When Outcomes are not Enough: An Examination of Abductive and Deductive Logical Approaches to Risk Analysis in Aviation

Matthew Stogsdill*

While airlines generate massive amounts of operational data every year, the ability to use the collected material to improve safety has begun to plateau. With the increasing demand for air travel, the aviation industry is continually growing while simultaneously being required to ensure the level of safety within the system remains constant. The purpose of this article is to explore whether the traditional analysis methods that have historically made aviation ultra-safe have reached their theoretical limits or merely practical ones. This analysis argues that the underlying logic governing the traditional (and current) approaches to assess safety and risk within aviation (and other safety critical systems) is abductive and therefore focused on creating explanations rather than predictions. While the current “fly-fix-fly” approach has, and will continue to be, instrumental in improving what (clearly) fails, alternative methods are needed to determine if a specific operation is more or less risky than others. As the system grows, so too does the number of ways it can fail, creating the possibility that more novel accidents may occur. The article concludes by proposing an alternative approach that explicitly adds temporality to the concepts of safety and risk. With this addition, a deductive analysis approach can be adopted which, while low in explanatory power, can be used to create predictions that are not bound to analyzing only outcomes that have occurred in the past but instead focuses on determining the deviation magnitude between the operation under analysis and historically commensurate operations.

KEY WORDS: abductive analysis; deductive analysis; risk analysis

1. INTRODUCTION

Since the death of Lt. Thomas Selfridge¹ in 1908, aviation has invested substantial effort and resources in to understanding every aviation accident and fatality to prevent a similar event from reoccurring in the future. Through such efforts the Air Transport System (ATS) has become the safest mode of transportation in the world (European Union Agency for Railways [ERA], 2016), with the likelihood of being involved in a fatal commercial aircraft accident being one in 2.25 million based on 2018 data (Aviation Safety Network [ASN], 2019). Even with this impressive record, the ATS continues to strive to become even safer with the goal being to reduce the number of accidents by 80% by 2020 (PROactive Safety PERformance for Operations [PROSPERO], 2015). However, while the last several years have seen a reduction in the number of aviation related fatalities, the International Civil Aviation Organization’s (ICAO) 2018a safety report shows that the rate of accidents and fatal accidents per one million departures has remained relatively constant, largely due to
the increasing amount of traffic (International Civil Aviation Organization [ICAO], 2018a). ICAO’s guiding principle, “[i]mproving the safety of the global air transport system….” (ICAO, 2018a, p. 2) is accomplished by focusing on the following four areas: policy and standardization, monitoring of key safety trends and indicators, safety analysis, and implementing programs to address safety issues (ICAO, 2018a, p. 2). Two of these principles (monitoring of key safety trends and indicators and safety analysis) directly relate to the development and utilization of methods used to assess and improve safety within the ATS.

The ability to analyze trends, develop indicators, and conduct safety analyses, are all contingent upon having access to operational data, as well as maintenance, weather, and other sources of information. The ATS, and especially airlines, have access to substantial amounts of operational data from Quick Access Recorders (QAR), but without an underlying sense of how to integrate this multimodal and multisource data, access to these vast databases becomes almost meaningless. In the EU FP7 funded project, PROSPERO, operational actors (airlines and airports), researchers, data analysts, and technology developers worked together to identify industrial needs and determine what steps would be necessary to meet those needs (Baranzini et al., 2013; Baranzini, McDonald, & Corrigan, 2014; Ulfvengren et al., 2013). Of the identified issues, the most noteworthy was the need for metrics that could produce more proactive and ideally predictive insights (Ulfvengren et al., 2013).

While there have already been extensive discussions regarding the ontology of safety (see Aven, 2014; Haavik, 2014; Hollnagel, 2014; Leveson, 2011) and risk (Aven, Renn, & Rosa, 2011; Roeser, Hillerbrand, Sandin, & Peterson, 2013; Roeser, Hillerbrand, Sandin, & Peterson, 2012) the discussions predominantly take one of three basic perspectives:

1. What is the probability/likelihood that event \( Y \) will occur given context \( X \)?
2. How can event \( Y \) be prevented given context \( X \)?
3. How can a highly unlikely/unknown event be predicted/prevented given context?

