Consciousness and Cognition 35 (2015) 33–41

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Vigilance impossible: Diligence, distraction, and daydreaming all lead to failures in a practical monitoring task

Stephen M. Casner [a,*], Jonathan W. Schooler [b]

[a] National Aeronautics and Space Administration, Mail Stop 262-4, Moffett Field, CA 94035, USA
[b] Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA 93106, USA

ARTICLE INFO

Article history:
Received 12 January 2015
Available online 16 May 2015

Keywords:
Vigilance
Monitoring
Mind wandering
Task-unrelated thought
Distraction
Divided attention
Cockpit
Airline pilots

ABSTRACT

In laboratory studies of vigilance, participants watch for unusual events in a "sit and stare" fashion as their performance typically declines over time. But watch keepers in practical settings seldom approach monitoring in such simplistic ways and controlled environments. We observed airline pilots performing routine monitoring duties in the cockpit. Unlike laboratory studies, pilots' monitoring did not deteriorate amidst prolonged vigils. Monitoring was frequently interrupted by other pop-up tasks and misses followed. However, when free from these distractions, pilots reported copious mind wandering. Pilots often confined their mind wandering to times in which their monitoring performance would not conspicuously suffer. But when no convenient times were available, pilots mind wandered anyway and misses ensued. Real-world monitors may be caught between a continuous vigilance approach that is doomed to fail, a dynamic environment that cannot be fully controlled, and what may be an irresistible urge to let one's thoughts drift.

Published by Elsevier Inc.

1. Introduction

In 1972, the crew of Eastern Air Lines Flight 401 became distracted when an indicator light designed to verify that the airplane's landing gear was down and locked failed to illuminate. As the crew worked together to resolve the problem, which in cruel irony turned out to be only a burned-out light bulb, the airplane continued a slow and unnoticed descent into the Florida Everglades (NTSB, 1973).

In response to this landmark accident, airlines redefined the roles of the two pilots in the cockpit. From then on, one pilot (called the *pilot flying*) would assume primary responsibility for operating the controls of the airplane, while the other pilot (the *pilot monitoring*) would serve as the "second set of eyes" in the cockpit, occasionally assisting the pilot flying in his or her duties but primarily performing the job of keeping watch for anything amiss. Moreover, airline companies devised explicit procedures to be used by monitoring pilots. Monitoring pilots were formally tasked with making verbal callouts as airplanes climbed or descended to or away from assigned altitudes in order to raise awareness among the crew about what the airplane was doing.

But despite these remedial steps, accidents related to lapses in monitoring continued to occur (NTSB, 1994). As recently as 2013, yet another crash occurred in San Francisco after the flight crew failed to notice that the airplane had slowed to an unsafe speed during approach (NTSB, 2014). An unsettling detail about the San Francisco accident was that, unlike most

* Corresponding author.
E-mail addresses: stephen.casner@nasa.gov (S.M. Casner), jonathan.schooler@psych.ucsb.edu (J.W. Schooler).

http://dx.doi.org/10.1016/j.concog.2015.04.019
1053-8100/Published by Elsevier Inc.

flights in which two pilots sit in the cockpit, this flight had four. None of the four pilots noticed the deteriorating airspeed that was shown on a simplistic cockpit instrument and positioned in plain view directly in front of them.

The persistence of airline crashes related to monitoring failures, despite the attention given to most every conceivable aspect of the problem, lead us to wonder if there is something fundamentally wrong with the very idea of placing a human in a supervisory role, however well-trained and experienced they may be. Do our troubles begin when we so much as *attempt* to watch over the shoulder of someone or something else while it works?

1.1. Why vigilance sometimes fails

Upon first glance, the problem of keeping watch for things gone amiss in an environment such as an airline cockpit seems simple enough. Unfortunately, countless studies have shown us that the process of paying attention while watching others work can go wrong in several ways.

1.1.1. Depletion

Thousands of laboratory experiments have demonstrated what happens when participants are asked to monitor a stimulus or process in a prolonged "sit and stare" fashion. What is invariably found is that a decline in performance eventually sets in (Mackworth, 1948; Teichner, 1972). Many argue that the cause of this vigilance decrement is the depletion of cognitive resources needed to perform what is an arduous, stressful, high-workload task (Dillard et al., 2015; Helton & Warm, 2008; Warm, Parasuraman, & Matthews, 2008). As participants work to maintain their focus on whatever they have been assigned to monitor, their stores of cognitive energy are depleted faster than they can be replenished.

1.1.2. External distractions

Early researchers such as Mackie (1987) were quick to point out that, in real-life monitoring tasks, outside the guarded door of the research laboratory, people seldom have the opportunity to sit and stare for extended periods of time. For example, monitoring pilots are periodically asked to answer radio calls from air traffic control or to read checklists. In separate studies of lifeguards and prison guards, Harrell (1999) and Tickner, Poulton, Copeman, and Simmons (1972) found that guards' scanning was also frequently interrupted by other duties such as responding to rule violations that they are tasked with monitoring.

