ORIGINAL ARTICLE

Now you see it, now you don't: dynamism amplifies the typicality effect

Richard Clewley[1] · **Jim Nixon**[1]

Abstract

Some safety events do not stabilise in a coherent state, presenting with transient or intermittent features. Such dynamism may pose problems for human performance, especially if combined with non-typical stimuli that are rarely encountered in everyday work. This may explain undesirable pilot behaviour and could be an important cognitive factor in recent aircraft accidents. Sixty-five airline pilots tested a real-world typicality gradient, composed of two cockpit events, a typical event, and a non-typical event, across two different forms of dynamism, a stable, single system transition, and an unstable, intermittent system transition. We found that non-typical event stimuli elicited a greater number of response errors and incurred an increased response latency when compared to typical event stimuli, replicating the typicality effect. These performance deteriorations were amplified when a form of unstable system dynamism was introduced. Typical stimuli were unaffected by dynamism. This indicates that dynamic, non-typical events are problematic for pilots and may lead to poor event recognition and response. Typical is advantageous, even if dynamic. Manufacturers and airlines should evolve pilot training and crew procedures to take account of variety in event dynamics.

Keywords Dynamic events · Typicality effect · Flight safety · Pilot response

1 Introduction

Everybody experiences intermittent problems with technology. From washing machines to smart-phones, problems and malfunctions do not always rest and stabilise in an intelligible state. Systems may remain in dynamic transition, appearing to work one moment, only to fail the next. In such cases we may resort to a repair behaviour common in the digital world, turning the object off, then back on. But what if that were not possible and it was necessary to capture the precise category of failure under these dynamic conditions?

1.1 Dynamic events

Humans are sensitive to transitional, dynamic information—we can detect change (Freyd 1987). Dynamic events change

state over time, regardless of human input (Cellier et al. 1997). Events have boundaries, they begin and they end. A thunderstorm starts with the first flash of lightening and perhaps ends with a distant rumble of thunder. The perceptual system is geared to detecting boundaries, and this has led to the idea that events are dynamic objects, bounded by discontinuities (Miller and Johnson-Laird 1976; Zacks and Tversky 2001; Clark 2013). Spatial discontinuities define the location and temporal discontinuities define the beginning and the end of an event (Zacks and Tversky 2001). Understanding the temporal structure of events helps us organise action and recall past similar events (Zacks et al. 2001, 2007; Speer and Zacks 2005). Without boundaries, to segment perceptual signals into tractable temporal units, we would experience the world as a continuous flow (Sargent et al. 2015).

Pilots need to capture event information, not be swept away in a river of continuous change that may threaten safe operation. Pilots are skilled at recognising and managing events and malfunctions, using assistive technologies that display a degree of diagnostic information and may suggest remedial actions (for instance the Electronic Centralized Aircraft Monitor [ECAM] on Airbus aircraft and Engine Indicating and Crew Alerting System [EICAS] on Boeing

✉ Richard Clewley
Richard.Clewley@cranfield.ac.uk

Jim Nixon
Jim.Nixon@cranfield.ac.uk

1 Safety and Accident Investigation Centre, Cranfield University, Cranfield MK43 0AL, UK

aircraft) (Ephrath and Young 1981; Thompson 1981). Pilots must identify events with enough precision so that the correct response protocol is followed. The real-world, however, can display a variety of dynamic events not captured in operating manuals or training encounters (Loukopoulos et al. 2009). For example, *periodic episodes* of stimuli, where cues do not show a continuous, intelligible system state or trend. This was experienced by a Virgin Atlantic Airbus A330 crew who received fifteen spurious cargo smoke warnings in 28 min, ranging in duration from 1 s to 173 s (see AAIB 2014). *Spike or transient indications*, where cues rapidly rise and decay, temporarily jump to extreme values or vacillate from normal to abnormal states. Such fluctuations can impinge on numerous systems, depending on the functions of the data, as experienced by a Qantas Airbus A330 crew when faulty data caused the flight control system to demand a descent (see ATSB 2011). *Unusual dynamism,* where cues are in rare dynamic configurations or combinations, as experienced by the crew of an Asiana Boeing 777 that crashed on landing following an approach without normal instrument aids and automation, combined with engine thrust and flight path parameter exceedances (see NTSB 2014). Each of these cases challenges the idea of a stabilised, coherent event, with boundaries that help the crew tame dynamism.

1.2 The typicality effect

Previous research suggests typicality may mediate pilot response to flight safety events (Clewley and Nixon 2019, 2020). Typical stimuli are known to be cognitively advantageous (Rosch et al. 1976). For the category '*primate*', the *chimpanzee* is a better, more typical member, than the *mongoose lemur*. Variations in rated typicality are known as gradients, and they have been reliably demonstrated across a range of categories (Rosch 1978; Barsalou 1987; Dry and Storms 2010). Typical category members are more rapidly verified, easily recognised, and readily learnt. This is known as the typicality effect (Rosch et al. 1976). Benefits are particularly pronounced in cases proximal to the prototype, the clearest and best cases, that act as cognitive reference points (Rosch 1975). Alongside subjective ratings and behavioural evidence of typicality effects, electrophysiological data show typical stimuli receive preferential processing (Lei et al. 2010; Wang et al. 2016) and display specific neural signatures in distinct brain regions (Iordan et al. 2016).

