2018

Manned Versus Unmanned Aircraft Accidents, Including Causation and Rates

Justin A. Jaussi
Embry-Riddle Aeronautical University, jaussij@my.erau.edu
Herbert O. Hoffmann
Embry-Riddle Aeronautical University, hoffm7d5@erau.edu

Follow this and additional works at: https://commons.erau.edu/ijaaa

Part of the Aviation Safety and Security Commons

In 2015, Ray Mabus, the U.S. Secretary of the Navy, stated that the F-35C Lightning II "should be, and almost certainly will be, the last manned strike fighter aircraft the Department of the Navy will ever buy or fly" (Myers, 2015, para. 2). This prophecy may provide an insight into the large-scale transition from manned to unmanned aircraft occurring within the U.S. Department of Defense (DoD), but what does this imply?

This paradigm shift is expected not only to revolutionize military flight operations, but also to be a multi-modal shift throughout aviation. A 2013 U.S. Department of Transportation report concluded that Unmanned Aircraft Systems operations will likely surpass manned commercial and military aircraft operations by 2035 (U.S. Department of Transportation, 2013, p. 7).

Aircraft reliability will be a significant hurdle to this expansion. Questions of unmanned aircraft reliability are often paired with a criticism of their safety record, including accident causation and rates. A thorough understanding of unmanned aircraft accident data is needed by policy makers to make accurate decisions regarding the development of regulatory requirements for unmanned aircraft. Because policy makers are influenced by their constituents, an accurate characterization of the data in the public forum is also needed.

Research Problem

The discussion of accident statistics (or "mishap" statistics, in military terms) of medium and large unmanned aircraft is often incomplete. The MQ-1 Predator and the MQ-9 Reaper are often part of these discussions because they have flown the bulk of the hours of U.S. Air Force unmanned aircraft of similar size (U.S. Air Force Safety Center, 2018).

Comparing the mishap rate of one aircraft to another is mostly accomplished by showing the current cumulative rate, or reporting the number of mishaps that have occurred over a recent period. For example, "during the 5-yr period ending in fiscal year 2005, Predator RPAs [Remotely Piloted Aircraft] were lost at a rate of 14.2 aircraft per 100,000 flight hours as compared to a rate of 3.0 for the U-2 and 1.6 for the A-10" (Tvaryanas & Thompson, 2008, p. 526). Although Tvaryanas and Thompson's statement might have been accurate with the data available at the time (and the article indicates the comparison was chosen due to similarities in mission type), it does not consider platform maturity. During the referenced period, U.S. Air Force Predator RPAs increased total flight hours from 17,476 to 140,735; the U-2 increased from 387,429 to 452,097 flight hours; and the A-10 from 3,768,311 to 4,350,290 flight hours (U.S. Air Force Safety Center, 2018). As an aircraft platform matures by accumulating more flight hours, the cumulative mishap rate generally

decreases (U.S. Air Force Safety Center, 2018). Therefore, a comparison of the mishap or loss rate of an aircraft with approximately 140,000 flight hours to one with more than 4 million does not provide a comprehensive picture of the data.

To better understand the comparison of manned versus unmanned aircraft safety data, this paper used a chi-square goodness of fit method to compare the initial progression of the MQ-9 Reaper Class A mishap rate to that of several U.S. Air Force manned aircraft. Although not a typical use of the chi-square approach, the method is used as an analytical tool to describe and quantify the comparisons, rather than make statistical conclusions.

Definitions

Aircraft accident. "An occurrence associated with the operation of an aircraft which takes place between the time any person boards the aircraft with the intention of flight and all such persons have disembarked, and in which any person suffers death or serious injury, or in which the aircraft receives substantial damage" (from Title 49 of the U.S. Code of Federal Regulations, Part 830). Part 830 indicates that in this definition, "aircraft" includes unmanned aircraft.

Mishap. An "unplanned occurrence, or series of occurrences, that results in damage or injury and meets Class A, B, C, or D reporting criteria" (U.S. Air Force, 2017b, p. 102).

Mishap rate. The average number of mishaps that have occurred over a defined number of flight hours, usually per 100,000 flight hours (U.S. Air Force Safety Center, 2018).

Class A mishap. A mishap resulting in direct costs of $2,000,000 or more, loss of aircraft, a fatality, or permanent total disability (U.S. Air Force, 2017b, p. 20).

Class B mishap. A mishap resulting in direct costs of $500,000 or more but less than $2,000,000, a permanent partial disability, or inpatient hospitalization of three or more personnel (U.S. Air Force, 2017b, p. 20).

Class C mishap. A mishap resulting in direct costs of $50,000 or more but less than $500,000, and other miscellaneous loss of work definitions (U.S. Air Force, 2017b, pp. 20–21).

Class D mishap. A mishap resulting in direct costs of $20,000 or more but less than $50,000, and other miscellaneous loss of work definitions (U.S. Air Force, 2017b, p. 21).

Unmanned Aircraft (UA). A broad spectrum from weather balloons to highly complex aircraft piloted from remote locations by licensed pilots or operators. The latter are group of the unmanned aircraft are often referred to as "remotely piloted aircraft" or RPAs (International Civil Aviation Organization, 2017, p. 1).

Unmanned Aircraft System (UAS). An aircraft and its associated elements which are operated with no pilot on board (International Civil Aviation Organization, 2017).

Review of Relevant Literature

The following literature was reviewed, highlighting unmanned mishap causation and rates. Two mishap case studies that demonstrate differences in manned vs. unmanned aircraft mishap causation are also included.

Progression Toward Increasing Autonomous Unmanned Aircraft Control

Aircraft control has become increasingly automated over the past century. The aircraft autopilot was first introduced in 1914 (Thompson, 2014); furthermore, automated landings began as early as 1937 (Larson, 2012) and have been used by commercial airlines since the mid-1960s (Bearup, 2015). The Automatic Carrier Landing System (ACLS) used by the U.S. Navy's F/A-18 Hornet fighter/attack aircraft was certified for fleet-wide use in 1984. ACLS improves safety by being able to land the aircraft in difficult circumstances such as low visibility, deck motion due to high seas, and high turbulence along the approach path (Urnes & Hess, 1985, pp. 289, 295).

This trend of increasing autonomous aircraft control will also be significant in the transition from manned to unmanned aircraft. When discussing the future of unmanned systems, the 2013 DoD *Unmanned Systems Integrated Roadmap* indicated that the long-term future state (beyond 2020) will include higher levels of automation (p. 72). Lee contended that autonomy will indeed be the driving force for new remotely piloted aircraft technology, and that distrust of autonomy will be a limiting factor of this progress (2011, p. 79).

Lee also concluded that military unmanned aircraft will require more autonomy to operate in the increasingly complex aviation environment and in contested military airspace (2011). To remain competitive in military aviation, increasing autonomous control of unmanned aircraft will likely be required. Straub concluded that law enforcement agencies might one day have no choice but to use armed unmanned aircraft to counter the equipment and tactics of criminals (2014).