The problem with these momentary distractions from the monitoring task is that they too can lead to monitoring misses. Studies of the deleterious effects of having to divide our attention between more than one task go beyond demonstrating that we are likely to miss something that lies to the left when our attention is momentarily diverted to the right. It has been demonstrated that we often "look but do not see" when cognitive resources needed to meaningfully process what we are monitoring are tied up by another task such as talking or listening (Strayer & Johnston, 2001). Einstein, McDaniel, Pagan, and Dismukes (2003) demonstrated that even momentary attentional diversions often lead to forgotten intentions associated with a monitoring task.

1.1.3. Internal distractions

Other researchers have shown us that, while people work, they sometimes turn their thoughts to matters that are unrelated to the task at hand, and that during these mental excursions, things can slip by unnoticed (Smallwood & Schooler, 2006).

Some researchers argue that mind wandering is the greater threat during a mundane vigilance task (Manly, Robertson, Galloway, & Hawkins, 1999). They argue that mind wandering gradually encroaches upon us, often without our awareness, when we perform monotonous or boring tasks over prolonged periods of time. Ariga and Lleras (2011) suggest that we may knowingly engage in mind wandering when we wish to take mental breaks from a tiring monitoring task. But Kurzban, Duckworth, Kable, and Myers (2013), and Thomson, Besner, and Smilek (2015) have yet a different take on mind wandering. These researchers suggest that we turn to mind wandering when we feel that our attentional resources might be more profitably directed elsewhere. After assessing the frequency at which events of interest might occur during a monitoring task, people may reallocate their limited attentional resources based on the likely reward for monitoring and the opportunity costs associated with not doing or thinking about something else. Casner and Schooler (2014) have already demonstrated that the successful use of highly reliable cockpit automation is associated with frequent reports of task-unrelated thoughts among airline pilots.

1.2. Is vigilance impossible?

These three reasons for which monitoring is suspected to fail, that have been so often studied independently in laboratory settings, combine to offer a rather bleak outlook for those who monitor in real-world settings. It seems that if we set out to actively monitor for long stretches of time, our performance predictably slumps. If we choose to (or are required to) perform other tasks that may arise, we get distracted and miss things for another reason. And if we allow our thoughts to drift onto other matters to take breaks, because we think our attention is not needed, because we find mind wandering irresistible, or simply because we think we can get away with it, we miss things for yet another reason. Whether we are diligent, distracted, or daydreaming, is monitoring doomed to fail?

1.3. An observational study of monitoring in the airline cockpit

We designed an observational study of airline pilots to see if this seemingly cheerless combination of psychological factors could help explain the persistent monitoring problems we see in the airline cockpit. If our theory about the inevitability of monitoring failures is accurate, we should expect monitoring failures to copiously occur, even when highly-experienced airline pilots perform under the watchful eyes of experimenters. And if we do witness a sample of cockpit monitoring lapses, the three reasons for monitoring failures should be able to account for at least a substantial proportion of them.

We asked sixteen airline pilots to act as pilot monitoring during routine flight operations in a full-motion flight simulator certified for airline training and testing. As our sole dependent measure, we noted whether or not pilots were successful in making the required altitude callouts that are central to the pilot monitoring job. We chose the altitude callout task for three reasons: (1) because it is one that is performed routinely during every flight; (2) because it is considered to be an indicator of overall monitoring performance by airline training departments (Sumwalt, 2003); and (3) because lapses in altitude awareness have been associated with numerous incidents and accidents (NTSB, 1994; Palmer, Hutchins, Ritter, & VanCleemput, 1991).

We asked pilots to monitor and make altitude callouts during two different phases of flight that differed in important ways. During the *arrival* phase, airplanes follow a published procedure that spells out in advance the series of altitudes that the airplane must cross. For the monitoring pilot, as soon as they monitor and call out one altitude, responsibility to monitor and call out the next one is immediately upon them. During an arrival, altitude monitoring is a continuous process that offers the monitoring pilot no breaks from the altitude callout task. During the *approach* phase, the environment is much less predictable. During the approach phase, the airplanes have arrived to the busy airspace that surrounds the airport, where air traffic control must concoct off-the-cuff plans for keeping airplanes separated. Airplanes will field directives from air traffic control who ask them to climb or descend. Once a climb or descent has been accomplished, the flight crew must wait for the air traffic controller to issue their next assignment. Thus, the approach phase offers the monitoring pilot breaks from the altitude callout task.