The typicality principle has been exploited in a variety of real-world contexts, including medical diagnosis (Dore et al. 2012), understanding sematic impairment following brain injury (Sandberg et al. 2012) and the influential recognition primed decision (RPD) model from the naturalistic decision making paradigm (Klein 1993). The RPD model suggests prototypical instances of situations allow actors to rapidly implement responses, making elaborate evaluation unnecessary (Klein 1998), accounting for efficient decision making seen in real-world, dynamic environments like firefighting (Klein et al. 2010).

Typicality confers cognitive advantage that translates to optimal behaviour. The corollary of this advantage, cognitive *dis*advantage for the non-typical, has received little attention (Clewley and Nixon 2021), but could offer better visibility of risk in human systems. Clewley and Nixon (2020) have recently described typicality gradients in the cockpit and view them as being potential proxies of cognitive (dis)advantage. Responses to typical flight safety events may be safer. For example, they describe a significant typicality gradient for aircraft fuel system events, locating candidate events for typicality effects. Fuel imbalances provide a typicality dividend, while fuel leaks do not, so risk poor recognition and response (Clewley and Nixon 2020).

1.3 Dynamic, non-typical events and contextual complexity

Dynamism places significant cognitive demands on prediction of future states and the planning of response steps (Zacks et al. 2001; Zacks et al. 2007; Clark 2013). If any system is in transition, the human will be challenged to describe, define, and forecast the 'next state'.

Dynamism experienced in the cockpit, such as intermittent, unstable event cues, is demanding, and may be augmented by the cognitive disadvantage of non-typical cues. Non-typical, dynamic flight safety events appear to pose particular problems for pilots. This accords with the idea of the complex problem space; multiple, dynamic events creating uncertainty (Walker et al. 2010). It is also supported by the notion of tractability (Hollnagel 2012). Tractable systems remain stable during description, in contrast to intractable systems, which continue to change *during* system description (Hollnagel 2012). This instability is fundamental to problematic dynamism. The rate of change, or degree of dynamism, is a form of contextual complexity, and there is evidence of its signature in recent aircraft accidents.

1.4 Dynamic instability in the real world

The signatures of contextual complexity are found in real-world accidents. In June 2009, an Air France Airbus A330, operating flight AF447, crashed into the Atlantic Ocean killing all 228 passengers and crew (BEA 2012). Ice had accumulated on speed sensors and made cockpit airspeed indications unreliable. In the first 99 s of the event there were approximately thirty system transitions, involving a wide array of cockpit indications, including the flight director guidance system, automatic thrust mode, warning tones and oral messages, flight control law changes, auto-flight mode annunciations and a mixture of reliable and unreliable

airspeed data (see BEA 2012, pp. 60–62 for graphical summary). Many of these indications were switching between credible and incredible readings; *now you see it, now you don't*.

The system state was changing during description and the crew were unable to adequately recognise the unreliable airspeed malfunction. Flight control inputs led to an aerodynamic stall and appropriate response protocols were not followed (BEA 2012). At least thirteen other flight crews, from five different airlines, had encountered in-flight airspeed events similar to AF447, and each case appeared to be mismanaged, lacking adequate diagnosis or response (BEA 2012).

Acute stress is known to adversely affect the cognitive performance of pilots (see NASA 2015, for a comprehensive review), and stress is a possible factor in the poor history of pilot management for these event types, especially as event-induced stress may be difficult to faithfully replicate in simulations. If dynamic states cannot be tracked and high value perceptual signals cannot be extracted, functional cognition could break down, perhaps leading to undesirable actions, where behaviour is not selected, but carried out instinctively (see BEA 2012, p. 174 for an explanation of an instinctive pull on the control stick to reduce speed if the crew suppose an overspeed is likely). Stressful, confusing, and dynamic proprioceptive cues, also difficult to replicate in simulations, could add further difficulty to event recognition, and undesirable actions are also more likely when pilots are surprised (Landman et al. 2017b; see EASA/NLR 2018, for an evaluation of training interventions).

The AF447 accident report describes two distinct signatures when discussing airspeed anomalies (BEA 2012). Firstly, erroneous indications may show a drop, followed by a levelling off at a failed value (BEA 2012). This 'classic system failure', involves a single transition to a new, stable, albeit failed, state. The key characteristic being the system rests in the degraded state and is thus coherent, recognisable and tractable (Flach 2012; Hollnagel 2012).

Secondly, erroneous indications may show intermittent drops, 'spiking' up or down, showing accurate or failed values, depending on *when* the indication was sampled (BEA 2012). This 'unstable system failure', experienced by the AF447 crew, involves intermittent, discontinuous system degradation. It is unruly and comprises multiple boundaries, as the 'object' appears, disappears, then reappears. This is the 'now you see it, now you don't' event structure. In the strict application of the definition of an event, this can be viewed as multiple events, and this has implications for the AF447 accident.