Previously it would have been extremely difficult if not impossible to differentiate these questions from one another empirically due to a lack of data. However, in the current digital era, our ability to collect, store, and analyze data has increased rapidly; the limiting factor is currently how such collected data is used, not the availability of the data itself. While this article will focus on the aviation domain, the same underlying reasoning examined in this article can, with little modification, be applied within both the medical and nuclear fields.

The purpose of this article is to advocate for an alternative and complimentary understanding of the safety and risk concepts which will require the compilation of risk information that can then be refined into predictive insights by those with the relevant skills. Only with such knowledge can the types of predictive insights needed by industry be developed, and only with such knowledge can patterns embedded within the data be properly detected and interpreted. The remainder of this article endeavors to demonstrate why the traditional contemporary risk and safety analyses fail to deliver the desired predictive insights despite having substantial amounts of operational data and then discusses a possible approach that could fulfill that need.

2. CURRENT APPROACHES

QAR data is generally analyzed using exceedance/event detection methods which are used to detect deviations in the data that exceed a specified threshold. These metrics can be either a group or single variable (ATR, 2016; Civil Aviation Authority [CAA], 2013; International Civil Aviation Organization [ICAO], 2013a). Though exceedance detection methods have proven invaluable when conducting investigations of accidents, incidents, or near misses, they have thus far struggled to produce either the proactive or predictive information needed by industry due to the many different contextual elements that cannot be controlled for in the analysis. Several methods have tried to compliment exceedance methods in the past (ARMS Working Group [ARMS], 2010; International Civil Aviation Organization [ICAO], 2013a). Although exceedance detection methods have proven invaluable when conducting investigations of accidents, incidents, or near misses, they have thus far struggled to produce either the proactive or predictive information needed by industry due to the many different contextual elements that cannot be controlled for in the analysis. Several methods have tried to compliment exceedance methods in the past (ARMS Working Group [ARMS], 2010; International Civil Aviation Organization [ICAO], 2013a), by focusing on the use of precursor events (near-misses) as proxies for how novel accidents could occur. However, while these approaches have made progress in highlighting potential system failure points, as yet they have not been able to provide the types of predictive insights desired by industry (Ulfvengren et al., 2013).

As the complexity of the ATS and the combination of possible interactions continue to increase, the ability to rely upon precursor events to provide insights about future accidents/incidents will...
continue to be reduced. This likely will occur despite even more widespread use of routine data measurements which use lower thresholds than exceedance detection thereby allowing for more subtle trends to be monitored (CAA, 2013, p. 27). Unfortunately, the resultant problem is not simply the need for more sensitive threshold values, but an alternative approach that can complement currently used methods; an approach that can be used to understand how the variables individually and collectively change based upon known inputs and unknown contextual conditions.

Recently, there have been advances in the development of synthetic data generation techniques (Lališ et al., 2018), which can, in theory, help to overcome some of the underlying limitations of exceedance detection and routine monitoring methods. The method relies upon resampling the limited number of events in an effort to balance the successfulness bias inherent in the raw data. However, even with synthetic data generation it seems clear that (Lališ et al., 2018, p. 160):

...the synthetic data used to fill the gaps of existing limitations will never contain anything outside of what is inserted in the very equations behind the simulation. Even though they are based on expert assumptions and account for randomness, it is not possible to include all the variables, which affect the values of measured aviation safety data.

Lališ et al. (2018) make a convincing case for the use of synthetic data despite its limitations. However, the problem here may be more fundamental, since as Lališ et al. (2018, p.160) points out, since “…aviation is a socio-technical system, it is unlikely that the system is deterministic. Therefore, there is no ultimate set of assumptions and equations, which describe the system completely…” This makes the ability to determine a particular outcome such as a hard landing quite difficult both practically and theoretically.

Others have approached the problem differently, choosing to look at larger data sets instead of creating synthetic data. Two successful approaches relied upon clustering the data into similar groups. Li, Hansman, Palacios, and Welsch (2016) created a Gaussian mixture model in order to detect underlying patterns in normal flights so that abnormal events could be more easily detected. Oehling and Barry (2019) continued in this direction when they applied a local outlier probability (an unsupervised learning approach) to the problem in order to analyze a data set consisting of over 1 million flights. Though several cases (see Das, Li, Srivastava, & Hansman, 2012; Fernández et al., 2019; Li, Das, John Hansman, Palacios, & Srivastava, 2015) have been able to detect abnormalities in large data sets and have even highlighted events that were missed using the standard industry approaches (Oehling & Barry, 2019), the dynamic environment and contextual subtleties within the aviation domain makes these value-based clustering approaches a difficult path to take.