To assess the effects of internal distractions on success at the callout task, as pilots worked, we verbally probed them about the current focus of their thoughts. Pilots were asked to report whether they were thinking about something related to or unrelated to the flight. The verbal thought probing technique is an adaptation of a technique introduced by Singer (1966) and has since been shown to be minimally intrusive (Ericsson and Simon, 1980) and correlate well with physiological measures of task-unrelated thought (Schooler et al., 2011; Smallwood et al., 2011; Smilek, Carriere, & Cheyne, 2010). Most importantly, the thought probe technique been successfully used in several recent studies (Casner, Geven, Recker, & Schooler, 2014; Casner & Schooler, 2014) that examined mind wandering among airline pilots. These studies have demonstrated pilots' routine willingness to acknowledge off-task thoughts when queried, and a relationship (at least in some situations) between such thoughts and performance detriments.

To assess the effects of external distractions on success at the callout task, at the time that each thought probe was given and again when each altitude callout was due, we noted whether or not the monitoring pilot was engaged in any non-monitoring activity such as communicating with air traffic control or the pilot flying, or whether they were interacting with any of the aircraft systems.

To explore the association between mind wandering and the monitoring role, we asked pilots to additionally act as the pilot flying during a portion of the study while we continued to probe them for their thoughts.

2. Method

2.1. Participants

Sixteen active Boeing 747-400 pilots, 7 captains and 9 first officers, participated in the study on a voluntary basis. The sample size of sixteen pilots was decided in advance based on a tradeoff between the statistical power needed for the analysis and the rather high cost of conducting each experimental session. The data from all sixteen pilots who participated in the study are reported here. The sixteen 747-400 pilots had an average of 17,844 h of total flight time (SD = 9736), and an average of 623 h during the past twelve months (SD = 198). In exchange for their participation, pilots received a standard hourly participant pay rate and reimbursement for any travel expenses incurred.

2.2. Apparatus

The Boeing 747-400 (Level D) flight simulator located at the NASA Ames Research Center was used for the experiment. A timer was used to time the delivery of the thought probes. A laptop computer was used to record pilots' responses to the probes. Pilots' performance on the altitude callouts was observed in real time then double-checked using video and audio recordings that were captured using video cameras and microphones installed throughout the cockpit.

2.3. Procedure

For the purposes of this study, each participant flew five flight segments: four arrivals and one approach. Participant pilots acted as pilot flying during two arrivals: one in which they used the highest level of automation available in the cockpit and one in which they used no automation at all. Participant pilots acted as pilot monitoring during two arrivals, again in each of two automation conditions. Participant pilots acted as pilot monitoring during one approach. Arrivals and approaches lasted an average of 16.5 min.

During the experiment, each pilot flew an additional eight flight segments that were designed to assess pilots' manual flying skills while acting in the pilot flying role. The manual flying performance data consist of a variety of different dependent measures that are reported in a separate paper (Casner et al., 2014). In total, pilots flew for approximately 4.5 h with a 1.5-h lunch break in the middle of the session.

Each participant pilot sat in their respective seat: captains occupied the left seat, first officers occupied the right seat. Occupying the other seat was a confederate pilot who is also rated in the 747-400 aircraft and is presently employed by an air carrier company. Pilots were told that the confederate pilot would not offer the assistance that a co-pilot normally would during a real flight. The co-pilot would comply with any requests made by the participant pilot, but would not volunteer any assistance. Air traffic control clearances were issued by another subject matter expert who acted as air traffic controller and who delivered each clearance via radio transmissions to the cockpit in the usual way. Sitting behind the two pilots was the experimenter whose function was to introduce each flight segment and then verbally prompt the experiment pilot for thought probes as they flew. Pilots were verbally prompted to categorize their thoughts at varying intervals (chosen randomly for each participant) averaging one probe every two minutes, ranging from approximately 1 min 40 s to 2 min 20 s.

The experimenter also noted whether or not either pilot was interacting with the automation, air traffic control, or the other pilot at the time that each thought probe was delivered and when each callout was due, and noted monitoring pilots' success at delivering each altitude callout.

Throughout the entire flight, pilots were asked to obey the "sterile cockpit rule" while they flew: to not engage in any sort of non-flight-related conversation.

3. Results and discussion

While serving in the role of pilot monitoring, each of the sixteen pilots had nine opportunities to respond with standard altitude callouts as the aircraft descended to nine different assigned altitudes. Two data points were lost due to experimenter error leaving us with 142 callout measures for the entire group of 16 pilots. Since the roughly nine callouts per pilot must be further broken down into our conditions of interest, we do not attempt to calculate means for individual pilots and use parametric statistics to test our hypotheses. Rather, we aggregate the callouts for all pilots and use a technique that will accommodate the correlated (repeated measures) nature of our data: generalized estimating equations.

Of the 142 total callouts, 107 callouts (75%) were made successfully while the remaining 35 callouts (25%) were missed. The missed callouts were well distributed across subjects: pilots averaged 2.4 missed callouts (SD = 1.4) and only one pilot made all nine callouts. These findings suggest that monitoring misses in the cockpit, even among highly-experienced pilots, are in no way rare occurrences.