The features of the event seen by the crew in the accident did not sufficiently overlap with the previous training encounters relating to airspeed anomalies. For example, training encounters in a simulator for airspeed anomalies are unlikely to include spiking and may exhibit contrasting temporal characteristics to a real-world event. Different temporal characteristics between training and real-world scenarios can hinder recall of training (Speer and Zacks 2005; Zacks et al. 2007), and prevent identification of the category of failure for response (see Clewley and Nixon 2019).

The interaction between typicality and dynamism, a signature of the AF447 accident, could be an important form of complexity in the cockpit, and an overlooked cognitive factor in aircraft accidents. Recent accidents involving the Boeing 737-MAX have brought further scrutiny on pilot recognition of malfunctions involving dynamic cues and sophisticated aircraft technology (AIB 2019; JATR 2019; KNKT 2019), underlining dynamism as a contemporary problem. Indeed, the test pilots in the Boeing 737-MAX certification programme had trouble responding to events that later proved beyond the capability of well-trained pilots (United States House Committee on Transportation and Infrastructure 2020). Further understanding may help evolve pilot training and response procedures, so that resilience and recovery are more easily achieved in the cockpit.

The aim of this research is to extend the typicality effect to the real-world dynamic task of pilot event recognition. We draw the same distinction as the BEA (2012) and operationalise two forms of dynamism: dynamism seen in a classic system failure involving a single transition, and dynamism in an unstable system failure involving multiple transitions. Typicality can be operationalised through gradients. This leads to the development of our hypotheses: We test a real-world typicality gradient comprising two cockpit events across the two levels of dynamism. Firstly, we predict a non-typical event will cause a greater number of response choice errors and elicit a greater response latency, when compared to a typical event. Secondly, we predict a high dynamism event will cause a greater number of response choice errors and elicit a greater response latency, when compared to low dynamism event. Finally, we predict event typicality and event dynamism will produce an interactive effect.

2 Method

2.1 Participants

Sixty-five airline pilots participated in the study. All participants worked at the same European short-haul airline on the same aircraft type. The sample comprised 33 Captains, 32 First Officers, mean age 36.18 years (SD = 7.7), 4 females and 61 males. Median flying experience 4250 h,

Table 1 Summary of experimental conditions

Typicality	Dynamism	
	Low	High
High	Typical/stable	Typical/unstable
Low	Non-typical/stable	Non-typical/unstable

range = 350–11,100 h. Median flying experience on the aircraft type 2000 h, range = 110–8000 h.

2.2 Design

A 2 × 2, fully within-subjects design was used. There were two factors, typicality and dynamism, each with two

levels, high and low (Table 1). The order of experimental conditions and response choices were randomized.

2.3 Independent variables

Typicality was operationalised at two levels, high (typical) and low (non-typical), to create a typicality gradient. We consulted a senior management Captain at the host airline to act as a subject matter expert (SME) on flight safety events. The SME proposed candidate event cues at the level of 'typical' and 'non-typical', according to Clewley and Nixon (2019). The ice protection event cue 'DEICE PRESS' fulfilled the criteria for 'typical' (see Fig. 2, below, for event stimuli); it is seen regularly in flight operations and frequently features in Safety Management System (SMS) data capture (e.g. crew reports). The event cue signals deice

Fig. 1 The two dynamism conditions. The dots indicate system transitions, from normal to event or event to normal
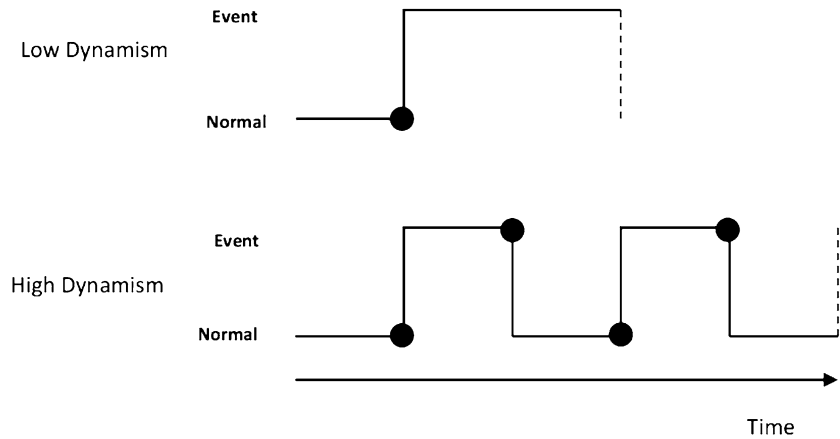
Fig. 2 Experimental stimuli. Event 'DEICE PRESS' was typical, 'ATT 1' was non-typical. Each were presented in two dynamic states, low dynamism (stable transition) and high dynamism (unstable transition)

system pressure is low. The instruments/auto-flight event 'ATT 1' fulfilled the criteria for 'non-typical'; it is rarely encountered in everyday work and seldom features in SMS data capture. The event cue signals abnormal configuration of the attitude/heading reference system after system failure. Both events feature in the evidence-based training matrix for large public transport aircraft (IATA 2013). This training matrix guides operators to develop pilot competencies in event management. The 'ATT 1' event falls under: "… System failures that require monitoring and management of the flight path using *degraded or alternative displays*" (IATA 2013, p. 115, author italics). The 'DEICE PRESS' event falls under "…Thunderstorm, heavy rain, turbulence, ice build up to include de-icing issues" (ibid. p. 111). The visual stimuli for each event were captured from actual aircraft systems.

