Theoretical Issues in Ergonomics Science, Volume 20, Issue 5, 2019, pp. 572-589
DOI:10.1080/1463922X.2019.1574929

Understanding Pilot Response to Flight Safety Events Using Categorisation Theory

Richard Clewley and Jim Nixon

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

Corresponding author: r.m.clewley@cranfield.ac.uk

Understanding Pilot Response to Flight Safety Events Using Categorisation Theory

Categorisation theory explains our ability to recognise events in terms of a similarity overlap between either a prototypical, ideal case or a stored exemplar derived from experience. Evidence from aviation accident reports indicate that pilots are not always able to recognise flight safety events in real-time and this can lead to undesirable pilot behaviour. Flight safety events may not always arise in recognisable formats, especially as rare and unusual combinations are possible. Correspondence with prototypes or exemplars may be weak, creating borderline cases and harming recognition. In this article we extend categorisation theory to develop a new framework which characterises flight safety events. We model three case studies using the new framework to demonstrate how categorisation theory can be used to understand flight safety events of different types. Finally we propose a roadmap for future research and discuss how categorisation theory could be applied to training or the organisation of flight crew reference material to improve response to inflight events.

Relevance to Human Factors/Ergonomics Theory

Categorisation theory is established in cognitive psychology as a leading explanation of variability in the human capacity to recognise objects and events. Several recent air transport accidents appear to show flight crew have difficulty recognising some events, and this article applies this theoretical framework to analyse such failures. We account for borderline cases, and argue that typicality and familiarity gradients are important drivers of pilot behaviour. We provide a theory-driven account of why some events are not recognised.

Keywords: flight safety events; categories; concepts; typicality; recognition-primed decision making

1. Introduction

In June 2009 an Airbus A330, operating as Flight AF447, crashed into the Atlantic Ocean 4 minutes and 23 seconds after the first indication of a problem. The accident report states that the crash resulted from a succession of events, including the loss of airspeed data and an aerodynamic stall (BEA 2012). The crew were unable to make sense of the situation and the appropriate remedial action was not taken resulting in the loss of the aircraft (BEA 2012). This type of loss of airspeed data, caused by ice accumulation on sensors, had been experienced by at least thirteen other Airbus A330 and A340 crews, from five different airlines. Four crews did not recognise the event, two crews concluded it was another type of event ('inconsistent angle of attack data') and none of the crews linked the event to the correct remedial response (BEA 2012).

Some flight safety events do not present as clear cases which map exactly on to training and experience. Even for very experienced pilots the world can be poorly structured and confusing. In the simplest information processing terms, an event is not directly diagnosable by the cues found in the world, and response is impaired. Some situations will be familiar, others not so much. In this article we apply categorisation theory to understand pilot response to different situations. Categorisation theory proposes that recognising entities in the real-world requires a similarity overlap between either a prototypical, ideal case or a stored exemplar derived from experience. Insufficient correspondence with a prototype or exemplar could damage recognition and pilot response. In categorisation theory real-world entities vary in how central they are to a concept, with clear cases exhibiting centrality, whilst other cases occupy the margins (Barsalou 1985; Dry and Storms 2010). These marginal, borderline cases can affect appropriate diagnosis and response to in-flight events.

For pilots, dynamic environments are grouped into categories, or types of events, which they must link to an adequate response. This is articulated by the BEA, (2012) in their final report into the crash of the AF447:

> '…the crew must analyse and confirm the type of failure before undertaking any failure processing action…the "adequate reaction" expected of the crew supposes immediate memory items with the purpose of stabilising the situation, then recourse to action instructions

> available on the [cockpit display] and/or recourse to procedures explained in the [checklist] and classified by category of diagnosed anomaly.'
> (BEA, 2012, p.102).

Pilots are expected to diagnose and confirm the type event, recall any memory items to stabilise the flight, then refer to expanded procedures. Response tools, such as abnormal and emergency checklists, order known flight safety events into families based on shared features, so at a practical level flight safety events are grouped and indexed by categories decided by the manufacturer or aircraft operator. Such indexes are a form of taxonomic order, a central principle of categorisation and knowledge structure (Rosch 1978). Underpinning this process is recognition, which demands a similarity overlap with an event concept. It is this categorisation that drives behaviour; insufficient overlap with a concept can lead to delayed or inappropriate pilot response.

In this article we demonstrate the application of categorisation theory to better understand the variability seen in event recognition and pilot response. We review categorisation theory, identifying the unique contributions that the theory can make to understanding flight safety events. A new framework is developed which describes three conceptions of events and the characteristics in relation to categorisation theory. Finally three case studies are analysed using the framework and a roadmap for future research is proposed.

2. Categorisation theory

Categories are the reduction of sensory experiences into equivalent groups on the basis of shared attributes (Pothos and Wills 2011). Table 1 contains a lexicon of key terms used in this primer. Categorisation, interchangeable with diagnosis, denotes the ability to determine if a new instance is a member of a category (Sloutsky 2003). It allows the cognitive system to treat similar entities equivalently, hence we can behave differently to different things and conserve cognitive effort by collapsing diversity (Bruner, Goodnow, and Austin 1956; Rosch 1978; Harnad 2005). The terms concept and category have also been used interchangeably, but here we are adopting the convention of Murphy and Medin (1985), which uses the term category to denote the set of entities in the world, and concept to denote the mental representation that supports the grouping. There is converging evidence from cognitive neuroscience and neuropsychology (e.g. category specific deficits) that humans process stimuli according to category criteria (Ashby and Ell 2001; Keri 2003; Shohamy et al. 2008; Mahon and Caramazza 2009). Categories may be organised into broad classes, which can be binary (safe/unsafe; edible/not edible), or more complex forms to reflect higher levels of discrimination (Hackett 2017). Many work environments of interest to ergonomics, such as police and security operations, healthcare and public emergency response, require specialist discrimination of events in order to tailor response. Additionally, such work systems are intolerant of delayed or inaccurate categorisation.

There is natural link between categorisation theory and the influential recognition-primed decision making (RPD) model (Klein, Calderwood, and Clinton-Cirocco 1986; Klein 1993). This model proposes mechanisms by which experienced professionals use prototypical patterns in the environment to make rapid judgements and implement response. Decision makers recall previous cases in order to assess the current situation and plan potential actions. The model has also been applied to the aviation context (O'Hare and Wiggins 2004; Wiggins et al. 2014). There has been little empirical development of the recognition strand of this model, especially in relation to professionals' abilities to acquire and use event knowledge in domains that produce both prototypical and indistinct situations. Applying the theory and associated methods can elucidate and define specific aspects of the RPD model through specification of the concepts that help and hinder recognition. Important event variability can be identified and this will clarify why recognition and decision making sometimes fail to follow the architecture of the model.