3.1. Effects of depletion

We observed no significant decrement in monitoring performance over time. Fig. 1 plots pilots' collective performance on the first, second, and third callout opportunities for all three pilot monitoring segments combined.

A model constructed using generalized estimating equations revealed no significant difference between pilots' performance on the first, second, and third callout opportunities. Although the average duration of our periods of watch were quite short (16.5 min), laboratory demonstrations of the vigilance decrement routinely observe most of the vigilance decrement unfold in the first 15 min, and sometimes less (Warm et al., 2008). Thus, if pilots engaged in any sort of "sit and stare"

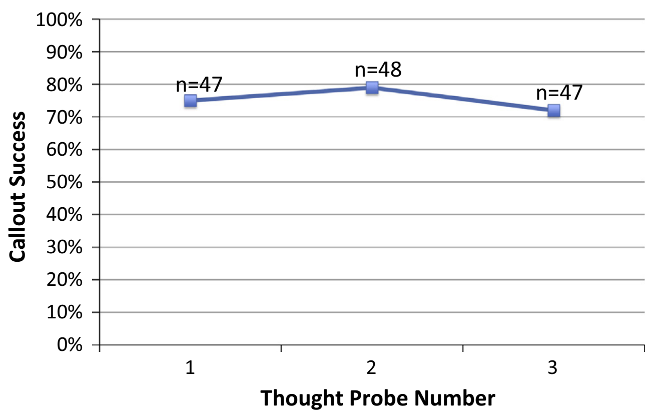

Fig. 1. Pilots' collective performance on the first, second, and third callouts.

monitoring and experienced deleterious effects, we would expect to see some drop in performance in the third callout that occurred in the closing seconds of each period of monitoring. But as next analyses will show, prolonged periods of watch did not occur because our pilots' monitoring was frequently punctuated by other activities.

3.2. Effects of external distraction

Recall that when we administered thought probes (roughly every two minutes), we also observed whether or not pilots were engaged in non-monitoring activities (e.g., interacting with the airplane or verbally communicating with others). Of the 311 thought probes for which we were able to discern pilots' activities, we found pilots engaged in these non-monitoring tasks a total of 69 times (22%). Fig. 2 illustrates how non-monitoring task activity was well distributed throughout the flight. Non-monitoring activities ebbed and flowed among pilots but never strayed far from the average of 22%.

The number and distribution of non-monitoring activities show that periods for which pilots were free to "sit and stare" seldom stretched for more than a few minutes at a time.

At the times in which altitude callouts were due, pilots were observed to be engaged in non-monitoring tasks 19% of the time (26 of 139 callouts observed). Comparing this percentage against the percentages in Fig. 2 suggests that pilots did not schedule their non-monitoring task activities to avoid overlap with altitude callout responsibilities.

As shown by Table 1, when monitoring pilots were neither talking nor interacting with the equipment at the time that an altitude callout was due, pilots' collective average success rate was 82% (93 of 113 callouts made). When monitoring pilots were observed engaging in non-monitoring activities at the time that a callout was due, pilots' collective success rate dropped to 42% (11 of 26 callouts made).

To test the significance of this apparent association between non-monitoring task activity and missed altitude callouts, we used generalized estimating equations to test the effect of being engaged in a non-monitoring task when a callout was due (a binary predictor variable) on whether or not that callout was missed (a binary outcome variable). A stable model resulted (Wald $X^2(1) = 22.72$, $p < .001$) that confirms what is suggested by means given in Table 1. Activities that pulled pilots away from monitoring had a significant effect on pilots' monitoring performance (Coef. = −.42, Std. Err. = .09, OR = .66, $p < .001$).

3.3. Effects of internal distractions

Pilots collectively reported engaging in off-task thoughts in response to 114 of the 318 thought probes that we administered (36%). Relating episodes of task-unrelated thought to missed altitude callouts is more difficult because we did not attempt to deliver thought probes at the same moment that altitude callouts were due, as this might have simply reminded pilots to make the callouts. Since the thought probes were delivered roughly every two minutes, a thought probe occurred less than two minutes before and after each time an altitude callout was due. Thus, we tried using the probes that preceded and succeeded each altitude callout opportunity to predict whether the callouts would be made or missed. The probes that occurred prior to the callout opportunities did not predict callout success. It may be that these probes served as reminders about the upcoming callouts. The probes that occurred after the callout opportunities were significant predictors of callout success and produced an unexpected result.

