5-7-2019

Pilot Evaluations of a Non-Verbal Startle and Surprise Management Method, Tested During Airline Recurrent Simulator Training

Annemarie Landman

Eric L. Groen

Marc Frank

Gunnar Steinhardt

M M. van Paassen

See next page for additional authors

PILOT EVALUATIONS OF A NON-VERBAL STARTLE AND SURPRISE MANAGEMENT METHOD, TESTED DURING AIRLINE RECURRENT SIMULATOR TRAINING

Annemarie Landman,
TNO, Delft University of Technology, The Netherlands
Eric L. Groen
TNO, The Netherlands
Marc Frank
Luxair, Luxembourg Airlines
Gunnar Steinhardt
Cargolux International Airlines, Luxembourg
M. M. (René) van Paassen
Delft University of Technology, The Netherlands
Adelbert W. Bronkhorst
TNO, The Netherlands
Max Mulder
Delft University of Technology, The Netherlands

Aviation safety organizations have recommended that airline pilots are trained for startle and surprise. However, little information is available on useful training interventions. Therefore, a training intervention trial was executed during airline recurrent simulator training. The method consisted of a slow visual scan from the side-window, over the instruments, ending with facing the other pilot. Following a recorded video instruction, 38 airline pilots in two-pilot crews performed a training scenario in which they could apply the method. Data on application and evaluation of the method were obtained from each pilot. Few pilots actually applied the method (18.4%), and many gave low ratings to applicability of the method in the scenario, as well as in operational practice. Results show that a startle management method, as well as manner in which it is trained, should be carefully evaluated before being implemented in training practice.

Aviation safety organizations have recently recommended that pilot training should include specific means to deal with startle and surprise. Although startle is commonly used to designate both startle and surprise, strictly seen, startle refers to a reflexive stress response, whereas a surprise occurs when information is encountered that does not fit within one's mental model of the situation (Rivera et al., 2014). If both are experienced simultaneously, there needs to be an adjustment of the mental model under high stress, which can be very difficult (Landman et al., 2017). This may result in panic, cognitive lockup and total confusion. Training interventions that have been proposed include teaching pilots certain actions to "break out" of this state. An example of this would be a checklist specifically focused on relaxation, problem analysis and decision-making. The application rate of such a method was high in an experimental setting and pilots generally appreciated it, however, some also indicated that it was too distracting or complex (Landman et al., 2019).

The current study tested a simpler startle management method, consisting of a slow scanning motion of the head from the side window, over the instrument panel, ending with facing one the other pilot. The reasoning behind this method were as follows. First, it may help one consider the overall situation, including the other pilot's state, instead of immediately zooming in on the problem. Looking out the side window, which is also used in initial training and aerobatics, can be used to obtain natural sense of the aircraft's attitude. Second, the method buys time and prevents intuitive reactions to a problem that is not fully understood yet. In a similar manner, standard procedures to recover from spatial disorientation include a first step of recognizing and confirming the spatial disorientation, before attempting to recover (e.g., Previc & Ercoline, 1999). Third, performing a slow, conscious motion may instill a sense of control and stimulate goal-directed processing, as high stress is known to shift attentional control towards being more stimulus-driven (Eysenck, Derakhshan, Santos & Calvo, 2007). Potential advantages that this method may have compared to a checklist, are that it is faster, simpler, more active and more specific (compared to e.g. the command to "Observe"). The current paper describes an early-stage trial of this method, to obtain data on its application and perceived usefulness in a representative sample of airline pilots in a standard training setting.

Method

Participants

Data were collected of 38 B737 pilots (18 captains and 18 first officers) and 18 Bombardier Q400 pilots (9 captains and 9 first officers). One dataset of a captain was excluded as this person was involved in the intervention method design. For privacy reasons, no other personal data was collected. The experience level of the B737 pilots was generally higher than that of the Q400 pilots, with circa 2,500-25,000 hours compared to 600-12,000 hours. Pilots were informed that their data would be processed anonymously. They were also free to refrain from filling in the questionnaire, but there were no refusals.

