning was triggered in the cockpit.