Lee further proposed that stages or degrees in the progression toward more autonomy, rather than an "all-or-nothing" approach, can mitigate doubts about autonomous unmanned aircraft operations (2011, p. 77). Much of Lee's discussion on this progression is based on a 1978 Massachusetts Institute of Technology study by Sheridan and Verplank that presented 10 levels of automation, ranging from no computer assistance to the computer performing the entire task (pp. 79–80).

More autonomous control of unmanned aircraft could result in a loss of jobs, particularly for pilots. In a 2013 University of Oxford study, Frey and Osborne concluded that most workers in transportation and logistics industries (for example, taxi drivers and delivery drivers) face a high risk of being replaced by computerization. The concept of self-driving commercial automobiles can logically be extended to self-driving commercial aircraft.

This movement toward unmanned commercial aircraft may also likely occur in stages. Lim, Bassien-Capsa, Ramasamy, Liu, and Sabatini (2017) contended that the current two-pilot requirement for commercial airlines may transition to single-pilot operations supported by a ground flight crew. The authors further concluded that this will likely begin with cargo aircraft operations, and that ultimately, we will likely see remotely piloted commercial airlines.

Ultimately, increasing autonomous aircraft control will of course affect how aircrew are trained. While automation can reduce crew workload and improve safety, it does not necessarily reduce operator training. In fact, one of the key "ironies of automation" described by Bainbridge is that the most successful automated systems may be the ones that need the greatest investment in operator training (Bainbridge, 1983).

Mishap Causation ("Why")

Pilot error has long been found to be a primary cause in a significant portion of aircraft accidents (Aircraft Owners and Pilots Association [AOPA] Air Safety Institute, 2017). In this section, first, how often pilot/crewmember errors contribute to manned and unmanned aircraft accidents will be discussed. Next, the general classifications or the types of crew errors will be analyzed, including which are the most common.

Pilot/crewmember error contributing to manned aircraft accidents. The AOPA Air Safety Institute's Joseph T. Nall Report is an industry-recognized resource on general aviation accident statistics. The most recent edition was released in 2017. It focuses on data from 2014 (AOPA Air Safety Institute, 2017). The statistics in the report include general aviation accidents, or all flight activities

in the NAS and flights to and from the United States and its territories, excluding those by the DoD or scheduled airlines. Also excluded are fixed-wing aircraft weighing more than 12,500 pounds, weight-shift control aircraft, powered parachutes, gyroplanes, gliders, airships, balloons, and unmanned aircraft.

The Nall Report (2017) provides a breakdown for commercial and non-commercial general aviation accident causes (commercial flights in this context include charter, cargo, crop-dusting, and external load flights, but not, as mentioned, scheduled airlines). Three broad categories of causation are used: pilot-related, mechanical/maintenance, and other/unknown. The Nall Report indicates that 75% of the 952 non-commercial and 63% of the 68 commercial accidents in 2014 were pilot-related. The report concludes that these numbers are almost identical to every year in recent memory.

Pilot/crewmember error contributing to <u>un</u>manned aircraft accidents. A review of the pilot/crewmember contribution to MQ-1 and MQ-9 unmanned aircraft accidents was conducted. The review focused on journal articles that employ the Human Factors Analysis and Classification System (HFACS) accident causation taxonomy. HFACS was developed by Dr. Douglas Wiegmann and Dr. Scott Shappell to determine why accidents happen and how to reduce accident rates (HFACS, 2014). The DoD implements a version of HFACS to identify hazards and risks that result in mishaps (U.S. Department of Defense, 2005a).

In a 2004 Department of Transportation/FAA study, Williams analyzed, among other unmanned aircraft accidents, 15 MQ-1 Predator Class A mishaps that occurred between December 1999 and December 2003. The study relied on executive summaries of the mishaps from the Air Force Judge Advocate General Corps website, Air Force Accident Investigation Board (AIB) reports, and previous work on Air Force accidents and human factors issues by Tvaryanas (Williams, 2004). Of the MQ-1 mishaps analyzed, 67% had human factors causes, 42% had aircraft causes, and 17% had maintenance causes. Of those attributed to human factors causes, 75% were due to procedural errors.

In a 2006 study, Tvaryanas, Thompson, and Constable conducted a 10-year (1993 to 2003) cross-sectional, quantitative HFACS analysis of several types of unmanned aircraft. The U.S. Air Force unmanned aircraft in the study (primarily 38 RQ-1 mishaps, but also 4 RQ-4 Global Hawk mishaps—maximum gross weight of 32,250 pounds, see U.S. Air Force, 2014) had 79.1% that involved human factors causes.

Tvaryanas and Thompson further analyzed unmanned aircraft mishap statistics in a 2008 study. This analysis differed from the 2006 study in that only

Air Force MQ-1 Predator mishaps were analyzed (for simplicity, unless otherwise stated, "MQ-1" is used to represent both the RQ-1 and MQ-1 throughout this paper). The study relied on U.S. Air Force Safety Center data and analyzed all MQ-1 mishaps with direct costs totaling $20,000 or more during fiscal years 1997 to 2005, a total of 54 mishaps. In addition, 41 "near misses" were included in the analysis, bringing the total events considered to 95. As in the 2006 study, the mishaps/incidents were analyzed using the HFACS taxonomy. The study found that 50.5% of the mishaps or incidents included crewmember failures.

In a 2013 study, Giese, Carr, and Chahl conducted a HFACS analysis categorizing 52 MQ-1 and MQ-9 Class A mishaps that occurred from 2004 to 2012. The study analyzed U.S. Air Force AIB mishap reports and found that 42% of the mishaps involved human error. Of the ones involving human error, 59% had operator error as the main cause.

Types of crewmember errors in manned aircraft accidents. HFACS is based in part on Reason's "Swiss cheese" accident causation model: when holes in the defense layers intended to prevent accidents line up, the mishap occurs (Reason, 2000, p. 769). Reason's model includes four levels of human error, with each level influencing the next: organizational influences, unsafe supervision, preconditions for unsafe acts, and unsafe acts (Wiegmann & Shappell, 2003). HFACS provides a framework to apply these four levels in real settings and further divides the four levels into 19 causal categories, as shown in Figure 1. Some derivations of the HFACS taxonomy include additional sublevels beyond these 19, for example, the "nanocodes" used in the DoD-HFACS (Cohen, Wiegmann, & Shappell, 2015, p. 731). The DoD's continued commitment to HFACS was demonstrated by the Naval Safety Center's recent launch of a DoD-HFACS mobile application ("app") via the Apple and Android app stores. The app is intended to replace the existing HFACS flipbook used to identify DoD-HFACS codes (Glover, 2018).
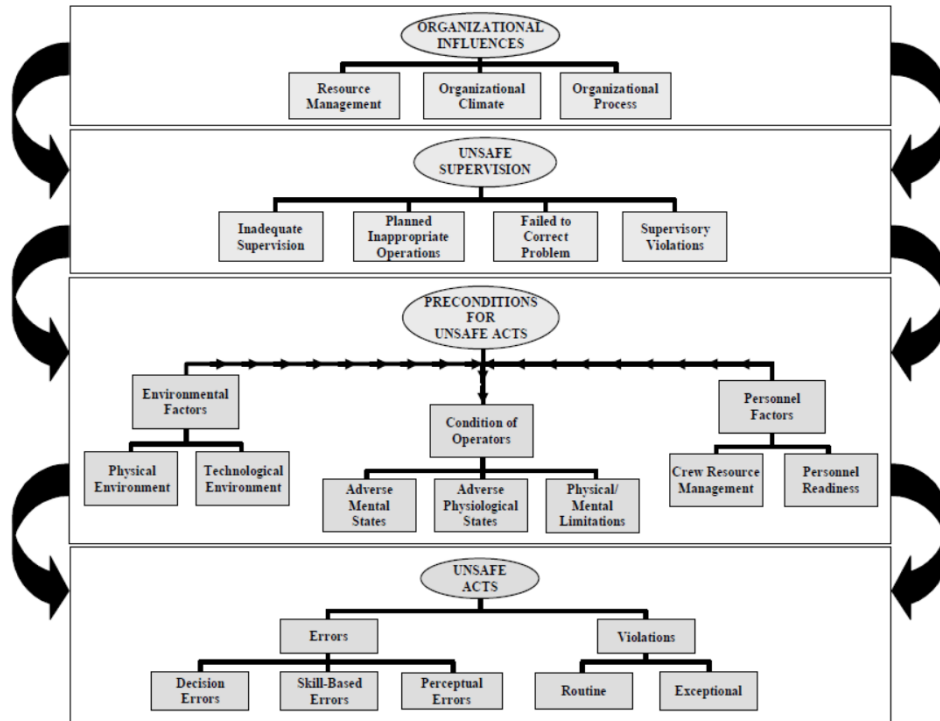
Figure 1. The 4 levels and 19 causal categories of the HFACS framework. Reprinted from Shappell & Wiegmann, 2004, with permission.