Table 2 shows the percentage of all callouts that were successfully made by pilots prior to reporting on-task and off-task thoughts. The percentages in Table 2 suggest a near-perfect reversal of the association between mind wandering and callout success across the two phases of flight. Even with our 142 callouts divided between the four cells appearing in Table 2, the apparent interaction between the two variables was confirmed by a model created using generalized estimating equations

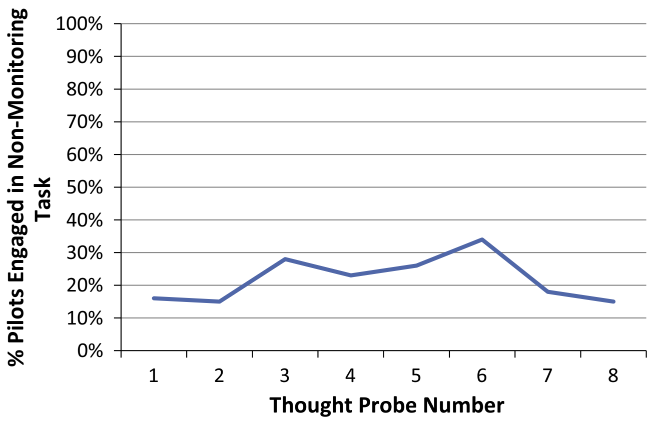

Fig. 2. Percentage of pilots who were engaged in non-monitoring activities when each successive thought probe was delivered (three flight segments collapsed).

Table 1
Number and percentages of the 142 callouts made when monitoring pilots were observed "just monitoring" or engaged in a non-monitoring task.

	Callouts made		
	Arrival	Approach	Combined
No	85% (65 of 77)	78% (28 of 36)	82% (93 of 113)
Yes	38% (6 of 16)	50% (5 of 10)	42% (11 of 26)
Interacting			

Table 2
Number and percentages of the 142 callouts made prior to reporting on-task and off-task thoughts.

	Callouts made		
	Arrival	Approach	Combined
On task	91% (53 of 58)	70% (16 of 23)	85% (69 of 81)
Off task	68% (15 of 22)	92% (12 of 13)	77% (27 of 35)
Thought			

(Wald $X^2(1) = 10.73$, $p < .05$). During the approach phase, callout *success* was significantly associated with subsequent reports of mind wandering (Coef. = .52, Std. Err. = .17, OR = 1.68, $p < .01$).

How can mind wandering be associated with poor performance in one phase of flight and with good performance in another? Recall that the arrival phase offers monitoring pilots few breaks from the altitude callout task. Of the 203 thought probes that were administered during the arrival phase (all subjects combined), only 62 of them (31%) happened during times in which pilots enjoyed a temporary break from monitoring for an upcoming altitude callout. Pilots engaged in 39% (29 of 75) of their reported off-task thoughts during these breaks. The remaining 61% of off-task thoughts that happened at less convenient times and led to the monitoring misses reflected in the bottom left corner of Table 2. On the other hand, the approach phase offered pilots a much greater number of breaks from the altitude callout task. Of the 110 thought probes administered during the approach phase, 81 of them (84%) occurred when pilots were temporarily free from making altitude callouts. Pilots seemed to take advantage of these breaks by fitting 84% (32 of 38) of their off-task thoughts into them and thereby avoid missing callouts. There was a significant effect (Coef. = −.12, Std. Err. = .05, OR = .89, $p < .05$) of these monitoring breaks on the likelihood that pilots would engage in off-task thoughts: Wald $X^2(1) = 5.56$, $p < .05$. Although we did not attempt to directly measure which mind wandering episodes were the result of deliberate planning (Seli, Carriere, & Smilek, 2014), this strong association between monitoring breaks and mental breaks provides at least circumstantial evidence that pilots were often aware of their mind wandering episodes and exercised conscious control over them.

3.3.1. Factors influencing mind wandering

Our results point to what is perhaps the greatest determinant of when pilots will mind wander: when they are given the opportunity to do so. Pilots reported mind wandering roughly 10% of the time (7 of 69 thought probes) when they were interacting with the airplane or others and roughly 43% of the time (104 of 242 thought probes) when they were not engaged in these tasks. Pilots reported mind wandering roughly 31% of the time (51 of 170 thought probes) when they were burdened with an upcoming altitude callout and roughly 43% of the time (61 of 143 thought probes) when they enjoyed a temporary break from altitude callouts. As predictors of the likelihood of mind wandering, a significant effect was observed for engaging in non-monitoring tasks (Coef. = −.31, Std. Err. = .06, OR = .73, $p < .001$) and a weaker effect for the burden of an upcoming altitude callout (Coef. = −.11, Std. Err. = .05, OR = .90, $p < .05$): Wald $X^2(2) = 32.91$, $p < .001$. Idle hands, mouths, and minds were associated with wandering thoughts.