2.1 Approaches to categorisation

There are two broad approaches to categorisation, the classical view and the probabilistic view (Medin 1989). The classical view asserts that there are defining features (cues) of category members, and these features are 'necessary and sufficient' to confer category status (Smith & Medin 1981). However, there are few concepts, if any, for which there is widespread expert agreement about the category-defining features (Medin 1989). The classical view would also predict that all category members are equal as they all possess the necessary properties. This position has been weakened by the apparent graded internal structure of categories – some members are better than others (Barsalou 1985; Dry and Storms 2010; Rosch 1978). In the classic example of graded structure, robins are reliably considered better examples of the category 'bird' than penguins (Nosofsky 1988).

The probabilistic approach to categorisation rejects the idea of definitive features, in favour of gradations of membership, with some category members being central and 'good', whilst others are borderline, less well defined and close to the category boundary (Rosch 1975; Mervis and Rosch 1981). The environment is ordered around probabilistic relationships, rather than certain relationships, so category judgements are really about what is likely, and membership is by degree (Bruner, Goodnow, and Austin 1956; Harnad 1990; Mervis and Rosch 1981). This explains variability in the human capacity to recognise objects and events – as the degree of membership varies, so does the ability of the human to discriminate. The roots of this approach are often attributed to Wittgenstein (1953) and his proposition that fuzzy boundaries and shared properties better define the world, not clear boundaries and essential properties. The perennial example of 'games' illustrates the point. Games share a family resemblance, but there is no single collection of necessary properties, merely sharing some properties, drawn from a cluster, is sufficient for an activity to be called a game (Lakoff 1987). Fuzzy category boundaries are typically used in decision support systems, such as medical diagnosis applications, as abrupt category distinctions fail to capture the subtleties of borderline cases (see Miranda and Felipe 2015 for a contemporary example).

2.2 Prototypes and exemplars

Eleanor Rosch and colleagues mounted the first challenge to the classical approach, proposing that categories are organised around the clearest and best cases, known as prototypes (Rosch and Mervis 1975; Rosch 1978, 1973, 1975). These act as cognitive reference points, and new instances are judged more or less similar to the generalised, prototypical case (Rosch 1975, 1978). Clusters of the best cues, known as high cue validity, indicate the clearest, prototypical cases of membership.

The exemplar view proposes that judgements are based on particular instances that have been encountered (Smith & Medin 1981). A new case is judged against experience of previous cases, and this accommodates metrics that are known to influence category judgements, such as familiarity, or how often a person has experienced an entity (Barsalou 1985; Nosofsky 1988). Exemplars are particularly important for more complex concepts for which there may be no prototype (i.e. no ‘best’ case) such as categories with rare members or categories that exhibit high diversity (e.g. the concept 'supercomputer'; Osherson and Smith 1997).

The view that exemplars are an important form of experiential knowledge is shared with the analogical, or case-based reasoning paradigm, where important similarities are detected between a current situation (the ‘target’) and a problem solved in the past (the ‘source’) (Schunn and Dunbar 1996; Buchanan, Davis, and Feigenbaum 2006). Case-based reasoning (CBR) places emphasis on the structure of problems, rather than recognition and discrimination, thus it often involves more elaborate knowledge of elements and relations between elements (De Mantaras et al. 2005; Krawczyk 2012; Kolodner 1992). CBR also proposes two levels of similarity matching, which is of interest to the current work (De Mantaras et al. 2005). Surface similarity refers to superficial resemblance, whilst structural similarity refers to deeper relations between elements (De Mantaras et al. 2005). Detection of structural similarities between entities is overcome by means of more elaborate abstract indexing, which draws on more sophisticated domain knowledge. In categorisation theory this is accomplished through taxonomic organisation and centrality (Mervis and Rosch 1981), where concepts are formally grouped, with central cases offering something akin to surface similarity. Discriminating deeper structural similarities, just like CBR, requires enhanced domain knowledge, often in the form of specialist taxonomies. Surface similarity tends to have better transfer in the problem-solving domain, and this is mirrored in categorisation by the superior verification of prototypes (Keane 1987; Blanchette and Dunbar 2000; Rosch, Simpson, and Miller 1976).

Allied to this approach, instance-based learning theory (IBLT) has been used to model how experience, stored as ‘instances’, drives learning and decision making (Gonzalez, Lerch, and Lebiere 2003). In this paradigm previously successful decisions are used to inform current decisions. Within IBLT instances have three elements; the attributes, or cues of the situation, the corresponding decision and the utility of the
34 outcome (Gonzalez 2013). The recognition phase matches past experiences with the

current environment, and as experience grows attention becomes more selective. This approach draws on cognitive knowledge of ‘typical’ and ‘atypical’ instances, with typical situations being deemed closely similar to previous encounters (Gonzalez 2013).

The complementary nature of analogical reasoning, CBR, IBLT and categorisation is illustrated by their shared assumption that specific knowledge structures arise from previous, concrete experience of a phenomenon, be they analogies, cases, instances or exemplars, and retrieval of this knowledge allows the actor to recognise entities. Indeed, the literatures share a substantial amount of terminology (e.g. Aamodt and Plaza 1994; De Mantaras et al. 2005). Each of these perspectives has different objectives, but in the context of this article we assume the exemplar to be the unit of direct experience, and we draw on notions of recognition and similarity discussed above.

Research in categorisation theory has advanced using both prototype and exemplar views, and current conceptions indicate they have complementary, supportive functions; for example, exemplars update malleable prototype knowledge (Murphy & Hoffman, 2012; Murphy, 2002; Smith, 2014; Smith & Minda, 1999), although theoretical controversies still exist (e.g. Murphy 2016). Both processes are unified by similarity, or the degree to which a new encounter matches, or overlaps with known concept distinctions, be they prototypes or exemplars (Medin 1989; Murphy and Hoffman 2012; Sloutsky 2003; Smith, Patalano, and Jonides 1998; Larkey and Markman 2005; Goldstone 1994). It is possible that pilots have weak, decayed or no prototypical event knowledge for certain event types, or they may never have seen particular events before (even in training), therefore have no exemplar.

2.3 The key gradients of typicality and familiarity

As noted above, categories appear to exhibit internal gradations, meaning that some members are reliably judged better than others. There are central cases and there are borderline cases. These variations are known as gradients. This article is focussed on the gradients of typicality and familiarity. Typicality is the degree to which a category member serves as a good, central instance (Rosch 1975). Rated typicality provides the strongest empirical evidence that categories exhibit gradients, and this has been demonstrated across a wide range of categories, including animals, metals, furniture, sports and even formal categories like numbers (Dry and Storms 2010). The typicality effect is associated with variations in category verification tasks (member yes/no), category learning and category member naming (Rosch, Simpson, and Miller 1976; Storms, De Boeck, and Ruts 2000; Sandberg, Sebastian, and Kiran 2012). Typical cases confer cognitive advantages, such as speed and accuracy. The best cases, proximal to the prototype, are powerful reference points for rapid and effective response.