HFACS founders Shappell and Wiegmann used their taxonomy in a 2004 study to compare human error across military, commercial carrier, and general aviation operations. The study comprised a meta-analysis of over 16,000 accidents from approximately 1990 to 1998. According to the authors, it was the first time that commercial and military data was compared beyond overall accident rates or the overall percentages of accidents associated with human error, i.e., beyond the discussion thus far in this report.

Shappell and Wiegmann's 2004 study focused on unsafe acts of aircrew—the bottom of the four HFACS tiers shown on Figure 1. HFACS divides unsafe acts into errors and violations. The errors are further divided into skill-based errors, decision errors, and perceptual errors (the bottom left portion of the HFACS diagram in Figure 1).

The study found that the three subcategories of crew error, from most common to least common, were the same across all three broad categories of manned aircraft aviation. The most common were skill-based errors, then decision errors, and finally perceptual errors.

Types of crewmember errors in <u>unmanned</u> aircraft accidents. The three unmanned aircraft HFACS studies previously mentioned—Tvaryanas, Thompson, and Constable (2006), Tvaryanas and Thompson (2008), and Giese et al. (2013)—also included an analysis of unsafe acts in MQ-1 and/or MQ-9 mishaps or incidents. Each of these three studies found the same order of most to least common in the HFACS subcategories of MQ-1/MQ-9 operator error as Shappell and Wiegmann's 2004 cross-segment study of manned aircraft: skill-based errors were most common, then decision errors, and then perceptual errors.

These studies also discussed the interrelatedness of crew errors to automation, HMI/control station design, and crew training (Tvaryanas & Constable 2006; Tvaryanas & Thompson, 2008; Giese et al., 2013). When addressing these issues, Tvaryanas and Thompson's 2008 study described how the Predator was initially acquired by the U.S. Air Force as an "advanced concept technology demonstration" and how the resulting rapid acquisition process affected ground control station design, procurement of technical publications, and the development of initial training programs (p. 529). Similarly, Giese et al. concluded that "the pressure of immediate usage for [the] Predator in combat theaters resulted in early Operational Test and Evaluation configurations of the system being ordered in significant numbers," resulting in a lack of maturity of certification requirements when compared to civil aircraft (2013, p. 1191).

Summary – mishap causation. In summary, a significant portion of manned aircraft mishaps have historically been attributed to pilot or crewmember error. A similar trend is evident in the unmanned aircraft data or those aircraft considered in this study (primarily the MQ-1 and MQ-9). Of the pilot/crewmember accidents, manned and the unmanned aircraft considered have the same breakdown of types of crew errors, from most common to least common. Skill-based errors were the most common.

Case Studies Highlighting Causation

While the general breakdown of (1) *percentage*, and (2) *type* of crewmember errors in manned and unmanned aircraft accidents appears similar, *how* these errors are manifested can be quite different. Further statistical analysis on causation is beyond the scope of this study. But, to demonstrate causation differences, two unmanned aircraft accidents will be analyzed in this section. The first was an MQ-1 Predator mishap that illustrates how the pilot not being onboard the aircraft might affect aircrew decision making. The second was an MQ-9 Reaper mishap that underscores the unique challenge of stall recognition pilots face when flying an unmanned aircraft.

U.S. Air Force MQ-1 mishap, tail number 05-3154, March 29, 2013. The first accident to be analyzed was an MQ-1 mishap that occurred on March 29, 2013, northeast of Kandahar, Afghanistan. All information about the mishap described here is based on the U.S. Air Force Accident Investigation Board (AIB) publicly releasable documents and report (U.S. Air Force, 2013).

MQ-1 overview. First, a brief description of the MQ-1 Predator. It has a wingspan of 55 feet and a maximum takeoff weight of 2,250 pounds. Powered by a turbocharged Rotax piston engine, the MQ-1 can reach altitudes up to 25,000 feet Mean Sean Level (MSL). A crew comprised of a pilot and sensor operator flies the aircraft from a ground control station via a line-of-sight datalink, or a satellite datalink for beyond line-of-sight operations. The MQ-1 began service with the U.S. Air Force in 1996 as the RQ-1. With the addition of AGM-114 Hellfire missiles in 2002, it was re-designated the MQ-1 (the "R" indicates a reconnaissance aircraft; "M" is for multi-role; "Q" indicates a remotely piloted aircraft system) (U.S. Air Force, 2015a). In March of 2018, the U.S. Air Force retired the MQ-1 (Losey, 2018). While in service, it was used for armed reconnaissance, airborne surveillance, and target acquisition, and could carry up to 450 pounds of payload (U.S. Air Force, 2015a).

Mishap sequence. The mishap flight was an armed Intelligence Surveillance and Reconnaissance (ISR) mission. The Launch and Recovery Element (LRE, for takeoff and landing) was operated from the airfield in Afghanistan. The Mission Control Element (MCE) was operated from Fargo, ND, via a satellite datalink. Aircraft control was transferred from the LRE to the MCE at 0535 Universal Time Coordinated (UTC), or approximately 25 minutes after takeoff (pp. 5, 6).

The weather forecast for the mission area included showers, few thunderstorms, and light icing. Nearly three hours into the flight, at approximately 0800 UTC, a weather technician from the Squadron Operations Center (SOC) informed the mishap crew of an inbound weather front and recommended they depart the aircraft from the mission area at 1000 UTC to avoid the approaching hazardous weather (pp. 6, 10).