There has also been increasing interest in using other Artificial Intelligence (AI) methods to attempt to find patterns within data sets that have previously been too large for humans to realistically evaluate. However, methods such as neural networks, known as black box algorithms, are not interpretable by humans which makes using them to make life and death decisions morally questionable (Knight, 2017). Luckily, there is a growing field which is looking to create Explainable Artificial Intelligence (XAI) (Gunning, 2017). The goal of XAI is to ensure that AI algorithms are transparent, causal, bias free, fair, and highly reliable (Hagras, 2018, p. 29). These elements are needed so that humans can understand how and why an algorithm arrived at a particular solution, which is critical if lives will be trusted to those decisions. However, some of the most powerful AI algorithms are so sensitive that it can lead to overfitting, especially if inaccurate indicators are included in the data set (Paltrinieri, Comfort, & Reniers, 2019).

2.1. Flight Safety Foundation Task Force

In the report authored by the Flight Safety Foundation’s Approach-and-Landing Accident Reduction Task Force (Flight Safety Foundation [FSF], 1999), 76 approach-and-landing accidents and serious incidents, occurring between 1984 and 1997 were examined (along with over 3,000-line audits) (Flight Safety Foundation [FSF], 2009, 1999). While 66% of those events were found to have had unstabilized approaches, the findings need to be interpreted carefully as they are from a highly biased sample (i.e., flights that resulted in some type of landing event). Not only did the results of the study show that 33% of the flights that ended in an event had stabilized approaches, but when conducting line audits, it also found that the same errors were made during flights that resulted in normal landings as those with event landings (FSF, 1999, pp. 15, 46). Furthermore, since only accidents and serious incidents were examined, the outcomes should only be generalized to other flights that resulted in approach and landing
accidents or serious incidents and not necessarily to flight operations in general.

Thus, while the conclusion reached by the Flight Safety Foundations Task Force may well have been correct, that is that unstable approaches increase the likelihood of an event, the logic used to arrive at that conclusion seems partially flawed. What makes this remarkable, irrespective of any methodological oversight, is that the group (i.e., normal outcome vs. accident/incident/near-miss outcome) could be categorized only once the flight was completed. Thus, the analysis could only be conducted after the flight’s conclusion which confounds the results. In other words, by defining a flight as either “normal” or “abnormal” after the fact, the conducted analysis is trying to explain why a particular outcome occurred, rather than trying to predict an outcome.

2.2. What Do We Mean by “Prediction”?

This raises an interesting question about what constitutes prediction. ICAO currently classifies operational data analysis in one of three ways: descriptive, inferential (previously proactive), and predictive (ICAO, 2018b, pp. 6–2; ICAO, 2013b, pp. 2–26). The descriptive and inferential methodologies for identifying hazards are clearly defined in ICAO’s Safety Management Manual (SMM):

Descriptive – …statistics include measures of central tendency such as mean (average), median and mode, as well as measures of variability such as range, quartiles, minimum and maximum, frequency distributions, variance and standard deviation (SD). (ICAO, 2018b, pp. 6–3)

Inferential – …statistics include methods for estimating parameters, testing of statistical hypotheses, comparing the average performance of two groups on the same measure to identify differences or similarities, and identifying possible correlations and relationships among variables. (ICAO, 2018b, pp. 6–3)

The definition advanced in describing predictive approaches is less detailed and therefore more difficult to both understand and operationalize:

Predictive – …analyses extract information from historical and current data and use it to predict trends and behaviour patterns. The patterns found in the data help identify emerging risks and opportunities. Often the unknown event of interest is in the future, but predictive analysis can be applied to any type of unknown in the past, present or future. The core of predictive analysis relies on capturing relationships between variables from past occurrences and exploiting them to predict the unknown outcome. Some systems allow users to model different scenarios of risks or opportunities with different outcomes (ICAO, 2018b, pp. 6–3)

What is interesting to note here is the emphasis on “capturing relationships between variables from past occurrences and exploiting them to predict the unknown outcome” because this focus implies the use of “variable-centered” analytic approaches which Howard and Hoffman (2018, p. 848) describe as “the traditional and dominant approach in the social sciences, and its purpose is to explain relationships between variables of interest in a population.” Furthermore, the language is outcome focused, which indicates that abductive logic is being used (Goel & Joyner, 2015), suggesting that the problem is being worked backward. Rather than trying to determine what specific event could occur given a starting condition and the theorized control rule, the analysis instead tries to determine what the starting condition was that resulted in the unwanted outcome using the theorized control rule.