Familiarity, or the subjective estimate of frequency of occurrence, has also been shown to be a potent driver of recognition and retrieval (Nosofsky 1988; Barsalou 1985). Frequency is analogous to repetition, so is influential in learning (Rosch, Simpson, and Miller 1976). This may be a particularly important gradient for pilots when rare and unusual cases present, possibly many years after brief exposure to a crude exemplar. Adverse familiarity gradients could damage recognition and response.

Term	***Definition***
Borderline case	Cases that are not central, or readily categorised, due to imperfect information or cue combinations. Cases proximal to category boundaries. (Genero and Cantor 1987).
Categorisation	The ability to determine if a new instance is a member of a known category. Interchangeable with *diagnosis*. (Sloutsky 2003). *Delayed categorisation* involves a temporal gap that interferes with response and safety. *Mis-categorisation* is assigning an entity to an incorrect category. *Non-categorisation* is failure to assign an entity to a category.
Category	A group of distinct items or entities that the cognitive system treats as equivalent for a particular purpose. See also *concept*. (Markman and Ross 2003).
Category structure	Real-world categories exhibit internal gradations (e.g. members vary in their *typicality* and *familiarity*). Some category members are more central, and therefore better at conveying the category distinction, than others. (Barsalou 1985).
Concept	The mental representation that supports the grouping of entities into categories. (Murphy & Medin, 1985).
Cue	A characteristic of the environment that can be used to resolve distinctions. The best cues provide clear cases of category membership. (Rosch 1978).
Exemplar	Particular stored example of a category member or particular instances that have been encountered. (Minda and Smith 2002; Smith and Medin 1981).
Familiarity	The subjective estimate of how often a person has experienced an entity. (Nosofsky 1988).
Prototype	Generalised case of the clearest and best examples of category membership. Used as a cognitive reference point to make category judgements. (Rosch 1975).
Similarity	The degree to which a new encounter overlaps with known category distinctions, such as exemplars or prototypes. (Goldstone 1994).
Typicality	With reference to a category member, the extent to which it serves as a good, central member. (Rosch 1975).

Table 1. Lexicon of categorisation theory across the literature.

3. Case Studies: Applying categorisation theory to flight safety events

In this section we present a new framework developed from categorisation theory. We describe three types of event and their characteristics (Table 2). We then demonstrate through case studies how the theory can assist in understanding events and reduce future risk of adverse outcomes.

This framework is compatible with recent developments in startle and surprise **(for example, Landman et al. 2017) and also supplements Clewley and Stupple's (2015) conception of procedural vulnerability**. Transport systems are heavily reliant on rules and procedures (Pélegrin 2013) and it is notable that events similar to AF447 show very poor compliance with response protocols (BEA 2012). Clewley and Stupple (2015) propose that rules and procedures become vulnerable under demanding conditions, and this can progress to two undesirable situations. Procedures become fragile, and at significant risk to being partially implemented or not achieving their aim, or they may fail completely, meaning they are not implemented at all. This offers a more nuanced progression which reflects the fact that procedures, response tools and protocols are used and misused in many ways. Case studies two and three elucidate. This connection between event structure and the effectiveness of response tools and protocols has important implications for training and pilot education.

We also wish this framework to reflect current thinking in event complexity. Walker and colleagues (2010) reviewed and translated the ideas of complexity to the field of ergonomics. The attribute view of complexity (see Walker et al. 2010) defines the ergonomic problem space as containing multiplicity (multiple interacting factors), dynamism (system state changes over time) and uncertainty (difficulty and vagueness in determining the final system state), with high levels of each producing the most complex situations. We view these attributes as important contextual factors in the structure of events. The case studies note how this informs our argument.

Type of Event	Status	Characteristics	Predicted Response	Example Flight Safety Events	Case Study Notes
Prototype or Exemplar	Friendly	Clearest and best cases. Typical and familiar. Clusters of the best cues. Good correspondence with previous encounters. Proximal to reference concepts. Strong information position. Exhibits stable features.	Adequate recognition. Appropriate response likely. Response tools used effectively to enhance safety. Startle unlikely. Procedural compliance robust. Pilots are well adapted, able to produce rule and skill-based response.	Events regularly seen in actual flight operations. Events that are well replicated in flight simulators and arise in a similar context in the real-world. Events where assistive technology conveys a clear, unambiguous concept. Events reinforced through repetition. Events that clearly correspond with checklists and protocols.	Case study 1. Engine failures are well replicated and often practised. Assistive technology conveys a clear concept.
Non-Typical or Unfamiliar	Alarm	Cues may be masked, contradictory, conflicting, intermittent, unusually dynamic, configured in unusual combinations, or unfamiliar. Limited correspondence to reference concepts.	Recognition may be delayed or inaccurate. Delayed or inappropriate response possible. Response tools may not be used effectively. Startle possible. Procedures vulnerable.	Events that have not been encountered before, but are similar to known, strong concepts. Events that have been encountered in the flight simulator but arise in a significantly different context or format. Events that exhibit high variability in terms of cue combinations and temporal characteristics. Events that are tenuously similar to previous brief encounters. Borderline events that are conceptually similar to others, causing confusion. Familiar events but mismanaged.	Case study 2. The familiar go-around manoeuvre is mismanaged. Unfamiliar flap malfunction causes confusion. Response tools not used effectively.
Weak or No Concept	Hostile	Cues may be masked, contradictory, conflicting, intermittent, unusually dynamic, configured in unusual combinations, or unfamiliar. Minimal or no correspondence with known category distinctions. Distal from reference concepts. Concepts resistant to prototypes. Weak information position.	Recognition may be delayed, inaccurate or absent. Delayed or inappropriate response likely. Response tools may not be used effectively. Startle likely. Procedures fragile. Pilots are poorly adapted.	No previous experience of event or anything similar. Severely decayed memory trace of an encounter. Events that exhibit high diversity and complexity in the real world, but are trained through brief, contrived exemplars. Complex events where cue processing in training/testing is inadequate to form robust foundation concept. Events where assistive technology fails to stabilise in a coherent state or cues are inadequately diagnostic. Events beyond published procedures. Unfamiliar events mismanaged. Multiple, dynamic, concurrent events.	Case study 3. Complex presentation involving multiple, dynamic events. Unfamiliar events mismanaged. Minimal overlap with reference concepts. Procedures and response tools failed.

Table 2. Three conceptions of events and predicted characteristics.

3.1 Case study 1, Airbus A330 engine shutdown: Friendly.

This case study describes a prototypical event structure with good correspondence to strong concepts. Categorisation theory would predict adequate recognition and response. It serves as a base line.