At 0945 UTC, the mishap aircraft was requested to support a Forward Operating Base (FOB) expecting an imminent hostile attack. To support the FOB, the crew and weather technician agreed to delay the departure from the mission area an additional 30 minutes, but no later than 1030 UTC (pp. i, 6).

At 0956 UTC, a lightning ground strike was observed within 10 nautical miles of the mishap aircraft. The crew began to depart the aircraft from the mission

area at 1028 UTC. Towering clouds began to surround the mishap aircraft, and it flew into clouds at 1035 UTC. The datalink from the aircraft ceased at 1039 UTC. Radar data indicated the aircraft departed controlled flight and entered a stall or spin. The aircraft was destroyed upon impact (pp. 6, 7, 10).

Mishap cause. The AIB concluded a lightning strike damaged the aircraft electrical circuits, resulting in the loss of the satellite datalink and loss of the aircraft. The aircraft manufacturer's report to the Air Force (a partially redacted version was included in the AIB documents) stated that returned hardware from other MQ-1 lightning strike events indicated that a lightning strike can damage the MQ-1's onboard electronics (pp. i, JJ-8).

Using DoD-HFACS, the AIB assessed the accident for contributing human factors. They found no evidence that human factors as described in the HFACS taxonomy were a factor in the mishap (p. 13).

This case study highlights how pilot decision making might be affected by not being onboard the aircraft. The decision to delay departing the aircraft from an area of deteriorating weather might have been influenced by the crew having no physical danger from a weather-induced crash. One could argue that "pushing the limits" with weather-related decisions—because the aircrew are not onboard the aircraft—could be influenced by the organizational climate as described in HFACS. The point could also be made that this type of decision is one of the five hazardous attitudes (possibly "invulnerability") that contribute to poor judgment, as described by the FAA (U.S. Department of Transportation, 2016b, p. 2-5). While this could be true in some scenarios, for this mishap, the AIB concluded that the human factors described in HFACS were not applicable.

U.S. Air Force MQ-9 mishap, tail number 10-4113, June 7, 2016. The second case study analyzed here was an MQ-9 mishap that occurred during a proficiency flight on June 7, 2016, in the Nevada Test and Training Range, located northwest of Las Vegas, NV. This summary is based on the U.S. Air Force AIB publicly releasable report (U.S. Air Force, 2016). This mishap highlights how the pilot not being onboard the aircraft can affect his or her ability to recognize a decreasing energy state of the aircraft and an impending stall.

MQ-9 overview. The MQ-9 is powered by a Honeywell turboprop engine and has a wingspan of 66 feet. It has a maximum takeoff weight of 10,500 pounds and a ceiling of 50,000 feet MSL. Like the MQ-1, the MQ-9 is operated by a pilot and sensor operator using a line-of-sight or satellite datalink. The MQ-9 has been in service with the U.S. Air Force since 2007. It can carry 3,750 pounds of payload

including AGM-114 Hellfire missiles, GBU-12 Paveway II laser-guided bombs, and GBU-38 Joint Direct Attack Munitions (U.S. Air Force, 2015b).

Mishap sequence. This was a stall/spin mishap that occurred approximately two minutes after aircraft control was transferred from the LRE to the MCE. During the transfer, the aircraft was flying at 8,500 feet MSL. The MCE crew did not realize the minimum acceptable autopilot altitude was set to 9,000 feet MSL. As the MCE crew completed their handover checklist, they unknowingly engaged the 9,000-foot minimum altitude setting, causing the aircraft autopilot to initiate a climb from 8,500 feet MSL.

The mishap pilot recognized the climb, but incorrectly attributed it to a malfunction or unexpected flight condition. To quickly take manual control of the aircraft, the pilot used the "landing configuration" command, which disables all autopilot modes and the aircraft automated stall protection system. The pilot initially reduced power to descend the aircraft toward 8,000 feet, then resumed completing the handover checklist, but did not adequately adjust aircraft pitch attitude.

As angle of attack continued to increase, the pilot failed to notice the audible and visual stall warnings. When the sensor operator correctly assessed the situation and advised the pilot that the aircraft was in a stall, the pilot increased engine power, but did not apply the stall recovery procedures outlined in the MQ-9 Flight Manual. Angle of attack increased until the aircraft entered a full aerodynamic stall and then spiraled to the ground.

Mishap cause. The AIB determined the cause of the mishap was a combination of (1) the pilot incorrectly prioritizing completing the handover checklist, and (2) the pilot's failure to observe the warnings of a reduced energy state and stall and then implement stall recovery procedures.

The AIB also used HFACS to categorize the human factors associated with this mishap. They concluded that the DoD-HFACS nanocode AE202 *Task Misprioritization* (a subset of Decision Errors, within Unsafe Acts) was applicable. The *Task Misprioritization* nanocode is when "the individual does not organize, based on accepted prioritization techniques, the tasks needed to manage the immediate situation" (U.S. Department of Defense, 2005a, p. 9).

Stall recognition. Regarding stall recognition, the FAA's *Airplane Flying Handbook* (U.S. Department of Transportation, 2016a) describes how manned aircraft pilots, over time, develop a "feel" for the airplane. This "feel" includes kinesthesis—the ability to sense movement—but also includes other sensory input

like hearing air moving across the airframe and feeling engine vibration. This "feel" can provide the manned aircraft pilot critical information about changes in aircraft speed and direction and warn the pilot of an impending stall (pp. 3-4, 4-5, 4-6). Developing this "feel" is a foundational skill emphasized early in pilot training and required for pilot certification. During the practical test to earn a private pilot certificate, the applicant must demonstrate knowledge of impending stall and full stall indications including how to recognize a stall by sight, sound, and feel (U.S. Department of Transportation, 2017).

Without these inputs, the unmanned aircraft pilot is relatively sensory deprived, lacking visual and auditory cueing (Tvaryanas et al. 2006). Williams concluded that unmanned aircraft pilots have the challenge of a general lack of sensory cues including delayed control feedback and a small field of view (2006). The MQ-9 stall/spin mishap described here highlights this sensory deprivation and the need for pilot training specifically tailored to the unmanned aircraft environment.

Lastly, this mishap highlights one of the key operational and physical differences in unmanned aircraft accident investigation identified by the International Society of Air Safety Investigators (ISASI): that the pilot/operator does not have direct feedback of the aircraft condition, trajectory, and surrounding airspace (2015).

Mishap Rates ("How Often")

Mishap rate comparisons are often made by considering the current cumulative mishap rate, or sometimes the mishap rate over a recent period. A comprehensive, analysis comparing the progression of the cumulative mishap rate during the maturity of the MQ-9 to manned aircraft was not found.

The U.S. Department of Defense has published several *Unmanned Aircraft Systems (UAS) Roadmaps* to provide a Defense-wide vision for UAVs and related technologies (U.S. Department of Defense, 2009). In the 2005, 2007, and 2009 Roadmaps, a basic comparison of several military unmanned aircraft to the F-16 and U-2 is provided. In addition to plotting the cumulative Class A mishap rates, the Roadmaps provide broad conclusions on platform maturity. An example from the 2009 Roadmap:

> Historically, UAS have suffered mishaps at one to two orders of magnitude greater than the rate (per 100,000 hours) incurred by manned military aircraft. In recent years, however, flight experience and improved technologies have enabled UAS to continue to track the reliability of early

manned military aircraft with their reliability approaching an equivalent level of reliability to their manned military counterparts. (p. 92)

Herz, in a 2008 study on Predator mishaps, makes a similar conclusion that while MQ-1 mishap rates might be high relative to more mature U.S. Air Force aircraft, they are similar to the rates seen in the early years of the F-16. But how similar? Herz's conclusion and the Roadmap discussions do not provide an analysis to answer this question. The next section of this paper provides such an analysis.