Training intervention

The experiment took place during a recurrent simulator training session at Luxair, Luxembourg Airlines. The training intervention consisted of an 8-minute instructional video, in which a type rating instructor gave information about startle and surprise, and outlined the intervention method: 1) Turn your head to the outside shoulder, look out of the side window. 2) Turn your head back in a continuous movement, check your flight instruments. 3) Continue turning and see your colleague's flight instruments. 4) Continue turning and have a look at your colleague. 5) Now turn back and evaluate the situation. The total duration of the method can be under 10 seconds. The video demonstrated execution of the method from a first-person view in the cockpit.

Scenarios

The B737 training session was a Line Oriented Flight Training (LOFT), which consisted of a complete flight from Tenerife (TFS) to Luxembourg (LUX). In cruise, the crew received warnings from ATC about an explosive device being on board. Sharing workload with the first officer, the commander would need to order a search and prepare the cabin for descent. During descend, the device would trigger, causing an elevator runaway. Since the explosion in the B737 scenario was most startling, this scenario was expected to be the most suitable for applying the startle management method.

The Q400 training session consisted of practicing several flight situations. The scenario that was used for the experiment involved a double engine malfunction, one after the other. The standard procedure in this case would not be adequate, as it would cause both engines to be shut down simultaneously. The inadequacy of standard procedures was expected to be surprising and stressful.

Dependent measures

During the debriefing of the training session, the pilots filled in questionnaires, which were collected in sealed envelopes. As a manipulation check of the scenarios, the following variables were rated on a 1-5 scale, ranging from *very little* (1) to *very much* (5): Surprise by the ATC warning (B737) or engine malfunctions (Q400) and Startle by the device explosion (B737) or engine malfunctions (Q400). Anxiety following the events was rated on a 10 cm horizontal visual-analogue scale ranging from *none at all* to *maximum* (Houtman & Bakker, 1983). Mental demand and perceived time pressure following the ATC message (B737) or engine malfunction (Q400) were rated on the NASA-TLX mental demand and temporal demand subscales (Hart & Staveland, 1988). Finally, pilots also indicated whether they were informed by colleagues about the events in the scenario.

Next, pilots were asked if they applied the training intervention during the scenario. If confirmed, they were asked at which moments they applied it, and to what extent they felt that it helped them, as rated from *very little* (1) to *very much* (5). On a similar scale, all pilots rated how useful the method would be in operational practice. If pilots did not apply it, they indicated if this was mainly because they forgot, because they didn't find it applicable to the situation, or because they used a different method to manage their startle.

Results

Manipulation check

The manipulation check shows that pilots found the scenarios moderately surprising and stressful, scoring on average around the midpoint on the scales (Table 1 and 2). It is interesting that startle and surprise scores spread from the lowest to the highest endpoints, indicating that pilots may experience the same scenario very differently. Anxiety levels are similar between the groups, while the Q400 group reported somewhat higher surprise and the B737 group more startle. In the B737 group, 45 % (17) of the pilots were informed about the scenario, whereas 54 % (20) were not, and one skipped the question. The Q400 pilots all reported not being informed.

Table 1.
Pilots' subjective experience of the B737 bomb threat scenario.

	Mean (*SD*)	**Minimum**	**Maximum**
Startle (explosion) (1-5)	3.05 (1.21)	1	5
Surprise (message) (1-5)	3.11 (1.13)	1	5
Mental demand (message) (5-100)	51.7 (16.1)	15	75
Time pressure (message) (5-100)	57.2 (20.0)	20	95
Anxiety (message) (0-10)	4.5 (2.3)	0.0	7.5
Anxiety (explosion) (0-10)	5.1 (2.3)	0.0	10.0

Table 2.
Pilots' subjective experience of the Q400 double engine malfunction scenario.