In January 2013 an Airbus A330-300 took off from Orlando Airport, Florida, and suffered a birdstrike at 530 feet (AAIB 2013). Shortly afterwards the Electronic Centralized Aircraft Monitor (ECAM, an assistive technology that suggests event titles and remedial actions) displayed ‘ENG 1 FAIL’ and ‘ENG 1 LO PR’, indicating that engine 1 had failed and the oil pressure was low. The crew subsequently shut down the damaged engine and returned to Orlando for a safe landing. The total time from first warning to engine shutdown was 1 minute and 46 seconds.

This event description closely matches the BEA’s expectations of pilot behaviour, as described above. The main cues available to the crew were [BIRDSTRIKE] + [ENG 1 FAIL ECAM] + [ENG 1 LO PR ECAM] + [ENG 1 PARAMETER VARIATIONS], and collectively they form a strong, non-contradictory cluster, known as high cue validity. The event content closely matches, or is equivalent to, a known concept, the often practised ‘engine failure on take-off’. The event is a good, central case, proximal to the prototype, and the report indicates the crew had no difficulty recognising or responding to it. These events are rarely encountered in flight operations, but often encountered in this context in training. Furthermore, the ECAM technology conveyed a highly descriptive, stable concept (simply ‘ENG 1 FAIL’), but this is not always the case (see case study 3, below). As Cantor and colleagues (1982) pointed out when they originally extended prototype theory to situations, prototypes allow an actor to plan and regulate behaviour. Well-structured events like this elicit desirable flight crew behaviour, and the crew used procedures and response tools effectively. This is the cognitive dividend of the typicality effect (Rosch, Simpson, and Miller 1976). This event was conceptually friendly due to its proximity to a strong cognitive reference point. We interact well with clear concepts, and they optimise adaptation, prediction and behaviour (Anderson 1991).

3.2 Case study 2, Boeing 757 go-around: Alarm

Systems do not always fail close to cognitive reference points. Pilots are not always able to use clear concepts to interpret and respond to flight safety events. A mismanaged go-around in a Boeing 757-200 (AAIB 2014) demonstrates the alarm case in Table 2. A go-around is a manoeuvre flown to discontinue an approach to land, and involves applying high power whilst changing the flight path to an upward trajectory. The crew were conducting an instrument approach, with landing gear and flaps extended, approximately 1250 feet above the ground, when they were instructed to go-around. The Captain disengaged the automatic thrust control, applied maximum power, but left the autopilot engaged in the programmed landing trajectory. Shortly after, with the speed increasing rapidly, the Captain disconnected the autopilot and manually flew the aircraft into a climb. The aircraft subsequently deviated from its cleared altitude and experienced two flap speed exceedances, which could cause structural damage to the aircraft. The crew mismanaged the flap malfunction and its response protocol, but later landed safely.

The report suggests go-around events most often occur during simulator exercises with one of the two engines on the Boeing 757 inoperative; they are usually initiated at or near the decision altitude of approximately 200 feet; rule-based weather criteria usually prime the decision; and they typically take place under examination, with high levels of anticipation. The Captain noted that ‘you know they [go-arounds] are coming’ (AAIB 2014 p.58) during such training encounters.

This instance deviated significantly from the prototypical event markers. It was initiated at higher altitude, with full thrust, was unexpected and included a procedural error. If flight crew knowledge of these events is organised around a well-executed prototype, the clear case in training, it is predictable that their knowledge and performance would deteriorate as they diverged from the typical instance. Go-arounds are demanding manoeuvres. The European Aviation Safety Agency report that 1 in 10 go-arounds result in hazardous outcomes similar to this case (EASA 2014). This case also suggests that even relatively minor procedural errors can lead to significant deviation from typicality, resulting in a ‘familiar event but mismanaged’ scenario. Trained prototypes may be difficult to override and may leave pilots poorly positioned to recover from non-typical instances. Events may deteriorate quickly, leaving the pilot in a flight regime for which they have little coherent conceptual knowledge.

The crew also exceeded the speed limit for the wing slat and flap devices during the go-around, and this caused a variety of unusual system behaviour. They were not sure if they had exceeded the flap speed limits, despite experiencing an 18 knot and 47 knot exceedance at two different flap settings. Flap speed exceedances are not practiced as they are usually a consequence of inadvertent mismanagement of energy and flightpath. As a result the crew had no exemplar events to compare with their predicament or help with recognition. They then experienced two separate caution messages, ‘LE (leading edge) SLAT DISAGREE’ and ‘TE (trailing edge) FLAP DISAGREE’, partly the result of the speed exceedance and partly due to the incorrect execution of the checklist. These failure messages are conceptually similar, so prone to confusion, a situation exacerbated by their subsequent disappearance brought about by switching to alternate flap control. This illustrates how connecting failure messages with correct checklist selection can be difficult in non-typical scenarios, especially where system behaviour is less transparent and informative than previous flight crew encounters. The difficulty the crew had in managing these events was attributed to their unfamiliarity with this type of malfunction and the associated checklist. The infrequency of exemplar slat/flap events leads to weaker conceptual knowledge; a classic gradient that damaged the crews understanding of the situation and their behaviour. Unfortunately a great number of flight safety events show this characteristic.

3.3 Case study 3, Airbus A330 unreliable airspeed data: Hostile

This case study examines the previously discussed AF447 accident. To re-cap, the aircraft experienced ice accumulation on the airspeed sensors, and this resulted in unreliable airspeed readings and eventually an aerodynamic stall (BEA 2012). Many other system indications were triggered during the event, including six separate episodes of the flight director vanishing, intermittent stall warnings and multiple system malfunction messages. The event was particularly dynamic, with approximately thirty system transitions in the first 99 seconds (BEA 2012). Neither the unreliable airspeed nor the stall were adequately recognised and the remedial actions were not carried out.

3.3.1 The role of exemplar events

The BEA report specifically refers to each crew member exposure to unreliable airspeed scenarios and stall scenarios. All the three crew members on AF447 had encountered such scenarios during routine simulator training. These encounters occur during the initial aircraft type training (type rating), or during routine re-validation of the pilot's licence and type-rating. These encounters are intended to serve as sample events that will lead to greater reliability of pilot performance during a real encounter. They are exemplars. Categorisation theory proposes that the real encounter must have sufficient overlap with the exemplar if it is to be recognised.

First, we consider exemplar unreliable airspeed events. The three crew on board AF447 had all experienced an unreliable airspeed event in the simulator training and testing programme for 2008-09. Additionally, the Captain had an additional exposure to such an event in 1997. This puts a lot of emphasis on minimal exposure. There is no telling how long these simulated events lasted. We cannot we be sure how well the pilots were able to process the patterns of cues during the session. We cannot be sure the cues exhibited in the simulator session corresponded with the cues in the real encounter. Given that the crew, with collectively over 19,000 hours, had difficulty decoding the event, we suggest that an exemplar effect damaged both recognition and response. Categorisation theory points to a weak or decayed exemplar, combined with unruly event content that had minimal overlap with reference concepts.