Methodology

The objective of the analysis portion of this study was to compare the progression of the cumulative Class A mishap rate of the U.S. Air Force MQ-9 to several U.S. Air Force manned aircraft during approximately the first one million flight hours. The approximate one million-hour range was selected because the U.S. Air Force MQ-9 had accumulated 1,378,992 flight hours as of FY17 (U.S. Air Force Safety Center, 2018). The standard metric of Class A mishaps per 100,000 flight hours was used.

A chi-square goodness-of-fit calculation was used to compare the cumulative mishap rates using designated points or benchmarks. Although each benchmark was a cumulative rate, each included new data in the rate calculation and was treated as a random sample for the chi-square calculations.

Comparison Aircraft

The choice of comparison aircraft for this type of analysis will vary depending on the goal of the researcher. For this study, the comparison aircraft were selected based on the data available on the U.S. Air Force Safety Center website and the following characteristics:

- The total flight hours. Aircraft with at least 250,000 flight hours, and up to approximately one million flight hours, were considered.
- The recency of the flight hours.
- A manned aircraft. Since the goal of the analysis was to compare the MQ-9 to U.S. Air Force manned aircraft, the RQ-4 Global Hawk was not considered for this analysis. In addition, the RQ-4 had only accumulated 206,974 hours as of FY17 (U.S. Air Force Safety Center, 2018).
- A fixed-wing aircraft. Helicopters and tilt rotor aircraft like the V-22 Osprey were not considered.
- An aircraft not primarily used for pilot training.

- Aircraft used in capacities other than air freight or cargo.
- Single-engine aircraft (like the MQ-9) were preferred, for similar propulsion system redundancy; however, because only a few single-engine aircraft were available, aircraft with more than one engine were also considered.
- Aircraft powered by turbine engine(s), for more similar powerplant reliability. This is why the MQ-9 was the focus of the comparison and not the MQ-1.
- The U.S. Air Force was the launch customer for the aircraft (or one of the launch customers). The A-7, for instance, was not selected. The U.S. Navy began flying the A-7 in 1965 and began flying them in Vietnam in 1967 (Schoeni, 1978, p. 4; Naval Aviation National Museum, 2016). The U.S. Air Force did not begin flying the A-7 until 1968 (National Museum of the Air Force, 2015). This will be evident in the A-7 data plotted later in this report.

Based on the criteria listed above, six comparison aircraft were selected: the F-16 Fighting Falcon, the F-15 Eagle, the A-10 Thunderbolt, the U-2 Dragon Lady, the B-1 Lancer, and the F-22 Raptor. As of FY17, the U-2, B-1, and F-22 had not yet achieved one million flight hours but had reached at least 250,000 hours. Like the MQ-9, the F-16 and U-2 are single-engine aircraft; the F-15, A-10, and F-22 have two engines; the B-1 has four.

Perhaps a comparison like this, between a remotely piloted aircraft (such as the MQ-1 or possibly the MQ-9) and an aircraft like the F-22 Raptor, do not adequately consider distinctions in mission; some instead recommend a comparison like the RQ-1 unarmed Predator to the MC-12 Liberty (Kreuzer, 2015). Although the MQ-9 and the comparison aircraft have somewhat different missions, the goal was to provide a more complete understanding of comparisons already being made.

Furthermore, while the MC-12's mission might be more aligned with the RQ-1's, as Kreuzer contends (reconnaissance), the U.S. military was not the launch customer for the MC-12. The MC-12 is a military version of the Beechcraft Super King Air 350ER (U.S. Air Force, 2017a). The Super King Air series was flown in civil aviation for several decades before entering service with the U.S. Air Force in 2009 (AOPA, 2018; U.S. Air Force 2017a). Therefore, the MC-12 does not fit the goal of this study—to compare the progression of the Class A mishap rate during initial maturity.

Included next is a general description of the steps used to complete the comparisons—as this methodology could be used to compare the cumulative mishap rates of other aircraft.

Step 1: Gather and Organize the Mishap Data

For each aircraft, obtain the data needed to calculate the cumulative mishap rate (usually mishaps per 100,000 flight hours), at some interval (usually annually), over the range to be compared.

Step 2: Determine the Comparison Points

Determine the comparison points or benchmarks in the following manner:

- The comparison points should begin at a point no lower than the frequency used to calculate the mishap rate. For example, if the often-applied metric of mishaps per 100,000 flight hours is used, the first comparison point should be no lower than 100,000 hours. Mishaps prior to this point will be reflected in the cumulative rate calculations.
- To generate the comparison points, determine an objective scale (meaning formula-based rather than arbitrary) that approximately fits the available data over the comparison range. For instance, if the aircraft with the fewest data points has 10, then develop a scale that generates approximately 10 points that generally align with the data.
- At least three comparison points are recommended.

Step 3: Determine the Cumulative Mishap Rates at the Comparison Points

Use linear interpolation or curve fitting to calculate the cumulative mishap rate for each aircraft at each comparison point.

Step 4: Conduct the Chi-Square Goodness-of-Fit Test Using the Comparison Points

A chi-square goodness-of-fit test involves a comparison between expected and observed values (often counts or frequencies). The expected and observed values usually sum to the same total—sometimes viewed as a requirement for a chi-square goodness-of-fit test. If they do not, then the expected values are usually adjusted (scaled as percentages of the total observed values) to make this so.

One reason for this adjustment is to ensure a common frame of reference for the counts or frequencies being compared (i.e., the same "per x"). For example, if the goodness-of-fit test is comparing categories of violent crimes in a particular city from one year to another, one year will be designated the "expected" values, the other the "observed." Because the total number of violent crimes across the categories is likely not the same in the two years compared, the expected counts or

frequencies are adjusted to have a common frame of reference with the observed values (the same "per x") (Weiss, & Weiss, 2017).

This type of adjustment would not work for the mishap rate comparison proposed here. However, this common frame of reference is accomplished by each comparison point being mishaps per 100,000 flight hours (or another shared "per x"). When the values (mishap rates) for the two aircraft being compared are totaled, though, they will most likely not be the same. Because of this, of the two aircraft being compared, the one with the lowest total will be treated as the expected values. Due to the nature of the chi-square calculation, this will yield a higher chi-square statistic (and lower *p*-value) than if the aircraft with the higher total were treated as the expected values.

While this might not be a typical use of a chi-square goodness-of-fit test, at its core, the goodness-of-fit test involves three essential elements:

- All frequencies or counts have the same frame of reference, as discussed.
- Each comparison in the overall calculation involves two frequencies or counts that are treated as if they should be equal—and the degree to which they are not is calculated using the chi-square formula.
- These *individual* chi-square calculations are summed, and it is the sum of these *individual* calculations that yields the chi-square "statistic"—or the *x*-axis value on the chi-square distribution to be checked against an *x*-axis "critical value," determined by the degrees of freedom in the comparison (or a *p*-value, representing the area under the distribution curve to the right of the chi-square statistic, is calculated and checked).