We found no evidence that mind wandering accumulated gradually as suggested by theories about tasks that pilots might find tiresome or boring (Szalma et al., 2004). Fig. 3 aggregates the first through eighth thought probes administered to pilots during all three pilot monitoring segments combined. Note that Fig. 3 does not show the expected increase in mind wandering activity as the flight segments progressed. Rather, the data show that pilots were at least as likely to jump into task-unrelated thought at the very outset of each flight segment, before the airplane started to descend and before altitude callouts were due. This result adds an interesting twist to the ideas set forth by Kurzban et al. (2013) and Thomson et al. (2015). These researchers explain the vigilance decrement as an artifact of the laboratory vigilance task in which novices learn the expected frequency at which their attention will be needed. Mind wandering then sets in as their attentional focus fades to the expected frequency. But unlike novices in a laboratory experiments, experienced pilots are already familiar with the expected frequency, leaving them free to start mind wandering immediately.

We directly tested the likelihood that monitoring pilots' mind wandering was triggered by the flying pilot's use of a highly-reliable automation system. Generalized estimating equations revealed that monitoring pilots did not engage in

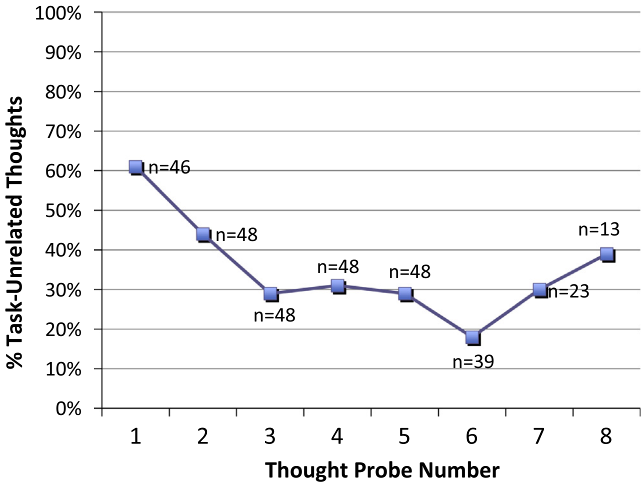

Fig. 3. Percentage of each successive thought probe during which pilots reported being engaged in mind wandering (three flight segments collapsed).

Table 3
Number and percentage of callouts explained by our analysis.

Total missed callouts	35
Depletion	0 (0%)
External distractions	9 (26%)
Internal distractions	8 (23%)
Both external and internal distractions	6 (17%)
Total explained	**23 (66%)**
Total unexplained	**12 (34%)**

significantly more or less task-unrelated thought when the pilot flying was using the highest level of automation available in the cockpit (37%, 34 of 102 probes) or flying the airplane entirely by hand (33%, 80 of 216 probes): a task that is well known to impose high workload and greater opportunity for error on the pilot flying (Casner, 2009). Monitoring pilots appeared to have little sympathy for the pilots flying when they worked harder to fly the airplane.

Lastly, since our participants also flew the airplane during two other flight segments in which they were probed, we can test the extent to which merely placing our pilots in the monitoring role would affect the amount of mind wandering they do. When the pilot flying was using the highest level of automation, both the pilot flying and pilot monitoring were seemingly placed in the position of monitoring the automation. During these times, when acting in each role, pilots spent similar amounts of time interacting (pilot flying = 30% vs. pilot monitoring = 22%) and thus had similar amounts of time available for monitoring (generalized estimating equations revealed no significant association between these two variables). Indeed, to the outside observer, it might be difficult to discern which role our pilots were performing. However, examining the thoughts of pilots in these two roles tells a different story. Pilots reported off-task thoughts for 19 out of 90 probes (21%) when acting as pilot flying, and 41 of 101 (41%) when acting as pilot monitoring. Indeed, assuming the role of pilot monitoring was a significant predictor (Coef. = .19, Std. Err. = .07, OR = 1.21, $p < .01$) of the likelihood that pilots would mind wander: Wald $X^2(1) = 8.59$, $p < .01$. Placing a pilot in the monitoring role may trigger a sense of reduced involvement in the flight.

3.4. Missed callouts associated with the three factors

Overall, the three factors we considered were implicated in 23 of the 35 (66%) missed callouts as shown in Table 3.

The remaining 12 callouts (34%) may have eluded our analysis in a number of ways. First, they may have happened for reasons that we have not considered here. Second, our ability to detect engagement in non-monitoring tasks was limited to observations of speech or other physical actions. Our strategy altogether misses engagement in cognitive non-monitoring tasks (e.g., mental calculations or other reasoning steps). Third, there is little data available that tell us about the typical duration of mind wandering episodes. Our crude thought probe technique allowed us to check pilots' thoughts once every few minutes but pilots may slip in and out of task-unrelated thought much more frequently than this, escaping our analysis undetected.

4. Conclusion

We studied monitoring performance in an important vigilance situation: airline pilots serving in a monitoring role. A first striking result was the number of monitoring lapses we observed while pilots performed a routine cockpit monitoring task.