Second, we consider exemplar stall events in the flight simulator. The Co-Pilot occupying the right-hand seat had experienced a simulated encounter five years previously, the Captain's most recent stall exercise was eight years previously and Co-Pilot in the left hand seat had not seen an exemplar for eleven years. The BEA note that simulator fidelity, flight envelope constraints and lack of startle effect pose problems for stall recreation in a simulator. In addition, they underline that in the cruise flight conditions at Mach 0.8, just a 1.5-degree change in wing angle of attack is sufficient to induce a stall warning, and this contrasts sharply with the classic low-speed stall seen during basic training. Accepting that typicality and familiarity confer a cognitive advantage, then the reverse is true. There was no cognitive dividend for the crew, they were in a weak information position remote from reference concepts and response was disrupted. We acknowledge that situations will occur beyond the reasonable prediction or resources of training departments (Dahlstrom et al. 2009), but we would also argue that a greater understanding of event structure could have safety benefits in cases such as this. Categorisation theory offers scope to train pilots to be sensitive to event variability and diversity, thereby reducing reliance on crude, possibly brief exemplars. Pilot education could be expanded to include the friendly, alarm and hostile levels of event structure, and this could make procedures and response tools less vulnerable – see section 4, below.

3.3.2 The role of assistive technology

Air data malfunctions, such as this unreliable airspeed encounter, do not always arise with descriptive and stable indications from the assistive technology. This is in stark contrast to case study 1, where a clear concept was conveyed to the crew via the ECAM display ('ENG 1 FAIL'). Air data events demand more elaborate diagnosis and can often involve a wide variety of system indications and cue diversity, requiring story building and collecting more diagnostic features. Many of the diagnostic cues available to the AF447 crew were intermittent. System transitions were almost continuous, so whilst it is seductive to think there was a single, recognisable condition (e.g. a stall), the multiplicity of conflicting cues make this remote. In terms of story-building, this event structure had no narrative arc. This is a two-fold cognitive burden – the assistive technology conveys no clear concept and the diagnostic cues do not come to rest (or near rest). This leads to variability in diagnosis. The BEA report that two other crews who experienced a similar unreliable airspeed encounter concluded it was a 'loss of angle of attack data'. In categorisation theory this is known as an adjacent concept and this can lead to 'mis-categorisation' (see the 'categorisation' entry in Table 1). It is conceptually similar and close to the correct distinction, but not accurate. These are known as boundary or borderline cases (Genero and Cantor 1987), and such exotic events appear to be demanding to recognise, even for experienced professionals. These are more hostile and complex forms of event structure.

3.4 Summary

There are several notable insights from categorisation theory. Event prototypes offer typicality and familiarity benefits, especially if they arise in similar contexts to simulated or previous encounters. These are relatively friendly situations to which pilots are well adapted. However, prototypical events may be conceptually narrow and very well primed, leading to deterioration in performance when significant deviations from typicality occur. These situations prompt alarm, and response protocols may become vulnerable. Complex events may exhibit considerable cue diversity and may have no clear correspondence with reference concepts. Crude or brief training events may be inadequate to promote recognition and satisfactory pilot response, especially during unfamiliar and exotic situations. This is a hostile cognitive environment to which pilots may be poorly adapted. The recognition of events appears deeply connected to typicality and familiarity gradients. A weak information position harms the connection between event and response, perhaps leading to delayed or inappropriate pilot behaviour. A key contribution of categorisation theory is its ability to describe the structure of the environment, as well as human cognition. In Table 2 we have defined three levels of this transaction and illustrated the potential of the framework through the case studies.

4. Application to flight safety

A number of exploitation routes are available for the application of categorisation theory to flight safety. Firstly, training could be a beneficiary of this approach. The commercial aviation industry recognises the need for updated training and testing of pilots, which has barely changed in over 30 years (Learmount 2017). Case studies 2 and 3, above, indicate a gap between trained situations and the variability and diversity exhibited in the real-world. This shortfall is often absorbed through failure, with significant human, social and economic costs.

We view current training/testing encounters (e.g. the simulator) as possibly being limited environments for pilots to acquire cue combinations and build conceptual knowledge. There are audit-based expectations on pilots, they are often under test conditions with other cognitive tasks and documentary evidence of their performance is being gathered (see Roth 2015). This may seriously limit their ability to process event cues, yet these encounters may be their only exposure to relatively complex events. Furthermore, several events may be presented in quick succession. Categorisation theory, its lexicon and its methods allow the event envelope to be expanded to take account of non-typical and complex event structure. We see two routes to this. Firstly, improving opportunities to process cue combinations may lead to better reference concepts, which we know to be important drivers of recognition. Better event recognition should lead to more reliable response. Secondly, additional emphasis on event variability and diversity could provide flight crew wider conceptual knowledge and strategies for dealing with unruly information, for example contradictory and intermittent data. We predict this will provide better pilot performance when faced with borderline events or events where assistive technology fails to convey a stable concept. Some event types, such as the air data malfunction of AF447, are profoundly more complex than others, and training needs to take account of this.

As an example, we will now describe how this could be done using our event framework found in Table 2. If we consider an 'Airspeed Mismatch' condition where pilot airspeed indications do not match, requiring flight crew intervention. Pilots could experience a prototypical event of this type, where cues are clear, stable and correspond well with the response tools, such as memory items and checklists. This is the routine event to which pilots are well accustomed (Casner, Geven, and Williams 2013). Using Table 2, a more complex form of this event could be generated. An instance involving intermittent cues that fails to stabilise in a particular system state, or a presentation that is more difficult to connect to the response tools, will offer training benefit. Pilots experience important variations on an event theme, and they also get exposure to non-typical and complex event characteristics. We feel this has the potential to improve pilot training in the simulator, or more broadly contribute to pilot education, by providing a theory-driven framework to devise scenarios or event characteristics that expand pilot knowledge. Indeed, using the Table 2 framework for educational tasks need not carry the cost burden of simulator training (see paragraph 5.2, below, where concept building is suggested). We acknowledge it is not possible to assemble exhaustive cue patterns for all types of scenario, but improvements to pilot knowledge that we propose should lead to greater cognitive flexibility in managing non-typical events. Overall, we feel this offers pilots opportunities that are currently beyond industry thinking.