Dynamism was operationalised to replicate a single, stable system transition (low dynamism), and discontinuous, intermittent system behaviour (high dynamism), reflecting the two signatures discussed in the AF447 report (BEA 2012). For both levels of dynamism, the event stimuli were dynamically animated and presented for 2 s (see Fig. 2, for event stimuli). In the low dynamism condition this comprised a single transition followed by 2 s of continuous presentation of the event stimuli (Fig. 1, top). In the high dynamism condition this comprised normal system indications punctuated with two separate episodes of the event stimuli, both of 1 s duration, creating four transitions (Fig. 1, bottom).

2.4 Dependent variables

Response choice accuracy was scored according to whether participants correctly identified the event stimuli from a list of six response choices. The list of six comprised the correct response, two alternative system states from the same section of the host airline expanded checklist and three unrelated event cues. The SME advised that this reflected the flight crew task with respect to recognition accuracy, checklist use and response selection.

Response latency was measured in seconds and was defined as the time taken to complete the event identification. The identification task was timed to replicate the temporal constraints seen in flight crew tasks. Participants were given 15 s to complete the task and a countdown timer was displayed in the bottom left-hand corner of the screen.

Rated typicality was measured on a 9-point scale, using the anchors 'not at all' (1) and 'very' (9). This is an established approach to measuring typicality (Barsalou 1987; Rothbart et al. 1996).

2.5 Materials and procedure

The research protocol was approved by the University Ethics Committee. Experimental materials were delivered by the Qualtrics survey platform (Qualtrics, Provo, Utah, USA). Response latency (time taken to complete the task) was captured by the software. Each respondent was sent a link to the Qualtrics survey platform and completed the research tasks on-line after giving informed consent.

Participants were presented with event stimuli, immediately followed by a list of six response choices. The on-screen instructions asked participants 'What did you see?'. Participants were instructed to move their choice to a box labelled 'My answer'. After the experiment we collected rated typicality data for the event stimuli so that we could construct a typicality gradient. Screenshots of the two event stimuli and the response display are shown in Fig. 2.

2.6 Data analysis

Response accuracy was binary (correct/incorrect), so we used Cochran's Q to report main effects and McNemar's Test for post-hoc comparisons. Alpha was Bonferroni corrected to limit type 1 error inflation. We used repeated measures analysis of variance (IBM SPSS version 25) to test for differences in response latency across the four conditions. Effect size was reported using n_p^2. For the typicality ratings we used paired t test to compare means, reporting effect size using r (Cohen 1992). An alpha of < 0.05 was considered significant.
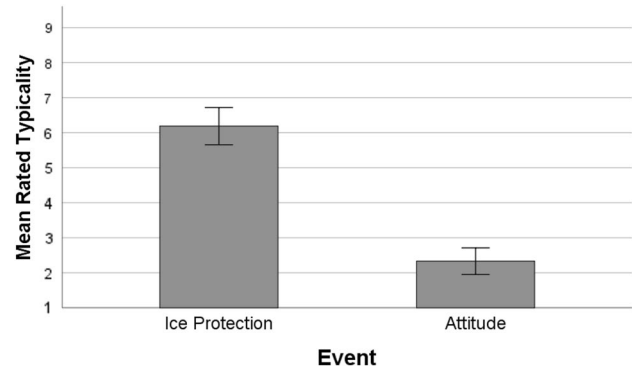
Fig. 3 The typicality gradient tested in this experiment, depicting two safety events. The 'Ice Protection' event served as the typical stimuli; the 'Attitude' event served as the non-typical stimuli. (95% confidence interval)

3 Results

3.1 Typicality gradient

Figure 3 shows the mean rated typicality for the two safety events used in the experiment. The ice protection event (M = 6.2, SD = 2.0) was rated more typical than the attitude event (M = 2.3, SD = 1.4) [t (55) = 12.41, $p < 0.001$, $r = 0.75$]. This is the typicality gradient we tested.

3.2 Response accuracy

We found significant differences in response accuracy (Q (3) = 78.10, $p < 0.001$, N = 65), as depicted in Fig. 4. Response accuracy for the typical event was equal in low and high dynamism conditions (98.47%). Post hoc comparison, using Bonferroni corrected McNemar's Test, revealed response accuracy for the non-typical event was significantly better in the low dynamism condition (64.62%) than high dynamism condition (41.54%) [$p = < 0.005$, N = 65].