Pilots missed 25% of the routine altitude callouts that they were tasked with making. Unlike laboratory experiments in which participants' performance gradually deteriorates as they endure long periods of watch, our pilots avoided long vigils and escaped their deleterious effects. One way in which our pilots avoided long periods of watch was to engage in other cockpit activities that are designed into the job of pilot monitoring. Unfortunately, while these breaks from the monitoring task may have helped pilots avoid one sort of failure (i.e., declines in vigilance over time) they seemed to have led them directly to another: the distractions introduced by these non-monitoring tasks were linked to their own monitoring misses. Another way pilots avoided long vigils was to engage in mind wandering: an activity over which pilots may have exercised notable conscious control. But when pilots could not find convenient times to mind wander, mind wandering sometimes led pilots to yet another source of monitoring failure. These results support our theory that whether pilots remain attentive or not, the task of monitoring may offer few paths to success. Engaging in diligent monitoring, mixing up the monitoring task with other sorts of activities, or taking mental breaks all seem to lead to monitoring lapses.

The implication that monitoring lapses are inevitable is hard to accept given that we rely on human monitors to keep us safe in a great many situations. While it is unlikely that we will eliminate the problems associated with long uninterrupted watches, our findings suggest that the two ways of avoiding them that we have studied here (internal and external distractions) might be better managed.

Pilots made no apparent attempt to avoid the deleterious effects of engaging in non-monitoring tasks when an altitude callout was soon due. It may be that pilots were often not free to postpone activities performed at the request of others, or that their confidence in their ability to multitask far outstrips their real ability to do so (Finley, Benjamin, & McCarley, 2014). We might try educating pilots about the "multitasking myth" (Loukopoulos, Dismukes, & Barshi, 2009) or designing a reminder system to help pilots deal with interrupted tasks (Einstein et al., 2003). Palmer and Degani (1991) have explored the idea of using electronic checklists to improve monitoring performance but discovered that adding more automation to the monitoring problem might further lessen pilots' engagement in the task.

Pilots devoted 43% of their available monitoring time to task-unrelated thought, and we observed many instances of mind wandering in the vicinity of monitoring misses. We might consider ways of reducing the overall amount of mind wandering. Less practically, MacLean et al. (2010) have demonstrated the benefits of meditation and mindfulness training on subjects' ability to perform a sustained attention task. Founded on the notion that a busy mind is an engaged mind, deBettencourt, Cohen, Lee, Norman, and Turk-Browne (2015) have demonstrated improved attention through modulating task difficulty when attentional lapses were detected using whole-brain neuroimaging. More practically, Sumwalt (2003) has suggested that monitoring performance might be improved by tasking pilots with actively monitoring a more comprehensive set of flight parameters. However, despite the existing formal requirement to monitor altitude, our pilots still mind wandered to great extent and sometimes missed callouts.

That our pilots may have demonstrated more control over their mind wandering than participants in other studies (Smallwood & Schooler, 2006) raises the question of whether or not the ability to minimize mind wandering or confine it to more appropriate times is a skill that can be selected, acquired, or even taught. By way of selection, Kane et al. (2007) found that subjects with higher working memory capacity were more successful in maintaining on-task thoughts when concentration and effort was needed. By way of training, we might consider using techniques to manage mind wandering similar to those used to manage fatigue (Rosekind et al., 1995).

In closing, we revisit the idea that monitoring is a task that requires the same tight-looped activity to be performed continuously, and where penalties for lapses can be severe. To accomplish this task, we often use a single human being who has limited attentional resources, who is easily interrupted by events both internal and external, and yet whose overall success may depend on all of that. In critical situations where lives are at stake, perhaps we do not have a good match here. But for as long as we continue to place humans in the monitoring role: in airline cockpits, beside swimming pools and baggage scanners, atop prison walls, and inside operating rooms, we should be grateful for the safety record we have today.

Acknowledgments

This work was internally funded by the National Aeronautics and Space Administration under the agency's Aviation Safety Program. Kevin Koehn assisted in analyzing the video data. Ed Hutchins provided valuable insights.

References

Ariga, A., & Lleras, A. (2011). Brief and rare mental "breaks" keep you focused: deactivation and reactivation of task goals preempt vigilance decrements. *Cognition, 118*(3), 439–443.

Casner, S. M. (2009). Perceived vs. measured effects of advanced cockpit systems on pilot workload and error: Are pilots' beliefs misaligned with reality? *Applied Ergonomics, 40*(3), 448–456.

Casner, S. M., Geven, R. W., Recker, M. P., & Schooler, J. W. (2014). The retention of manual flying skills among airline pilots. *Human Factors, 56*(8), 1506–1516.

Casner, S. M., & Schooler, J. W. (2014). Thoughts in flight: Automation use and pilots' task-related and task-unrelated thought. *Human Factors, 56*(3), 433–442.

deBettencourt, M. T., Cohen, J. D., Lee, R. F., Norman, K. A., & Turk-Browne, N. B. (2015). Closed-loop training of attention with real-time brain imaging. *Nature Neuroscience, 18*(3), 470–475.