Finally, we believe response protocols and procedures, such as those specified in abnormal and emergency checklists and aircraft handling manuals, could be enhanced to include recognition guidance that reflects the ambiguous nature of some events. Response protocols pre-suppose that a situation has been recognised adequately. If we examine the thirteen other crews, from five different airlines, who experienced an

AF447 type-event, it is notable that the BEA found no evidence any had used the remedial ‘memory actions’ or emergency checklist procedure. This suggests a systematic weakness in pilot capability to connect these events to the correct procedure. Assistive technologies such as the Quick Reference Handbook (an abnormal and emergency checklist) could include likely cue combinations and sources of the best and strongest data. This also raises design issues, particularly around presentation and display of failure messages and diagnostic cues. For example, in the case of AF447, displaying angle of attack data may have provided the crew with strong, supportive help in recognition (see The Boeing Company 2000 for a discussion on the benefits and drawbacks of providing angle of attack data on large commercial aircraft).This reorganisation of cockpit materials could provide algorithm-based support to recognition. Such information could be used by pilots when briefing abnormal situations or during the real-time event management. A recognition-focussed approach may reduce the fragility of procedures and response tools. In principle these improvements to training and response protocols could be used in other domains, such as security and police operations, medicine, nursing and social work.

5. Future directions in research

Application and expansion of this theory can provide greater precision in describing flight safety events, deliver insights into pilot behaviour and inform future best practice The key objective of this research programme is to offer solutions that provide tangible safety improvements. We have identified three themes that are amenable to empirical study and draw on gaps in current knowledge in flight safety.

5.1 Concept gradients and strength

The gradients of typicality and familiarity are measurable, so may be used as a proxy for concept strength. This will allow us to identify which event types may challenge appropriate crew response. One interesting application of this would be the apparent difficulty flight crew have recognising unstable approaches to land (e.g. NTSB 2014a; NTSB 2014b; AAIB 2014b; DSB 2010). For example, pilots may have weak conceptual knowledge of unstable approaches, compared to prototypical successful approaches.

These gradients may also indicate which events are resistant to prototypes, and so rely heavily on exemplar exposure. The source of event concepts may also be important, as the repetition of real flight operations may form concentrations of knowledge, whereas flight simulator encounters may involve weaker concepts or over-learning a response, at the expense of recognition. This is compatible with recent developments in surprise and startle, where dominant concepts, fed by the typical and familiar, bias response (Landman et al. 2017).

5.2 Acquisition, maintenance and decay of concepts

Pilots have the opportunity to acquire and maintain event knowledge through a variety of channels. Over time this knowledge can decay. Recent empirical data suggests that abnormal events presented outside the usual training and testing paradigm pose problems for pilot response (Casner, Geven, and Williams 2013), and this certainly mirrors the case studies presented above. Pilots may acquire the ability to respond to an event in the context of a training encounter, but do not necessarily acquire cue patterns that will lead to real-world recognition. It may be possible to devise vignettes for 'concept building', outside of the training/testing paradigm, in order to boost recognition power. Allied to this, there does not appear to be a simple connection between some events and their response protocols, particularly 'memory actions'. Following AF447 the BEA interviewed a sample of flight crew and found a prominent reluctance or inability to use these procedures. We need a better understanding of how procedures, response tools and protocols are acquired and maintained, and crucially, how they remain connected to event concepts. This procedural vulnerability and fragility needs to be better understood.

5.3 Signatures of complex events

The AF447 accident demonstrates that multiple, overlapping events may poorly correspond to exemplar events that flight crew see in simulator exercises. The real-world also presents cases that are beyond published procedures (NASA 2005). This suggests some events exhibit a complexity for which pilots are ill-prepared. If we expect crew to recognise these events then we need a greater understanding of these signatures. For example, pilots do not routinely or intentionally practice recovering from mismanaged events, yet paradoxically this may extend conceptual event knowledge. Even minor procedural errors during a go-around, for instance, can escalate into a more complex situation, as case study 2 indicated. We propose a research programme to understand the interaction between concept gradients and the three key attributes of complexity, dynamism, uncertainty and multiplicity (described in Walker et al. 2010). This may shed light on the most challenging forms of event structure that pose the greatest risk of delayed or inappropriate pilot behaviour, and procedural fragility. Capturing complexity, or a useful proxy, means it can be embedded in pilot training and education.

6. Conclusions

Categorisation theory proposes that recognising events in the real-world requires a similarity overlap between either a prototypical, ideal case or a stored exemplar derived from experience. Pilots are not always able to recognise events and this often leads to undesirable pilot behaviour. Using categorisation theory we have introduced a stable lexicon, and established a framework to understand how pilots recognise and respond to flight safety events. We have significantly extended the idea of event prototypes and analysed the cognitive benefits and costs of such knowledge structures. The framework also accounts for important event variety, describing both the typical and familiar, and troubling divergence away from reference concepts. We have linked this event diversity to predictions of pilot response, specifying key environments that pose risks to procedures and response tools. We have suggested a research programme to refine the application of the theory. Such research can contribute to broader themes in ergonomics science, such as capturing the enigma of complexity. We believe categorisation theory can account for important variability in event recognition and pilot behaviour, and ultimately contribute tangible improvements to safety.

References

AAIB (Air Accidents Investiagtion Branch). 2013. “AAIB Bulletin: 9/2013 G-VKSS EW/C2013/01/03.” London.

———. 2014a. “AAIB Bulletin: 10/2014 G-TCBC EW/G2013/08/19.” London.

———. 2014b. “AAIB Bulletin: 10/2014 VP-CKY.” London.

Aamodt, A., and E. Plaza. 1994. “Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches.” *AI Communications*. https://doi.org/10.3233/AIC-1994-7104.

Anderson, J. R. 1991. “The Adaptive Nature of Human Categorization.” *Psychological Review* 98 (3): 409–29. https://doi.org/10.1037/0033-295X.98.3.409.

Ashby, F. G., and S. W. Ell. 2001. “The Neurobiology of Human Category Learning.” *Trends in Cognitive Sciences* 5 (5): 204–10. https://doi.org/10.1016/S1364-6613(00)01624-7.

Barsalou, L. W. 1985. “Ideals, Central Tendency, and Frequency of Instantiation as Determinants of Graded Structure in Categories.” *Journal of Experimental Psychology. Learning, Memory, and Cognition* 11 (4): 629–54. https://doi.org/10.1037/0278-7393.11.1-4.629.

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*. Paris: BEA.

Blanchette, I., and K. Dunbar. 2000. “How Analogies Are Generated: The Roles of Structural and Superficial Similarity.” *Memory and Cognition*. https://doi.org/10.3758/BF03211580.

Bruner, J., J. Goodnow, and G. Austin. 1956. *A Study of Thinking*. New York: John Wiley.

Buchanan, B. G., R. Davis, and E. A. Feigenbaum. 2006. “Expert Systems: A Perspective from Computer Science.” In *The Cambridge Handbook of Expertise and Expert Performance*, edited by K. A. Ericsson, N. Charness, P. J. Feltovich, and R. R. Hoffman, 87–103. Cambridge: Cambridge University Press.

Cantor, N., W. Mischel, and J. C. Schwartz. 1982. “A Prototype Analysis of Psychological Situations.” *Cognitive Psychology* 14: 45–77. https://doi.org/10.1016/0010-0285(82)90004-4.