	Mean (*SD*)	Minimum	Maximum
Startle (1-5)	2.61 (.92)	1	4
Surprise (1-5)	3.33 (.91)	1	5
Mental demand (5-100)	58.6 (17.2)	35	100
Time pressure (5-100)	54.4 (17.2)	25	90
Anxiety (0-10)	5.1 (1.9)	1.7	8.0

Application of the startle management method

In the B737 group, 9 out of 38 pilots (24 %) applied the method in the scenario. Eight when the explosion occurred, and one as an extra scan to check for issues. Of those not applying the method, most indicated that they forgot (37 %), or found it not applicable (37 %). Others reported they used a different method to manage startle (26 %).

In the Q400 group, 4 out of 18 pilots (22 %) applied the method in the scenario. Most pilots did not find in applicable in the scenario (56 %), some forgot (16.7 %) and one used a different method (5.6 %). All in all, the application rate of the method was low and it was similar in the different scenarios.

Perceived usefulness of the startle management method

The perceived usefulness of the method in the scenarios is shown in Figure 1. As can be seen in the figure, there were many in the B737 group who rated the method of very little use, whereas those in the Q400 group rated it little to moderately useful.

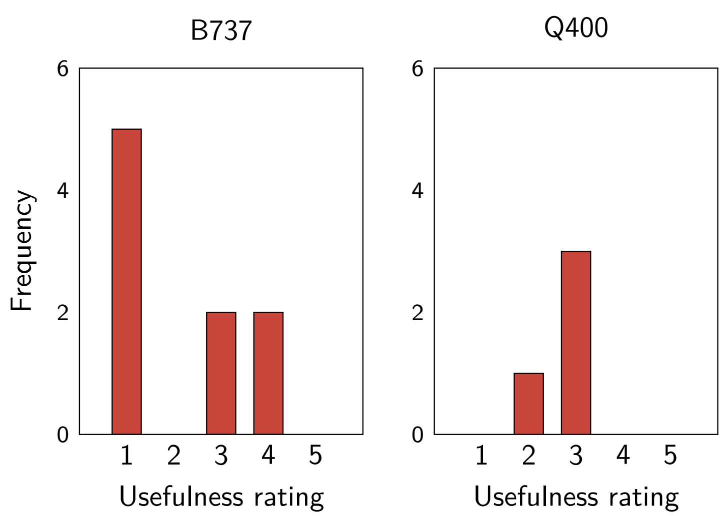

Figure 1.
Pilots' perceived usefulness of the startle management method in the scenario. Only those who applied the method are included.

The perceived usefulness of the method in operational practice is shown in Figure 2. It was similar to the ratings of usefulness in the scenario. Although the Q400 group seemed a little more positive towards the method, both groups included a relatively large proportion of pilots who rated the method of “very little” or “little” use in operational practice.

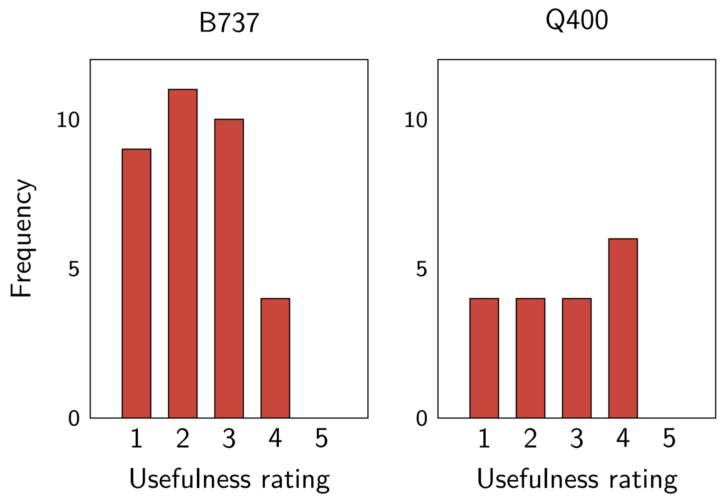

Figure 2.
Pilots’ perceived usefulness of the startle management method in operational practice.