3. FROM ABDUCING EXPLANATIONS TO DEDUCING PREDICTIONS

While variable-centered abductive methods have helped make the ATS the mostly accident-free system we know today, such approaches have focused on making the system more reliable rather than safer (Leveson, 2011). To compliment the current approaches, it is beneficial to adopt the abstract and philosophical set of definitions for prediction and explanation put forward by Shanahan (1989):

Prediction: …is projection forwards from causes to effects (Shanahan, 1989, p. 1005), meaning that for a given rule a cause and effect can be calculated, and verified over time (e.g., hypothesis and theory testing)

Explanation: …is projection backward from effects to causes (Shanahan, 1989, p. 1005), and based upon known rules and known outcomes the most likely starting cause can be determined (e.g., accident investigation, medical diagnoses)

In defining prediction and explanations in terms of cause and effects, causality is explicitly introduced as a prerequisite for the development of prediction with temporality being strongly implied (Shanahan, 1989). The depiction of the explanation and prediction concepts in Fig. 1 help to illustrate the difference between the two definitions. Explanations are created using abductive logic (e.g., accident analyses and medical diagnoses) in which the unknown cause is adduced using known rules (theories) and outcomes
(effects), with the determined cause presumably being the best explanation for the given facts.

However, if the goal is to produce methods that can predict unwanted outcomes and to avoid hindsight bias, a deductive approach based only upon the information available to the operator at the time must be used. This perspective shift requires not only a change in the underlying logic being used, but also how individual sequences are viewed and compared with previously collected data. Howard and Hoffman (2018) describe such approaches as “person-specific” which are used to “…investigate effects that may be idiosyncratic to specific subjects” (Howard & Hoffman, 2018, p. 851). This is an important shift in thinking because it moves away from the idea that there are specific variable values that can be used to predict outcomes and instead begins to adopt a more contextually nuanced understanding of individual operational sequences as well as how risk increases or decreases over time depending upon the relationship between the operational context and the environment.

3.1. Flight Specific Perspective

In adopting a deductive flight-specific (i.e., person-specific) perspective, the flight’s evolution is examined over time. Thus, the temporal operational patterns become the unit of analysis rather than the operational outcome which allows for predictions to be made based upon the current operational context and the applicable control rules. This requires three temporal points to be identified: \( t_0 \) is when the operator makes the go/no-go decision; \( t_i \) is the current moment; and \( t_n \) is the anticipated time of the operation’s end.

3.1.1. Decision Gate

Of these temporal points, the most important is \( t_0 \), since it is at that moment when resources have been fully committed to an operation and discontinuing is no longer a simple task. It is also at this point that the operator has concluded that the operation can proceed safely. This sentiment is very much in keeping with Hollnagel’s intuitive understanding of safety, in which he argues (Hollnagel, 2014, p. 3):

> What we mean in general by ‘being safe’ is that the outcome of whatever is being done will be as expected. In other words, that things will go right, that the actions or activities we undertake will meet with success.

Thus, any operation that is started is considered (at least in the operator’s mind) to more likely succeed than fail. However, focusing only on the starting conditions will remove the possibility of hindsight bias but will be unlikely to provide differentiable results when comparing flights that end “normally” versus those ending in “events” (as was found in FSF, 1999). It is for this reason that the other two time points (\( t_i \) and \( t_n \)) are needed.

3.1.2. Current Moment

This next temporal point can take one of two forms depending upon whether the analysis is being done in real time or not. If done in real time, \( t_i \) represents the current moment, making it represent the maximum amount of information that can be processed in the analysis at that moment in time (i.e., the information known between \( t_0 \) and \( t_i \)). On the other hand, if this approach is being done after the flight has already been completed, then \( t_i \) represents what information was available at that moment during the flight. The number of possible future paths that end with \( t_n \) dramatically decrease as \( t_i \) approaches \( t_n \).

3.1.3. Anticipated Completion Time

The final temporal point is \( t_n \), which is the anticipated completion time for the operation (i.e.,
touchdown). This continually updated point in time is important because it gives a limit to the analysis preventing an infinite recursive loop. The amount of variation between the planned $t_n$ and the actual touchdown time is also of interest since, when aggregated, these values can provide a rough approximation as to how well the flight plan represents reality.