Casner, S. M., R. W. Geven, and K. T. Williams. 2013. “The Effectiveness of Airline

Pilot Training for Abnormal Events." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 55 (3): 477–85. https://doi.org/10.1177/0018720812466893.

Clewley, R., and E. J. N. Stupple. 2015. "The Vulnerability of Rules in Complex Work Environments: Dynamism and Uncertainty Pose Problems for Cognition." *Ergonomics* 58 (6): 935–41. https://doi.org/10.1080/00140139.2014.997804.

Dahlstrom, N., S. Dekker, R. van Winsen, and J. Nyce. 2009. "Fidelity and Validity of Simulator Training." *Theoretical Issues in Ergonomics Science* 10: 305–14. https://doi.org/10.1080/14639220802368864.

Dry, M. J., and G. Storms. 2010. "Features of Graded Category Structure: Generalizing the Family Resemblance and Polymorphous Concept Models." *Acta Psychologica* 133 (3): 244–55. https://doi.org/10.1016/j.actpsy.2009.12.005.

DSB (Dutch Safety Board). 2010. "Crashed during Approach, Boeing 737-800, near Amsterdam Schiphol Airport, 25 February 2009: Project Number M2009LV0225_01." The Hague.

EASA (European Aviation Safety Agency). 2014. "EASA SIB No: 2014-09 Aeroplane Go-Around Training." Cologne.

Genero, N., and N. Cantor. 1987. "Exemplar Prototypes and Clinical Diagnosis: Toward a Cognitive Economy." *Journal of Social and Clinical Psychology* 5 (1): 59–78. https://doi.org/10.1521/jscp.1987.5.1.59.

Goldstone, R. L. 1994. "The Role of Similarity in Categorization: Providing a Groundwork." *Cognition* 52 (2): 125–57. https://doi.org/10.1016/0010-0277(94)90065-5.

Gonzalez, C., J. F. Lerch, and C. Lebiere. 2003. "Instance-Based Learning in Dynamic Decision Making." *Cognitive Science*. https://doi.org/10.1016/S0364-0213(03)00031-4.

Gonzalez, C. 2013. "The Boundaries of Instance-Based Learning Theory for Explaining Decisions from Experience." *Progress in Brain Research* 202: 73–98. https://doi.org/10.1016/B978-0-444-62604-2.00005-8.

Hackett, P. M. W. 2017. "Editorial: Conceptual Categories and the Structure of Reality: Theoretical and Empirical Approaches." *Frontiers in Psychology* 8 (601): 1–2. https://doi.org/10.3389/fpsyg.2017.00601.

Harnad, S. 2005. "To Cognize Is to Categorize: Cognition Is Categorization." In *Handbook of Categorization in Cognitive Science*, edited by H. Cohen and C.

Lefebvre, 19–43. London: Elsevier.

Harnad, S. 1990. “Psychophyscial and Cognitive Aspects of Categorical Perception: A Critical Overview.” In *Categorical Perception: The Groundwork of Cognition*, edited by S. Harnad, 1–25. Cambridge: Cambridge University Press.

Keane, M. 1987. “On Retrieving Analogues When Solving Problems.” *The Quarterly Journal of Experimental Psychology Section A*. https://doi.org/10.1080/02724988743000015.

Keri, S. 2003. “The Cognitive Neuroscience of Category Learning.” *Brain Research Reviews* 43 (1): 85–109. https://doi.org/10.1016/j.neubiorev.2007.11.002.

Klein, G. 1993. “A Recognition-Primed Decision (RPD) Model of Rapid Decision Making.” *Decision Making in Action: Models and Methods*, 138–47. https://doi.org/10.1002/bdm.3960080307.

Klein, G., R. Calderwood, and A. Clinton-Cirocco. 1986. “Rapid Decision Making on the Fire Ground.” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 30 (6): 576–80. https://doi.org/10.1177/154193128603000616.

Kolodner, J. L. 1992. “An Introduction to Case-Based Reasoning.” *Artificial Intelligence Review*. https://doi.org/10.1007/BF00155578.

Krawczyk, D C. 2012. “The Cognition and Neuroscience of Relational Reasoning.” *Brain Research*. https://doi.org/10.1016/j.brainres.2010.11.080.

Lakoff, G. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press. https://doi.org/10.2307/415440.

Landman, A., E. L. Groen, M. M. van Paassen, A. W. Bronkhorst, and M. Mulder. 2017. “Dealing With Unexpected Events on the Flight Deck: A Conceptual Model of Startle and Surprise.” *Human Factors: The Journal of the Human Factors and Ergonomics Society* 59 (8): 1161–72. https://doi.org/10.1177/0018720817723428.

Larkey, L. B., and A. B. Markman. 2005. “Processes of Similarity Judgment.” *Cognitive Science* 29 (6): 1061–76. https://doi.org/10.1207/s15516709cog0000_30.

Learmount, D. 2017. “Training for Tomorrow Today.” *Aerospace* 44 (9): 38–40.

Mahon, B. Z., and A. Caramazza. 2009. “Concepts and Categories: A Cognitive Neuropsychological Perspective.” *Annual Review of Psychology* 60 (1): 27–51. https://doi.org/10.1146/annurev.psych.60.110707.163532.

Mantaras, R. L. De, D. Mcsherry, D. Bridge, D. Leake, B. Smyth, S. Craw, B. Faltings, et al. 2005. “Retrieval, Reuse, Revision and Retention in Case-Based Reasoning.”

Knowledge Engineering Review. https://doi.org/10.1017/S0269888906000646.

Markman, A. B., and B. H. Ross. 2003. “Category Use and Category Learning.” *Psychological Bulletin* 129 (4): 592–613. https://doi.org/10.1037/0033-2909.129.4.592.

Medin, D. L. 1989. “Concepts and Conceptual Structure.” *The American Psychologist* 44 (12): 1469–81. https://doi.org/10.1037/0003-066X.44.12.1469.

Mervis, C. B., and E. Rosch. 1981. “Categorization of Natural Objects.” In *Annual Review of Psychology*, 32:89–115. https://doi.org/10.1146/annurev.ps.32.020181.000513.

Minda, J. P., and J. D. Smith. 2002. “Comparing Prototype-Based and Exemplar-Based Accounts of Category Learning and Attentional Allocation.” *Journal of Experimental Psychology. Learning, Memory, and Cognition* 28 (2): 275–92. https://doi.org/10.1037/0278-7393.28.2.275.

Miranda, G. H. B., and J. C. Felipe. 2015. “Computer-Aided Diagnosis System Based on Fuzzy Logic for Breast Cancer Categorization.” *Computers in Biology and Medicine*, no. 64: 334–46. https://doi.org/10.1016/j.compbiomed.2014.10.006.

Murphy, G., and A. Hoffman. 2012. “Concepts.” In *The Cambridge Handbook of Cognitive Science*, edited by K. Frankish and W. M. Ramsey. Cambridge: Cambridge University Press.