Discussion

This experiment shows that pilots found the tested startle management method generally of little to moderate use in the scenarios and in operational practice. It is most notable that the intervention had a very low application rate (22-24 %), and a relatively large proportion of pilots (22-24 %) rated the method to be of very little use in practice. This is in contrast to a checklist-based startle management method, which was applied by all, and rated, on average, to be of very little use by 8 % of the participating pilots (Landman et al., 2019). There are some methodological aspects of the current study that may have caused lower ratings and application compared to the other study. First, there was very little time available in the experimental training session to explain the reasoning behind the method. With more time, the reasoning behind the method can be explained and there would be more room for discussing the method beforehand, which may improve the pilots’ openness towards it. Second, the tested startle management method was introduced during a mandatory training session. On the one hand, this mandatory setting could make the pilots more resistant to accepting the method. On the other hand, the current sample group is more representative of the general pilot population, compared to a group who participated in an experimental study based on invitation (Landman et al., 2019). Third, many pilots, especially those in the Q400 engine malfunction group indicated that the scenario was not startling enough for the method to be applicable. One remarked that it might be more useful “in cruise, when not mentally prepared for a malfunction, as we are in the simulator.”

Besides the manner in which the instructions of the method were given, there are some aspects of the method itself that can be adapted to improve it. First, most pilots who applied the method in the experiment, applied it together with their fellow pilot, indicating that if one pilot takes the initiative to execute the method, the other pilot is likely to join. The application rate might thus be improved by adding a callout at the start (e.g., "Let's do a scan"). Second, pilots indicated that they particularly experienced looking out of the side window as unhelpful. Some remarked that they thought it would be disorienting in-flight; that it seemed senseless; and that it caused them to lose time. Some of these objections can potentially be tacked with an explanation of the purpose behind the "senseless" and counter-intuitive actions. However, these objections may also indicate that the tested method may benefit from including actions that are more task-focused. Task-focus is known as an effective coping mechanism against performance stress (see e.g., Baumeister 1984, Matthews, Hillyard, & Campbell, 1999). Consciously working on (part of) a solution to the stressful problem, even if that means systematically gathering information or simplifying the situation, may give a sense of control and instill confidence. Perhaps placing more emphasis on a structured scan of the instruments and checking verbally with one's fellow pilot would improve the acceptance and effectiveness of the method.

In conclusion, whereas the current experiment had a strong practical approach, this made it difficult to accurately measure pilots' evaluations of the method. In order to obtain a more accurate picture, pilots could be tasked with executing and evaluating a method in a more experimental setting. Also, the experiment shows the importance to reserve time and resources for the development, training and testing of a startle management method, so that the end product is an effective method that pilots will apply in practice.

References

Baumeister, R. F. (1984). Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of personality and social psychology*, *46*(3), 610-620.

Eysenck, M. W., Derakhshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion*, *7*(2), 336-353.

Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *In P. A. Hancock and N. Meshkati (Eds.) Human Mental Workload. Amsterdam: North Holland Press.*

Landman, A., Groen, E. L., Van Paassen, M. M., Bronkhorst, A. W., & Mulder, M. (2017). Dealing with unexpected events on the flight deck: a conceptual model of startle and surprise. *Human factors*, *59*(8), 1161-1172.

Landman, A., Van Middellaar, S. H., Groen, E. L., Van Paassen, M. M., Bronkhorst, A. W., & Mulder, M. (2019). Testing the applicability of a checklist-based startle management method in the simulator. *In Proceedings of the 20th International Symposium on Aviation Psychology*, Dayton, OH.

Matthews, G., Hillyard, E. J., & Campbell, S. E. (1999). Metacognition and maladaptive coping as components of test anxiety. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice*, *6*(2), 111-125.

Previc, F. H., and Ercoline, W. R. (Eds.), (2004). *Spatial Disorientation in flight.* Reston, VA: American Institute of Aeronautics and Astronautics.

Rivera, J., Talone, A. B., Boesser, C. T., Jentsch, F., & Yeh, M. (2014). Startle and surprise on the flight deck: Similarities, differences, and prevalence. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 58, No. 1, pp. 1047-1051). Sage CA: Los Angeles, CA: SAGE Publications.