Dillard, M. B., Warm, J. S., Funke, G. J., Funke, M. E., Finomore, V. S., Jr., Matthews, G., et al (2015). The sustained attention to response task (SART) does not promote mindlessness during vigilance performance. *Human Factors, 56*(8), 1364–1379.

Einstein, G. O., McDaniel, M. A., Pagan, J. L., & Dismukes, R. K. (2003). Forgetting of intentions in demanding situations is rapid. *Journal of Experimental Psychology: Applied, 9*(3), 147–162.
Finley, J. R., Benjamin, A. S., & McCarley, J. S. (2014). Metacognition of multitasking: How well do we predict the costs of divided attention? *Journal of Experimental Psychology: Applied, 20*(2), 158–165.
Harrell, W. A. (1999). Lifeguards' vigilance. Effects of child-adult ratio and lifeguard positioning on scanning by lifeguards. *Psychological Reports, 84*, 193–197.
Helton, W. S., & Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica, 129*(1), 18–25.
Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science, 18*(7), 614–621.
Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences, 36*(6), 661–679.
Loukopoulos, L., Dismukes, R. K., & Barshi, I. (2009). *The multitasking myth: Handling complexity in real-world operations*. Burlington, VT: Ashgate Publishing Company.
Mackie, R. R. (1987). Vigilance research – Are we ready for countermeasures? *Human Factors, 29*(6), 707–723.
Mackworth, J. F. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology, 1*, 16–21.
MacLean, K. A., Ferrer, E., Aichele, S. R., Bridwell, D. A., Zanesco, A. P., Jacobs, T. L., et al (2010). Intensive meditation training improves perceptual discrimination and sustained attention. *Psychological Science, 21*(6), 829–839.
Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia, 37*, 661–670.
NTSB. (1973). Aircraft accident report: Eastern Air Lines, Inc., Lockheed L-1011, N310EA, Miami, Florida, December 29, 1972. NTSB/AAR/73-14. Washington, DC: National Transportation Safety Board.
NTSB. (1994). Safety study: A review of flightcrew-involved, major accidents of U. S. air carriers, 1978 through 1990. NTSB/SS-94/01. Washington, D.C.: National Transportation Safety Board.
NTSB (2014). *Descent below visual glidepath and impact with seawall: Asiana Airlines Flight 214, Boeing 777–200ER, HL7742, San Francisco, California, July 6, 2013*. Washington, D.C.: National Transportation Safety Board.
Palmer, E. A., & Degani, A. (1991). Electronic checklists: Evaluation of two levels of automation. In *Proceedings of the Sixth International Aviation Psychology Symposium* (pp. 178–183). Columbus, OH.
Palmer, E. A., Hutchins, E. L., Ritter, R. D., & VanCleemput, I. (1991). Altitude deviations: Breakdowns of an error tolerant system. NASA Technical Memorandum DOT/FAA/RD-92/7.
Rosekind, M. R., Smith, R. M., Miller, D. L., Co, E. L., Gregory, K. B., Webbon, L. L., et al (1995). Alertness management: Strategic naps in operational settings. *Journal of Sleep Research, 4*, 62–66.
Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling, and the wandering mind. *Trends in Cognitive Science, 15*, 319–326.
Seli, P., Carriere, J. S. A., & Smilek, D. (2014). Not all mind wandering is created equal: Dissociating deliberate from spontaneous mind wandering. *Psychological Research*, 1–9. October 2014.
Singer, J. L. (1966). *Daydreaming: An introduction to the experimental study of inner experience*. New York: Random House.
Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek, M. D., et al (2011). Pupillometric evidence for the decoupling of attention from perceptual input during offline thought. *PLoS One, 6*, e18298.
Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132*(6), 946–958.
Smilek, D., Carriere, J. S., & Cheyne, J. A. (2010). Out of mind, out of sight: Eye blinking as indicator and embodiment of mind wandering. *Psychological Science, 21*(6), 786–789.
Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science, 12*(6), 462–466.
Sumwalt, R. (2003). Cockpit monitoring: Using procedures to enhance crew vigilance. *Professional Pilot*, 2–6. August 2003.
Szalma, J. L., Warm, J. S., Matthews, G., Dember, W. N., Weiler, E. M., Meier, A., et al (2004). Effects of sensory modality and task duration on performance, workload, and stress in sustained attention. *Human Factors, 46*(2), 219–233.
Teichner, W. H. (1972). Predicting human performance III: The detection of a simple visual signal as a function of time of watch. NMSU-ONR-TR-72-1. Las Cruces, NM: New Mexico State University.
Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. *Perspectives on Psychological Science, 10*(1), 82–96.
Tickner, A. H., Poulton, E. C., Copeman, A. K., & Simmons, D. C. V. (1972). Monitoring 16 television screens showing little movement. *Ergonomics, 15*(3), 279–291.
Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors, 50*(3), 433–441.