### 3.2. A Complimentary Approach

By defining the individual operation in terms of $t_0$, $t_i$, and $t_n$, a causal (information limited) system can be defined. This approach uses the information collected between $t_0$ and $t_i$ to determine how well it fits with previously recorded commensurate flight paths. In assessing the similarity (or difference) between the flight path being analyzed and the historical flight paths, a prediction can be made as to how likely it is that the analysis flight will conclude with an a known, and ideally desired, outcome.

For example, an aircraft approaching a runway with a pattern that is very similar to previously recorded flights approaching the same runway in similar conditions is unlikely to have a novel event occur. This is based on the logic that the pilots of the current flight are trained to a specified standard, and so long as the flight is within boundary conditions that tend to result in a desired outcome, odds are that this flight will also result in a desired outcome. However, if the flight begins to deviate too much from previously successful paths, the odds of an undesirable outcome will increase. Furthermore, if the path were to deviate too far beyond previously recorded paths, it should be seen as highly risky, since there would be no comparison points with which it could be compared. In this case the risk should be seen as a proportional increase given the amount of deviation.

In the proposed approach, a flight plan would be generated and then compared against a database of similar flights. As the flight progresses, the flight under analysis will continually be compared to the commensurate flights, with the historical flights that did not proceed in a similar way being constantly trimmed away, reducing the number of comparison flights in the next iteration. Fig. 2 shows the ranges of similar historically relevant paths (dotted lines), the historical average path (long short short-dashed line), the actual path (solid line), and potential future paths (dashed lines). The two future paths are depicting separate scenarios with the upper one showing a potential path leading away from the known context (making that path riskier), while the lower path is returning to the average path (making it the less risky option). In this way a flight’s deviation from a given norm can be seen as the flight evolves allowing for the individual events that may occur along its path to be analyzed using only the data that was available to the pilot at the time. While accident/incidents already examine the aircrafts path and parameter values in great depth, the proposed analysis will need to be operationalized in such a way to allow for the aircraft’s onboard computer to process and compare the aircraft flight path with historically commensurate ones.

### 4. DISCUSSION

The approach discussed above is different from other causal approaches because it remains focused on the similarity between sequences rather than on differences in outcomes based upon a specific
Fig. 3. New perspective of risk as a vector (left) as compared to traditional perspectives that view risk as a scalar (right).

decision point. This approach is felt to be inherently different from a treatment comparison approach like Rubin’s causal model (Rubin & John, 2016). In Rubin’s approach, the difference between a fact and counterfactual state can be determined by comparing the outcomes of subjects that had similar starting conditions; however, this is still focused on the existence (or nonexistence) of a decision (or treatment). These types of causal models remain focused on the outcome of the operation rather than on the process used to arrive there, suggesting that it is applying abductive logic to extrapolate if a decision did/did not have an impact on the outcome of interest.

Traditionally, statistical and scientific approaches have generally attempted to minimize false positives (Type-I) errors; however, due to the trans-scientific nature of the safety and risk concepts it seems that false negatives (Type-II) errors are of greater concern (Hansson & Aven, 2014). Since anticipating an accident and thereby averting one is far preferable to falsely believing everything is fine up until an accident occurs, it therefore seems more beneficial to shift the burden of proof from showing that a difference does exist to showing that a difference does not exist. This shift will require the reworking of some foundational ideas, in particular that safety and risk are temporal sequence-based concepts rather than being outcome based.

This proposed perspective, in the case of temporal analyses, is more conservative (and arguably more useful) since it precludes the need for a mutually exclusive and collectively exhaustive outcome taxonomy. Instead, the approach uses the simpler, though less precise, method of comparing the current operation to historically commensurate situations. This comparison is then evaluated to determine if the current and historical approaches are sufficiently alike to justify extrapolating a likely outcome. Using this perspective risk becomes a “vector” within a series distribution of previously observed paths at similar operational stages. This contrasts with the more traditional “scalar” interpretation currently used in many different industries including aviation (ICAO, 2018b, pp. 2–16). If a risk vector begins to diverge from the known (dark grey band in Fig. 3), then the amount of risk within the system increases as the operation enters a lesser-known context. In such a case, the governing relationships which normally work may become less effective or cease working altogether.