These three requirements are met in the steps proposed here.

Yet, because this is not a typical use of the chi-square goodness-of-fit test, an optional supplement to Step 4 is to expand upon the chi-square results with an additional test—for example, the Mann-Whitney test. The Mann-Whitney test is often used to compare means of non-parametric distributions (Leedy & Ormrod, 2016).

Step 5 (Optional): If Some Mishap Rates at the Comparison Points Are Less Than 5, Adjust the Rate and Repeat the Test

Some references on statistics contend that none of the expected counts or frequencies in a chi-square goodness-of-fit test can be less than 5. Some say that no more than 20% can be less than 5. Still others contend that these requirements are too restrictive (Weiss & Weiss, 2017).

Because of this consideration, if some of the mishap rates at the comparison points are less than 5, consider the following optional step:

(a) Alter the rate calculation (the "per x") to the next comparison point, until the rates are above 5. For instance, instead of mishaps *per 100,000 flights hours*, if the next comparison point is 150,000 flight hours, adjust the rate to *per 150,000 flight hours*. This would increase all mishap rates by a factor of 1.5.

(b) Remove initial comparison point(s) if needed, as required in Step 2. For example, if the rate is increased from *per 100,000* to *per 150,000 flight hours*, remove the *per 100,000 flight hours* comparison point. The mishaps prior to the aircraft achieving 150,000 flight hours would still be included in the comparison because the mishap rate calculations are cumulative values.

(c) Repeat the test with the updated data points.

Results

Using the five-step method described above, the MQ-9 cumulative mishap rate progression during approximately the first one million flight hours was tested against the comparison aircraft.

Step 1: Gather and Organize the Mishap Data

The historical mishap data for the MQ-9 and each comparison aircraft were obtained from the U.S. Air Force Safety Center's public website (U.S. Air Force Safety Center, 2018) and plotted by aircraft platform. Then, the cumulative mishap rate was calculated at the available data points (in this case, each year) as shown on Figure 2. The Appendix of this report lists all values of the plotted data. Note: Class A mishaps due to direct action of an enemy or hostile force are not included in the Air Force Safety Center data (U.S. Air Force, 2017b; R. Greenwood, personal communication, March 26, 2018).
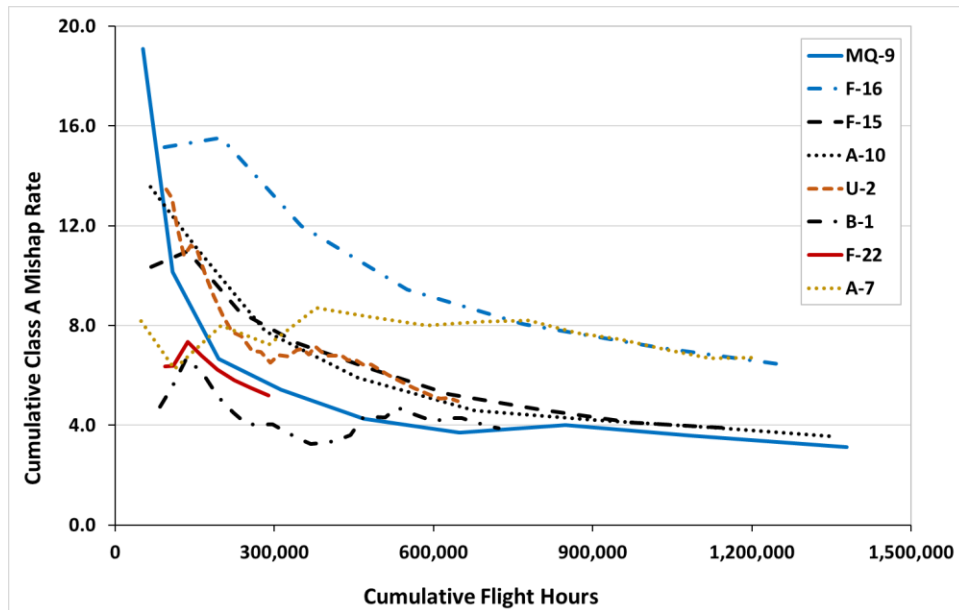
Figure 2. Cumulative mishap rates of the aircraft discussed in this study. The tables in the Appendix include the plotted data values. Note how the A-7 does not have the initial decrease in the cumulative rate seen in the other aircraft, likely at least in part due to the U.S. Air Force not being the launch customer.

Step 2: Determine the Comparison Points

As described in the methodology section, the comparison points for this analysis were determined using an objective, formula-based scale that approximately matched the number of data points for the MQ-9 and most comparison aircraft. Since the metric used was mishaps per 100,000 flight hours, the first comparison point was 100,000 hours. Then, six additional comparison points were determined by successively increasing the flight hours by 50%, resulting in a seventh and final comparison point of 1,139,063 hours. This seemed a good fit for the available data and produced three groups, based on the number of comparison points: Group A (all seven comparison points: the F-16, F-15, and A-10); Group B (five comparison points: the U-2 and the B-1); and Group C (three comparison points: the F-22).

Step 3: Determine the Cumulative Mishap Rates at the Comparison Points

Linear interpolation was used to determine the cumulative mishap rate for each aircraft at each comparison point. The tables in the Appendix show how the comparison points fit into the raw data.

Step 4: Conduct the Chi-Square Goodness-of-Fit Test Using the Comparison Points

   Group A. The comparison points for the Group A aircraft are shown in Figure 3. Noted details of the plot include:

- Each plot has a similar shape.
- The F-15 and A-10 plots have similar values.
- The MQ-9 has the lowest rate at all but the first comparison point.
- The F-16 plot is quite higher than the other three.

Figure 3. Comparison points for Group A (seven comparison points).

   The chi-square goodness-of-fit test found a significant difference between the mishap rate progression of the MQ-9 and the F-16 (p-value of 0.00). See Table 1.

   The MQ-9 data points, although generally lower than the F-15 and A-10, were found to be similar (p-values of 0.736 and 0.721, respectively).

   The Mann-Whitney tests (H_0: $\mu_1 = \mu_2$) using the same comparison points for Group A yielded results consistent with the chi-square tests. See Table 1.