3.3 Response latency

We found significant differences in response latency (Fig. 5; Table 2). Typicality showed a significant main effect (F (1, 64) = 133.21, $p < 0.001$, η_p^2 0.68). Dynamism showed a significant main effect (F (1, 64) = 38.63, $p < 0.001$, η_p^2 0.38). There was also a significant interaction effect (F (1, 64) = 26.42, $p < 0.001$, η_p^2 0.29). Non-typical/high dynamism event stimuli elicited the greatest mean response latency (M = 11.34 s, SD = 1.12 s), while typical/low dynamism event stimuli elicited the shortest response latency (M = 6.96 s, SD = 2.65 s).
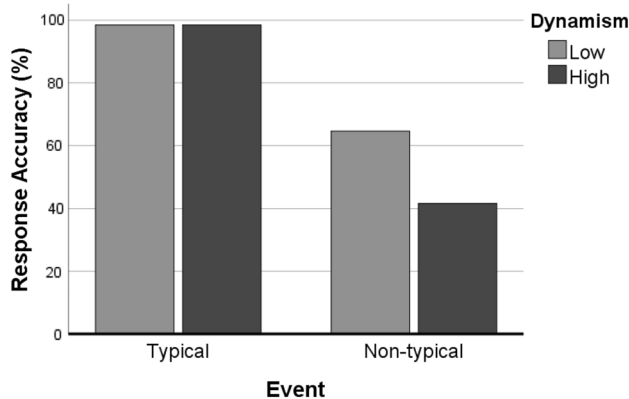
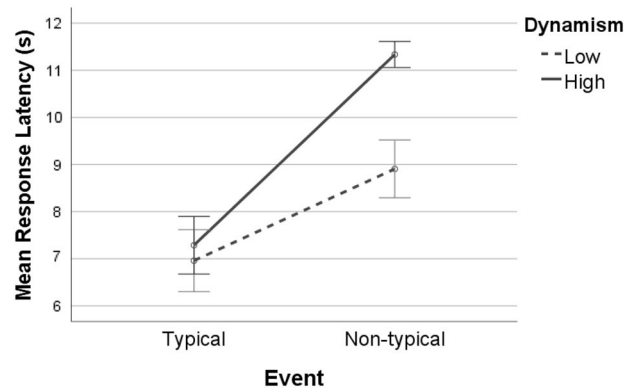**Fig. 4** The response accuracy for the four conditions

Fig. 5 The mean response latency (s) for the four conditions, showing an interaction effect between typicality and dynamism (95% confidence interval)

3.4 Summary

Overall, pilots suffered a decline in performance when confronted with non-typical event stimuli. Response accuracy declined and response latency increased. This typicality effect is amplified when the stimuli are presented in a dynamic, intermittent form, so pilots suffer further decrement of response accuracy and additional increases in response latency. Conversely, for typical stimuli, recognition is stable despite system dynamism and only minor processing delays are incurred.

4 Discussion

The present study tested a real-world typicality gradient, composed of two cockpit events, across two different forms of dynamism: a single, low dynamism transition, and an unstable, high dynamism system transition. We have found that non-typical event stimuli elicit a greater number of response errors and incur an increased response latency when compared to typical event stimuli, replicating the typicality effect and supporting our first hypothesis. These performance deteriorations were amplified when a form of system dynamism was introduced, indicating dynamic, intermittent event cues could be problematic in the cockpit when

Table 2 The mean, standard deviation and 95% confidence intervals for response latency (s) in each of the conditions

Condition	Mean (SD)	95% CI
Typical/dynamism low	6.96 s (SD = 2.65)	6.30–7.62
Typical/dynamism high	7.29 s (SD = 2.47)	6.68–7.90
Non-typical/dynamism low	8.91 s (SD = 2.48)	8.29–9.52
Non-typical/dynamism high	11.34 s (SD = 1.12)	11.06–11.61

combined with non-typical stimuli. Cognitive performance with typical event stimuli appears to remain intact, despite system dynamism. Pilots are subject to typicality effects and these appear amplified during dynamic event encounters, supporting our hypothesis that dynamism and typicality exhibit an interactive effect.

This study supports the axiom that pilot knowledge tends to be concentrated around typical events (Clewley and Nixon 2019, 2020). Dynamism had no notable effect on pilot behaviour for the typical event, indicating typicality may provide a protective cloak; cognition is geared to typicality (Clewley and Nixon 2020). Improving the quality of pilot exposure to non-typical events remains an important strategy to mitigate the typicality effect. Typical stimuli receive preferential processing (Lei et al. 2010; Wang et al. 2016), and in the cockpit this appears to be a mechanism that can tame dynamism, leading to preferred response. For typical events, dynamism does not degrade cognitive performance.

Dynamic, unstable event features carry risk of delayed and inadequate recognition. Pilots in this study faced with a high dynamism, non-typical event took longer to make poorer recognition choices, and this event structure exhibited an interactive effect. In real-world encounters this could promote two types of situations. Firstly, pilots may not adequately recognise event stimuli, making it less likely that appropriate checklists and procedures are carried out. The accuracy of just 41.54% for the non-typical, high dynamism condition indicates clear cognitive problems for pilots in establishing basic event verification. This would seem to be particularly relevant to cases where the key event features do not remain extant for prolonged periods, as was the case for the AF447 crew (BEA 2012). Unstable, intermittent event features, that change during system description, may present significant cognitive challenges.