5. CONCLUSIONS

While the current methods can be used to create explanations as to the most likely reason an event did (or could) occur, this is insufficient for producing the predictive information necessary to ensure that the ATS can remain ultra-safe as the system continues to evolve at an exponential pace. This article has argued that for predictive information to be produced, a new perspective needs to be adopted which is based on the starting contexts of a specific process so that possible paths can be derived, instead of basing the “predictions” on the aggregated differences of historical events.

There are numerous potential benefits in separating the prediction and explanation methods. First the terminology issue that has been plaguing the safety and risk communities for years can be somewhat mitigated, as differentiating the two approaches gives rise to an opportunity to create a more concise set of nonconflicting terminologies. Second,
this separation also permits the occasionally subtle but always important distinctions between the following processes to be more clearly delineated: safety/risk assessment, information/knowledge collection, qualitative and quantitative modeling, and coping with uncertainties (Aven & Zio, 2014). Hopefully by creating separate but codependent approaches, some current semantic and conceptual debates which Boholm (2017) discusses can be reevaluated.

By arguing that the current approaches are theoretically ill suited for predicting either novel events or events for a single operation, this article highlights a potential path forward that can more clearly discriminate between what can and cannot be done based upon the theory underpinning current methodologies. It is worth reiterating that the current approaches have served aviation (and other safety critical systems) well in the past and will continue to play an invaluable role in protecting people in the future. However, the current methods do have limits which the industry is interested in overcoming thereby allowing the system to improve without relying primarily on the “fly-fix-fly” approach of the past (Leveson, 2011).

ACKNOWLEDGMENT

The author would like to thank the reviewers for helping to clarify the paper’s message, and the members of the FP7 PROSPERO project for their insights which helped formulate the ideas presented above.

REFERENCES

ARMS Working Group (ARMS). (2010). The ARMS methodology for operational risk assessment in aviation organisations. Aviation Safety Network (ASN). (2019). Aviation safety network releases 2018 airliner accident statistics. Retrieved from https://news.aviation-safety.net/2019/01/01/aviation-safety-network-releases-2018-airliner-accident-statistics/

ATR. (2016). Flight Data Monitoring: On ATR Aircraft. Aven, T., & Zio, E. (2014). Foundational issues in risk assessment and risk management. Risk Analysis, 34(7), 1164–1172. https://doi.org/10.1111/risa.12132

Aven, T. (2014). What is safety science. Safety Science, 67, 15–20. https://doi.org/10.1016/j.ssci.2013.07.026

Aven, T., Renn, O., & Rosa, E. A. (2011). On the ontological status of the concept of risk. Safety Science, 49(8–9), 1074–1079. https://doi.org/10.1016/j.ssci.2011.04.015

Baranzini, D., McDonald, N., Corrigan, S., Kontogiannis, T., Zanin, M., Ureta, H., … Nugent, T. (2013). Functional specification for the design and preparation of PROSPERO approach to safety performance management (D3.6).

Boholm, M. (2017). The semantic field of risk. Safety Science, 92, 207–216. https://doi.org/10.1016/j.ssci.2016.10.011

Civil Aviation Authority (CAA). (2013). CAP 739: Flight data monitoring.

Das, S., Li, L., Srivastava, A., & Hansman, R. J. (2012). Comparison of algorithms for anomaly detection in flight recorder data of airline operations. Proceedings from 12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Virgina: Reston.

European Union Agency for Railways (ERA). (2016). Railway safety performance in the European Union 2016. Retrieved from http://books.google.se/books?id=sZBLvQEACAAJ&hl=&source=gbs_api

Fernández, A., Martínez, D., Hernández, P., Cristóbal, S., Schwaiger, F., Nunez, J. M., & Ruiz, J. M. (2019). Flight data monitoring (FDM) unknown hazards detection during approach phase using clustering techniques and AutoEncoders. Paper presented at 9th SESAR Innovation Days December 2–5, Athens, Greece.

Flight Safety Foundation (FSF). (1999). Killers in aviation: FSF task force presents facts about approach-and-landing and Controlled-Flight-into-terrain accidents. Flight Safety Digest, 17, 11–12.