Table 1

Chi-Square Goodness-of-Fit and Mann-Whitney Results for Comparing the Cumulative Mishap Rate Progression of the MQ-9 to the Group A Aircraft (Seven Comparison Points)

Hours	MQ-9	F-16	MQ-9	F-15	MQ-9	A-10
100,000	11.5	15.2	11.5	10.6	11.5	12.6
150,000	8.5	15.3	8.5	10.6	8.5	11.2
225,000	6.4	14.9	6.4	8.8	6.4	9.3
337,500	5.2	12.3	5.2	7.4	5.2	7.2
506,250	4.1	10.0	4.1	6.1	4.1	5.6
759,375	3.9	8.1	3.9	4.8	3.9	4.4
1,139,063	3.5	6.8	3.5	3.9	3.5	3.9
Total	43.1	82.7	43.1	52.2	43.1	54.2
Chi-square	\	/	\	/	\	/
Statistic	43.79		3.56		3.67	
p-value	0.000		0.736		0.721	
Null (H_0)	Reject		Accept		Accept	
Mann-Whitney						
p-value	0.011		0.405		0.337	
Null (H_0)	Reject		Accept		Accept	

Note: Degrees of freedom = 6; chi-square critical value = 12.59; alpha = 0.05. Aircraft with the lower total were used as the expected value for chi-square calculations. Mishap rates shown are rounded values; precise values were used in the statistical calculations. The *p*-values and null hypotheses are included to describe and quantify the comparisons.

Group B. The comparison points for the Group B aircraft are shown in Figure 4. As seen on the figure, the MQ-9 cumulative rate was generally higher than the B-1, but lower than the U-2.
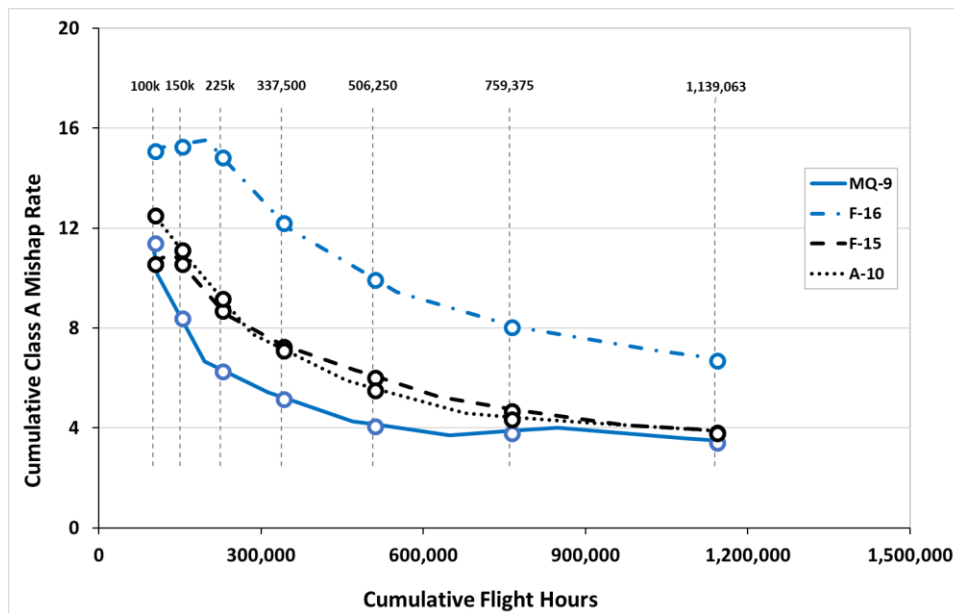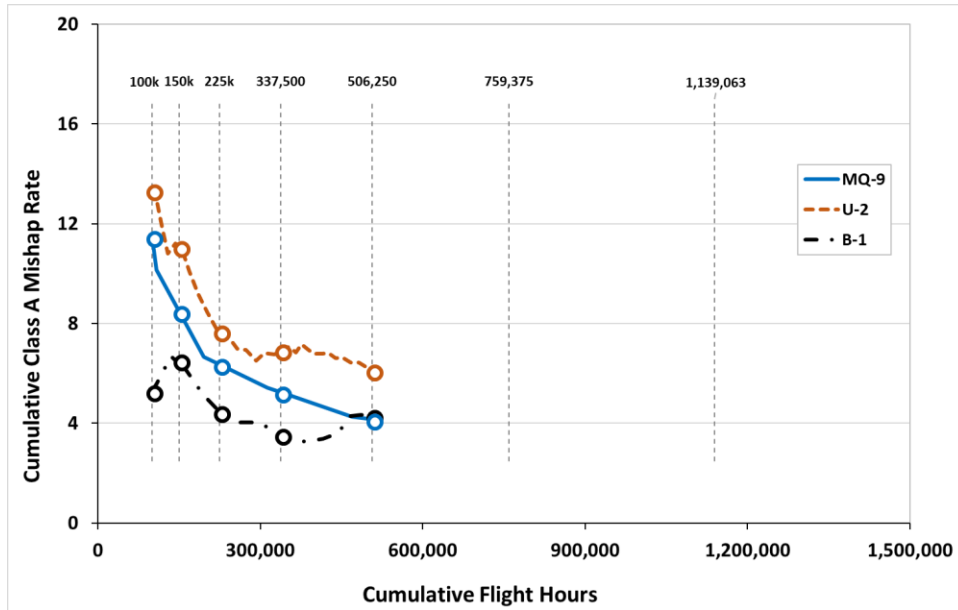
Figure 4. Comparison points for Group B (five comparison points).

The chi-square goodness-of-fit test did not find a significant difference between the MQ-9 and the U-2, or between the MQ-9 and the B-1. The MQ-9 to B-1 comparison yielded a chi-square statistic near, but less than the critical value (and a corresponding *p*-value just above than 0.05 limit). See Table 2.

To get an additional sense of the limits of this method—for what would be considered a significant difference—the U-2 and B-1 progressions were compared to each other using the same method. A significant difference was the result, as shown on Table 2.

As in Group A, the Mann-Whitney test results were consistent with those found by the chi-square goodness-of-fit test. See Table 2.

Table 2

Chi-Square Goodness-of-Fit and Mann-Whitney Results for Comparing the Cumulative Mishap Rate Progression of the Group B Aircraft (Five Comparison Points)

Hours	MQ-9	U-2	MQ-9	B-1	U-2	B-1
100,000	11.5	13.3	11.5	5.3	13.3	5.3
150,000	8.5	11.1	8.5	6.5	11.1	6.5
225,000	6.4	7.7	6.4	4.5	7.7	4.5
337,500	5.2	6.9	5.2	3.6	6.9	3.6
506,250	4.1	6.1	4.1	4.3	6.1	4.3
Total	35.7	45.2	35.7	24.2	45.2	24.2
Chi-square	\	/	\	/	\	/
Statistic	2.88		9.42		24.20	
p-value	0.579		0.051		0.000	
Null (H_0)	Accept		Accept		Reject	
Mann-Whitney						
p-value	0.421		0.310		0.016	
Null (H_0)	Accept		Accept		Reject	

Note: Degrees of freedom = 4; chi-square critical value = 9.49; alpha = 0.05. Aircraft with the lower total were used as the expected value for chi-square calculations. Mishap rates shown are rounded values; precise values were used in the statistical calculations. The *p*-values and null hypotheses are included to describe and quantify the comparisons.