Secondly, dynamic, non-typical events induce a greater response latency, and in real-world encounters this could lead to unacceptable delays in desirable pilot behaviour, such as flight path interventions or response selection. Additionally, events could escalate during these delays. In this study we tested two short bursts of dynamism, which we consider to be mild in comparison to the AF447 accident, where multiple systems remained in dynamic states for the first 3 min (BEA 2012). Dynamic, non-typical event stimuli carry risk in the cockpit, and the interactive effect explains undesirable pilot behaviour seen in some aircraft accidents.

We suggest pilots receive education and training on the temporal variety of dynamic events. It is unclear whether pilots currently receive training to manage cues and indications that do not stabilise in a coherent state. Such dynamic, contextual complexity may be met for the first time in a high risk, real-world event. This is particularly prescient given that recall of similar, past events, important in pilot response, may be damaged if a starkly different temporal structure is met (Zacks et al. 2001). The extent to which a training encounter has the same temporal characteristics as a real-world event may be important in some cases. If an event does not look like training, the training may not be recalled.

Dynamic variety could be introduced into aircraft 'type ratings' (training to fly a specific aircraft) and pilot recurrent simulator training. Both of these forums have acknowledged limitations, such as providing brief, contrived or predictable event examples (Casner et al. 2013; Clewley and Nixon 2020). System spikes, episodes of intermittent system indications and failures that do not stabilise in a coherent state (*dynamic variety*) could be added to EASA Part-FCL, Subpart H, Section 1, AMC1 FCL.725(a), requirements for the issue of class and type ratings; Section 2, AMC2 FCL.735.A, multi-crew cooperation training courses (aeroplanes), systems abnormal and emergency operations; and GM1 to Appendix 9 of Annex I, Training, skill test and proficiency check for MPL, ATPL, type and class ratings, and proficiency check for IRs (see EASA 2020). Pilots may then have knowledge of, and simulator experiences relating to, a variety of dynamic event characteristics. This is a shift towards educating pilots about the *anatomy* of events and event dynamics.

Additional resilience could be built into the cockpit through better pilot materials, checklists and procedures, that offer improved guidance on intermittent, transient cues, for example. Training a single system transition may not prepare pilots for intermittent cues. In this research we introduced two periods of event stimuli in the high dynamism condition, and that generates *four* system transitions. We believe that pilot recognition and response could be improved if flight crew education explicitly trains the three types of dynamism we have identified in Sect. 1.1, above: *periodic episodes*, *spikes or transient cues* and *unusual dynamism*. Real world events will continue to present these characteristics, they will continue to challenge the predictive faculties of humans. As currently trained, some real-world events may be beyond pilot knowledge.

Dynamic events pose clear problems for frame selection, complementing recent work on startle and surprise (Landman et al. 2017a). This approach suggests new system states require revised 'frames', or knowledge structures, to guide processing and provide context and meaning. The model specifies 'surprise', which requires cognitive effort to select a new, appropriate frame. Associated stress can also affect performance, leading to wide ranging effects on cognition, especially attentional tunnelling (see Vanderhaegen et al. 2020, for a contemporary view of attentional resources in dynamic events and heartbeat synchronisation).

In a highly dynamic event, the frame 'mis-match' may endure as the system state vacillates. This is a possible explanation for the delay and effort seen in reframing event stimuli, as multiple event transitions require tracking. Dynamic states present a simple conundrum for the pilot or operator: which frame?

Naturally, this research has limitations. We tested one typicality gradient, composed of two cockpit events, and two forms of system dynamism. This serves as an initial platform and we think the principle could be extended to examine pilot performance on a greater range of events and dynamics. We used real cockpit cues, animated dynamically, but we feel this work would benefit from moving to a flight simulator environment to further validate the approach. These findings have potential applications in other complex, dynamic environments, such as medicine (Perry and Wears 2012), firearms events (Mitchell and Flin 2007), crowd and stadium disasters (Challenger and Clegg 2011) and firefighting (Grenfell Tower Inquiry 2017), where non-typical events can exhibit dynamic variety. Other dynamic activities that employ simulations (for example, see Crichton 2017, for a discussion on simulator exercises in drilling operations) could benefit from considering dynamic variety as a training variable. Dynamic events may not resemble trained for or anticipated events, and if paired with a non-typical situation, human performance may decline, compromising safety.

5 Conclusion

Some aircraft accidents, like the crash of flight AF447 (BEA 2012), involve non-typical events that fail to stabilise in a coherent, intelligible state. Such events present a deluge of intermittent cues: now you see it, now you don't.

In this article we have outlined the role of typicality and event dynamism in aircraft accidents. We have extended the typicality effect to a real-world dynamic task. This study has indicated it is important that pilots experience events as coherent, intelligible entities, not a continuous ebb and flow of change, so we have suggested improvements to pilot training and cockpit materials. Event boundaries are important in pilot response. We have presented evidence from experienced airline pilots that dynamism, when combined with non-typical stimuli, decreases response accuracy and increases response latency. Dynamism and typicality are axiomatic variables in aircraft accidents.