Flight Safety Foundation (FSF). (2009). Stabilized approach (FSF ALAR Briefing Note 7.1). Alexandria, VA: FSF

Goel, A., & Joyner, D. (2015). Knowledge-based artificial intelligence/cognitive systems. Retrieved from https://classroom.udacity.com/courses/ud409/lessons/1951908693/concepts/19221891400923

Gunning, D. (2017). Explainable artificial intelligence (xai). Defense Advanced Research Projects Agency (DARPA), nd Web, 2, 2.

Haavik, T. K. (2014). On the ontology of safety. Safety Science, 67, 37–43. https://doi.org/10.1016/j.ssci.2013.09.004

Hagras, H. (2018). Toward human-understandable, Explainable AI. Computer, 51(9), 28–36. https://doi.org/10.1109/MC.2018.3629065

Hansson, S. O., & Aven, T. (2014). Is risk analysis scientific. Risk Analysis, 34(7), 1173–1183. https://doi.org/10.1111/risa.12230

Hollnagel, E. (2014). Safety-I and Safety-II. Ashgate Publishing, Ltd. Retrieved from http://books.google.se/books?id=8naUBAAQBAJ&hl=&source=gbs_api

Howard, M. C., & Hoffman, M. E. (2018). Variable-centered, person-centered, and person-specific approaches. Organizational Research Methods, 21(4), 846–876. https://doi.org/10.1177/1094428117744021

International Civil Aviation Organization (ICAO). (2013a). Flight data analysis programme manual. Quebec, Canada: International Civil Aviation Organization.

International Civil Aviation Organization (ICAO). (2013b). Safety management manual. Quebec, Canada: International Civil Aviation Organization

International Civil Aviation Organization (ICAO). (2018a). ICAO safety report 2018. Quebec, Canada: International Civil Aviation Organization

International Civil Aviation Organization (ICAO). (2018b). Safety management manual. Quebec, Canada: International Civil Aviation Organization

Knight, W. (2017). The dark secret at the heart of AI. MIT Technology Review, 120(3). Retrieved from https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/

Lališ, A., Socha, V., Krmen, P., Vittek, P., Socha, L., & Kraus, J. (2018). Generating synthetic aviation safety data to resample
or establish new datasets. *Safety Science, 106*, 154–161. https://doi.org/10.1016/j.ssci.2018.03.013

Leveson, N. G. (2011). *Engineering a safer world: Systems thinking applied to safety (engineering systems)*. Cambridge, MA: The MIT Press.

Li, L., Das, S., John Hansman, R., Palacios, R., & Srivastava, A. N. (2015). Analysis of flight data using clustering techniques for detecting abnormal operations. *Journal of Aerospace information systems, 12*(9), 587–598.

Li, L., Hansman, R. J., Palacios, R., & Welsch, R. (2016). Anomaly detection via a Gaussian mixture model for flight operation and safety monitoring. *Transportation Research Part C: Emerging Technologies, 64*, 45–57. https://doi.org/10.1016/j.trc.2016.01.007

Oehling, J., & Barry, D. J. (2019). Using machine learning methods in airline flight data monitoring to generate new operational safety knowledge from existing data. *Safety Science, 114*, 89–104. https://doi.org/10.1016/j.ssci.2018.12.018

Paltrinieri, N., Comfort, L., & Reniers, G. (2019). Learning about risk: Machine learning for risk assessment. *Safety Science, 118*, 475–486.

PROactive Safety PERformance for Operations (PROSPERO). (2015). PROSPERO Homepage. Retrieved from http://www.prosperop7.eu

Roeser, S., Hillerbrand, R., Sandin, P., & Peterson, M. (Eds.). (2012). *Handbook of risk theory*. Heidelberg, Germany: Springer.

Roeser, S., Hillerbrand, R., Sandin, P., & Peterson, M. (Eds.). (2013). *Essentials of risk theory*. Heidelberg, Germany: Springer.

Rubin, D. B., & John, L. (2016). Rubin causal model. In M. Lovric (Ed.), *International encyclopedia of statistical science* (pp. 1263–1265). Heidelberg, Germany: Springer.

Shanahan, M. (1989). Prediction is deduction but explanation is abduction. *Proceedings of the 11th international joint conference on Artificial intelligence, 2*, 1055–1060.

Ulfvengren, P., McDonald, N., Kontogiannis, T., Malakis, S., Cacciabue, C., Cassani, M., … Frau, G. (2013). Requirements for Proactive Safety Performance for Operations at Organizational & Systemic Levels (D2.2). Retrieved from https://cordis.europa.eu/project/id/314822/reporting