Murphy, G. L. 2016. “Is There an Exemplar Theory of Concepts?” *Psychonomic Bulletin & Review* 23 (4): 1035–42. https://doi.org/10.3758/s13423-015-0834-3.

Murphy, G. L., and D. L. Medin. 1985. “The Role of Theories in Conceptual Coherence.” *Psychological Review* 92 (3): 289–316.

Murphy, G L. 2002. *The Big Book of Concepts*. Cambridge, MA: MIT Press. https://doi.org/10.1017/S030500090300597X.

NASA. 2005. “The Challenge of Aviation Emergency and Abnormal Situations - NASA/TM—2005–213462.” Ames Research Center, CA.

Nosofsky, R. M. 1988. “Similarity, Frequency, and Category Representations.” *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14 (1): 54–65. https://doi.org/10.1037/0278-7393.14.1.54.

NTSB (National Transportation Safety Board). 2014a. *Crash During a Nighttime Nonprecision Instrument Approach to Landing UPS Flight 1354 Airbus A300-600, N155UP Birmingham, Alabama August 14, 2013*. Washington, DC: NTSB.

———. 2014b. *Descent Below Visual Glidepath and Impact With Seawall Asiana*

Airlines Flight 214 Boeing 777-200ER, HL7742 San Francisco, California July 6, 2013. Washington, DC: NTSB.

O'Hare, D., and M. Wiggins. 2004. "Remembrance of Cases Past: Who Remembers What, When Confronting Critical Flight Events?" *Human Factors: The Journal of the Human Factors and Ergonomics Society*. https://doi.org/10.1518/hfes.46.2.277.37333.

Osherson, D., and E. E. Smith. 1997. "On Typicality and Vagueness." *Cognition* 64 (2): 189–206. https://doi.org/10.1016/S0010-0277(97)00025-5.

Pélegrin, C. 2013. "The Never-Ending Story of Proceduralization in Aviation." In *Trapping Safety into Rules: How Desirable or Avoidable Is Proceduralization?*, edited by C. Bieder and M. Bourrier, 13–25. Farnham: Ashgate.

Pothos, E. M., and A. J. Wills. 2011. "Introduction." In *Formal Approaches in Categorization*, edited by E. M. Pothos and A. J. Wills, 1–17. Cambridge: Cambridge University Press.

Rosch, E. 1975. "Cognitive Reference Points." *Cognitive Psychology* 7 (4): 532–47. https://doi.org/10.1016/0010-0285(75)90021-3.

———. 1978. "Principles of Categorization." In *Cognition and Categorization*, edited by E. Rosch and B. B. Lloyd, 27–48. Hillsdale, NJ.: Erlbaum.

Rosch, E., and C. B. Mervis. 1975. "Family Resemblances: Studies in the Internal Structure of Categories." *Cognitive Psychology* 7 (4): 573–605. https://doi.org/10.1016/0010-0285(75)90024-9.

Rosch, E., C. Simpson, and R. S. Miller. 1976. "Structural Bases of Typicality Effects." *Journal of Experimental Psychology: Human Perception and Performance* 2 (4): 491–502. https://doi.org/10.1037/0096-1523.2.4.491.

Rosch, E. 1973. "On the Internal Structure of Perceptual and Semantic Categories." In *Cognitive Development and the Acquisition of Language*, edited by T. E. Moore, 111–44. New York: Academic Press.

Roth, W-M. 2015. "Flight Examiners' Methods of Ascertaining Pilot Proficiency." *International Journal of Aviation Psychology* 25 (3–4): 209–26. https://doi.org/10.1080/10508414.2015.1162642.

Sandberg, C., R. Sebastian, and S. Kiran. 2012. "Typicality Mediates Performance during Category Verification in Both Ad-Hoc and Well-Defined Categories." *Journal of Communication Disorders* 45 (2): 69–83. https://doi.org/10.1016/j.jcomdis.2011.12.004.

Schunn, C. D., and K. Dunbar. 1996. "Priming, Analogy, and Awareness in Complex Reasoning." *Memory and Cognition*. https://doi.org/10.3758/BF03213292.

Shohamy, D., C. E. Myers, J. Kalanithi, and M. A. Gluck. 2008. "Basal Ganglia and Dopamine Contributions to Probabilistic Category Learning." *Neuroscience and Biobehavioral Reviews* 32 (2): 219–36. https://doi.org/10.1016/j.neubiorev.2007.07.008.

Sloutsky, V .M. 2003. "The Role of Similarity in the Development of Categorization." *Trends in Cognitive Sciences* 7 (6): 246–51. https://doi.org/10.1016/S1364-6613(03)00109-8.

Smith, E. E., and D. L. Medin. 1981. "Categories and Concepts." Cambridge, MA.: Harvard University Press. https://doi.org/10.2307/414206.

Smith, E. E., A. L. Patalano, and J. Jonides. 1998. "Alternative Strategies of Categorization." *Cognition* 65 (2–3): 167–96. https://doi.org/10.1016/S0010-0277(97)00043-7.

Smith, J. D. 2014. "Prototypes, Exemplars, and the Natural History of Categorization." *Psychonomic Bulletin & Review* 21 (2): 312–31. https://doi.org/10.3758/s13423-013-0506-0.

Smith, J. D., and J. P. Minda. 1998. "Prototypes in the Mist: The Early Epochs of Category Learning." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24 (6): 1411–36. https://doi.org/10.1037/h0090333.

Storms, G., P. De Boeck, and W. Ruts. 2000. "Prototype and Exemplar-Based Information in Natural Language Categories." *Journal of Memory and Language* 42 (1): 51–73. https://doi.org/10.1006/jmla.1999.2669.

The Boeing Company. 2000. "Operational Use of Angle of Attack on Modern Commercial Airplanes." *Aero* 12 (4): 11–22.

Walker, G. H., N. A. Stanton, P. M. Salmon, D. P. Jenkins, and L. Rafferty. 2010. "Translating Concepts of Complexity to the Field of Ergonomics." *Ergonomics* 53 (10): 1175–86. https://doi.org/10.1080/00140139.2010.513453.

Wiggins, M. W., D. Azar, J. Hawken, T. Loveday, and D. Newman. 2014. "Cue-Utilisation Typologies and Pilots' Pre-Flight and in-Flight Weather Decision-Making." *Safety Science* 65: 118–24. https://doi.org/10.1016/j.ssci.2014.01.006.

Wittgenstein, L. 1953. *Philosopical Investigations*. New York: Macmillan.

Understanding pilot response to flight safety events using categorisation theory

Clewley, Richard

2019-02-27

Clewley R, Nixon J. (2019) Understanding pilot response to flight safety events using categorisation theory. Theoretical Issues in Ergonomics Science, Volume 20, Issue 5, 2019, pp. 572-589
https://doi.org/10.1080/1463922X.2019.1574929