Declarations

Conflict of interest Author A (lead author) declares that he/she has no conflict of interest. Author B (secondary author) declares that he/she has no conflict of interest. The authors have no relevant financial or non-financial interests to disclose.

Ethical approval This article does not contain any studies with animals performed by any of the authors. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

AAIB (Air Accidents Investigation Branch) (2014) Aircraft accident report 1/2014, report on the accident to Airbus A330–343, G-VSXY, London Gatwick Airport, 16 April 2012. AAIB, London

AIB (Aircraft Accident Investigation Bureau of Ethiopia) (2019) Aircraft accident investigation preliminary report: B737-8 (MAX) registered ET-AVJ 28 NM South East of Addis Ababa, Bole International Airport. AIB

ATSB (Australian Transport Safety Bureau) (2011) In-flight upset—Airbus A330-303, VH-QPA, 154 km west of Learmonth, WA, 7 October 2008. ATSB, Canberra

Barsalou LW (1987) The instability of graded structure: implications for the nature of concepts. In: Neisser U (ed) Concepts and conceptual development: ecological and intellectual factors in categorization. Cambridge University Press, Cambridge, pp 101–140

BEA (Bureau d'Enquêtes et d'Analyses) (2012) Final report on the accident 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France flight AF447 Rio de Janeiro—Paris. BEA, Paris

Casner SM, Geven RW, Williams KT (2013) The effectiveness of airline pilot training for abnormal events. Hum Factors 55(3):477–485. https://doi.org/10.1177/0018720812466893

Cellier J-M, Eyrolle H, Marine C (1997) Expertise in dynamic environments. Ergonomics 40(1):28–50. https://doi.org/10.1080/001401397188350

Challenger R, Clegg CW (2011) Crowd disasters: a socio-technical systems perspective. Contemp Soc Sci 6(3):343–360. https://doi.org/10.1080/21582041.2011.619862

Clark A (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci 36(3):81–204. https://doi.org/10.1017/S0140525X12000477

Clewley R, Nixon J (2019) Understanding pilot response to flight safety events using categorisation theory. Theor Issues Ergon Sci 20(5):572–589. https://doi.org/10.1080/1463922X.2019.1574929

Clewley R, Nixon J (2020) Penguins, birds and pilot knowledge: can an overlooked attribute of human cognition explain our most

puzzling aircraft accidents?. Hum Factors. https://doi.org/10.1177/0018720820960877

Clewley R, Nixon J (2021) A new facet of category theory: cognitive disadvantage and its implications for safety in the cockpit. Int J Aerosp Psychol (In press)

Cohen J (1992) A power primer. Psychol Bull 112:155–159. https://doi.org/10.1037/0033-2909.112.1.155

Crichton MT (2017) From cockpit to operating theatre to drilling rig floor: five principles for improving safety using simulator-based exercises to enhance team cognition. Cogn Technol Work 19:73–84

Dore KL et al (2012) Influence of familiar features on diagnosis: instantiated features in an applied setting. J Exp Psychol Appl. https://doi.org/10.1037/a0026539

Dry MJ, Storms G (2010) Features of graded category structure: generalizing the family resemblance and polymorphous concept models. Acta Physiol (oxf) 133(3):244–255. https://doi.org/10.1016/j.actpsy.2009.12.005

EASA (2020) Easy access rules for flight crew licensing (Part-FCL) (1178/2011-Annex-1). EASA, Cologne

EASA/NLR (2018) Startle effect management: NLR-CR-2018-242. European Aviation Safety Agency

Ephrath AR, Young LR (1981) Monitoring vs. man-in-the-loop detection of aircraft control failures. In: Rasmussen J, Rouse W (eds). Proceedings of a NATO symposium on human detection and diagnosis of system failures. Plenum, New York, pp 143–169

Flach JM (2012) Complexity: learning to muddle through. Cogn Technol Work 14:187–197. https://doi.org/10.1007/s10111-011-0201-8

Freyd JJ (1987) Dynamic mental representations. Psychol Rev 94(4):427–438. https://doi.org/10.1037/0033-295X.94.4.427

Grenfell Tower Inquiry (2017) The Grenfell Tower Inquiry. https://www.grenfelltowerinquiry.org.uk/. Accessed 1 Feb 2019

Hollnagel E (2012) Coping with complexity: past, present and future. Cogn Technol Work 14(3):199–205. https://doi.org/10.1007/s10111-011-0202-7

IATA (2013) Evidence-based training implementation guide. International Air Transport Association, Montreal

Iordan MC et al (2016) Typicality sharpens category representations in object-selective cortex. Neuroimage 134:170–179. https://doi.org/10.1016/j.neuroimage.2016.04.012

JATR (Joint Authorities Technical Review) (2019) Boeing 737 MAX flight control system. JATR

Klein G (1998) Sources of power: how people make decisions. MIT Press, Cambridge

Klein G (1993) A recognition-primed decision (RPD) model of rapid decision making. In: Klein G et al (eds) Decision making in action: models and methods. Ablex, Norwood, pp 138–147

Klein G, Calderwood R, Clinton-Cirocco A (2010) Rapid decision making on the fire ground: the original study plus a postscript. J Cogn Eng Decis Mak 4(3):186–209. https://doi.org/10.1518/155534310X12844000801203

KNKT (Komite Nasional Keselamatan Transportasi) (2019) Final report: PT. Lion Mentari Airlines Boeing 737-8 (MAX); PK-LQP Tanjung Karawang, West Java, Republic of Indonesia 29 October 2018. KNKT, Republic of Indonesia

Landman A et al (2017a) Dealing with unexpected events on the flight deck: a conceptual model of startle and surprise. Hum Factors 59(8):1161–1172. https://doi.org/10.1177/0018720817723428

Landman A et al (2017b) The influence of surprise on upset recovery performance in airline pilots. Int J Aerosp Psychol 27(1–2):2–14. https://doi.org/10.1080/10508414.2017.1365610

Lei Y et al (2010) How does typicality of category members affect the deductive reasoning? An ERP Study. Exp Brain Res. https://doi.org/10.1007/s00221-010-2292-5

Loukopoulos LD, Dismukes RK, Barshi I (2009) The multitasking myth: handling complexity in real-world operations. Ashgate, Farnham

Miller GA, Johnson-Laird PN (1976) Language and perception. Harvard University Press, Cambridge

Mitchell L, Flin R (2007) Shooting decisions by police firearms officers. J Cogn Eng Decis Mak 1(4):375–390. https://doi.org/10.1518/155534307X264861

NASA (2015) Effects of acute stress on aircrew performance: literature review and analysis of operational aspects. NASA technical memorandum 2015-218930. NASA

NTSB (National Transportation Safety Board) (2014) Descent below visual glidepath and impact with seawall, Asiana Airlines Flight 214, Boeing 777-200ER, HL7742, San Francisco, California July 6, 2013. NTSB, Washington

Perry SJ, Wears RL (2012) Underground adaptations: case studies from health care. Cogn Technol Work 14:253–260. https://doi.org/10.1007/s10111-011-0207-2

Rosch E (1975) Cognitive reference. Cogn Psychol 547:532–547. https://doi.org/10.1016/0010-0285(75)90021-3

Rosch E (1978) Principles of categorization. In: Rosch E, Lloyd BB (eds) Cognition and categorization. Erlbaum, Hillsdale, pp 27–48

Rosch E, Simpson C, Miller RS (1976) Structural bases of typicality effects. J Exp Psychol Hum Percept Perform 2(4):491–502. https://doi.org/10.1037/0096-1523.2.4.491

Rothbart M, Sriram N, Davis-Stitt C (1996) The retrieval of typical and atypical category members. J Exp Soc Psychol 32(4):309–336. https://doi.org/10.1006/jesp.1996.0015

Sandberg C, Sebastian R, Kiran S (2012) Typicality mediates performance during category verification in both ad-hoc and well-defined categories. J Commun Disord 45(2):69–83. https://doi.org/10.1016/j.jcomdis.2011.12.004

Sargent JQ, Zacks JM, Bailey HR (2015) Perceptual segmentation of natural events: theory, methods, and applications. Camb Handb Appl Percept Res. https://doi.org/10.1017/CBO9780511973017.029

Speer NK, Zacks JM (2005) Temporal changes as event boundaries: processing and memory consequences of narrative time shifts. J Mem Lang. https://doi.org/10.1016/j.jml.2005.02.009

Thompson DA (1981) Commercial air crew detection of system failures: state of the art and future trends. In Rasmussen J, Rouse W (eds). Proceedings of a NATO symposium on human detection and diagnosis of system failures. Plenum, New York, pp. 37–48.

United States House Committee on Transportation and Infrastructure (2020) Final committee report: the design, development and certification of the Boeing 737 MAX. United States House Committee on Transportation and Infrastructure, Washington

Vanderhaegen F, Wolff M, Mollard R (2020) Non-conscious errors in the control of dynamic events synchronized with heartbeats: a new challenge for human reliability study. Saf Sci. https://doi.org/10.1016/j.ssci.2020.104814

Walker GH et al (2010) Translating concepts of complexity to the field of ergonomics. Ergonomics 53(10):1175–1186. https://doi.org/10.1080/00140139.2010.513453

Wang X et al (2016) Categorization method affects the typicality effect: ERP evidence from a category-inference task. Front Psychol 7(JAN):1–11. https://doi.org/10.3389/fpsyg.2016.00184

Zacks JM, Tversky B (2001) Event structure in perception and conception. Psychol Bull 127(1):3–21. https://doi.org/10.1037/0033-2909.127.1.3

Zacks JM, Tversky B, Iyer G (2001) Perceiving, remembering, and communicating structure in events. J Exp Psychol Gen 130(1):29–58. https://doi.org/10.1037/0096-3445.130.1.29

Zacks JM et al (2007) Event perception: a mind/brain perspective. Psychol Bull 133(2):273–293. https://doi.org/10.1037/0033-2909.133.2.273