Group C. The final comparison group was the MQ-9 and the F-22, with three comparison points. Figure 5 shows that the F-22 had a lower cumulative Class A mishap rate at each point.
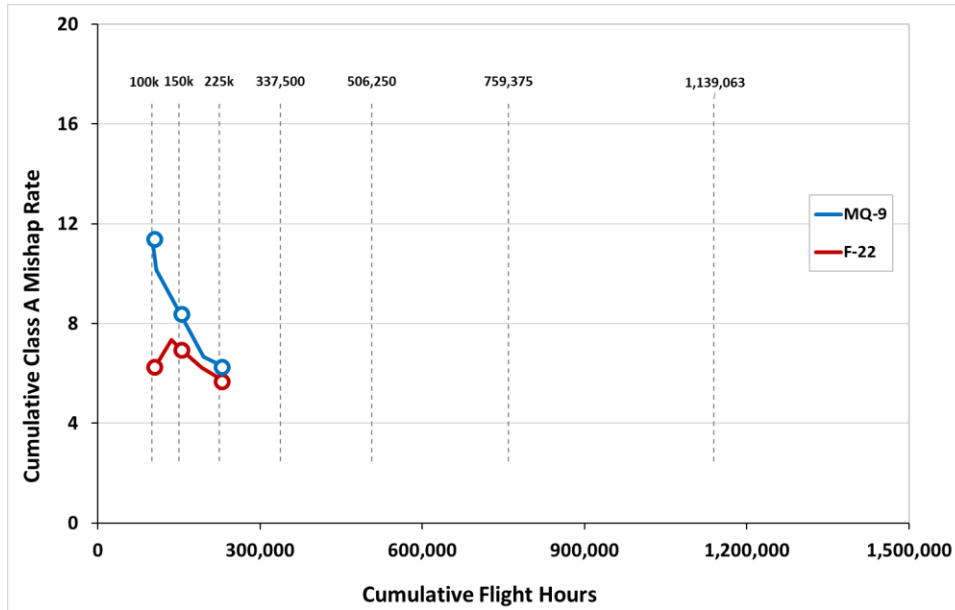
Figure 5. Comparison points for Group C (three comparison points).

Although the F-22 had a lower mishap rate at each point, the chi-square goodness-of-fit test did not find a significant difference in this comparison. This conclusion was once again supported by a Mann-Whitney test. See Table 3.

Table 3

Chi-Square Goodness-of-Fit and Mann-Whitney Results for the MQ-9 to the F-22 Mishap Rate Progression (Group C, Three Comparison Points)

Hours		MQ-9	F-22
100,000		11.5	6.4
150,000		8.5	7.0
225,000		6.4	5.8
	Total	26.3	19.2
Chi-square		\	/
Statistic		4.48	
p-value		0.107	
Null (*H*$_0$)		Accept	
Mann-Whitney			
p-value		0.268	
Null (*H*$_0$)		Accept	

Note: Degrees of freedom = 2; chi-square critical value = 5.99; alpha = 0.05. Aircraft with the lower total were used as the expected value for chi-square calculations. Mishap rates

shown are rounded values; precise values were used in the statistical calculations. The *p*-values and null hypotheses are included to describe and quantify the comparisons.

Step 5 (Optional): If Some Mishap Rates at the Comparison Points Are Less Than 5, Adjust the Rate and Repeat the Test

Because some of the comparison points in Groups A and B had values less than 5, the analysis was repeated for these groups using a per 150,000 flight hours mishap rate. This adjustment increased all rates above 5. This also resulted in the first comparison point being 150,000 flight hours, as described in the Methodology section. The comparison of the MQ-9 to the F-22 did not have rates below 5, so this optional step was not applicable for Group C.

When the chi-square goodness-of-fit test was repeated with the new Group A and Group B comparison points, the results were the same as those found using the standard per 100,000 flight hour rate. Table 4 shows the updated comparisons of the MQ-9 to the F-16, F-15, and A-10.

Table 4

Chi-Square Goodness-of-Fit and Mann-Whitney Results for Group A Aircraft Using Class A Mishaps per 150,000 Flight Hours

Hours	MQ-9	F-16	MQ-9	F-15	MQ-9	A-10
150,000	12.7	23.0	12.7	16.0	12.7	16.8
225,000	9.5	22.4	9.5	13.2	9.5	13.9
337,500	7.9	18.4	7.9	11.0	7.9	10.8
506,250	6.2	15.0	6.2	9.1	6.2	8.4
759,375	5.8	12.2	5.8	7.2	5.8	6.7
1,139,063	5.3	10.2	5.3	5.9	5.3	5.8
Total	47.4	101.2	47.4	62.3	47.4	62.4
Chi-square	\	/	\	/	\	/
Statistic	63.92		5.25		5.35	
p-value	0.000		0.386		0.375	
Null (*H*$_0$)	Reject		Accept		Accept	
Mann-Whitney						
p-value	0.009		0.240		0.262	
Null (*H*$_0$)	Reject		Accept		Accept	

Note: Degrees of freedom = 5; chi-square critical value = 11.07; alpha = 0.05. Aircraft with the lower total were used as the expected value for chi-square calculations. Mishap rates shown are rounded values; precise values were used in the statistical calculations. The *p*-values and null hypotheses are included to describe and quantify the comparisons.

Table 5 shows the comparisons with the MQ-9, U-2, and B-1. Note how the U-2 to B-1 comparison using the per 150,000 flight hour rate was the only comparison in this study that the Mann-Whitney results differed from the chi-square goodness-of-fit test. The chi-square test again found a significant difference between the U-2 and B-1. The Mann-Whitney test, however, produced a *p*-value of .057 (just above the .05 alpha), indicating there was not a significant difference between the U-2 and B-1 using a per 150,000 flight hour rate. See Table 5.

Table 5

Chi-Square Goodness-of-Fit and Mann-Whitney Results for Group B Aircraft, and the U-2 to the B-1, Using Class A Mishaps per 150,000 Flight Hours

Hours		MQ-9	U-2		MQ-9	B-1		U-2	B-1
150,000		12.7	16.6		12.7	9.8		16.6	9.8
225,000		9.5	11.5		9.5	6.7		11.5	6.7
337,500		7.9	10.4		7.9	5.3		10.4	5.3
506,250		6.2	9.2		6.2	6.5		9.2	6.5
Total		36.3	47.8		36.3	28.3		47.8	28.3
Chi-square		\	/		\	/		\	/
Statistic		3.87			3.30			14.26	
p-value		0.276			0.348			0.003	
Null (H_0)		Accept			Accept			Reject	
Mann-Whitney									
p-value		0.343			0.486			0.057	
Null (H_0)		Accept			Accept			Accept	

Note: Degrees of freedom = 3; chi-square critical value = 7.81; alpha = 0.05. Aircraft with the lower total were used as the expected value for chi-square calculations. Mishap rates shown are rounded values; precise values were used in the statistical calculations. The *p*-values and null hypotheses are included to describe and quantify the comparisons.

Summary, Mishap Rate Comparisons

A analysis comparing the U.S. Air Force MQ-9 cumulative Class A mishap rate progression to that of several U.S. Air Force manned aircraft during approximately the first one million flight hours was completed. To summarize, the MQ-9 was found to have:

- a lower rate than the F-16 (a significant difference);
- a lower rate than the F-15 and the A-10 (but not a significant difference);
- a lower rate than the U-2 but a higher rate than the B-1 (neither were significant differences); and
- a higher rate than the F-22 (but not a significant difference).

Conclusion

This research has shown that there is a well-established, significant percentage of manned aircraft accidents attributed to crew error. A similar trend appears to be the case in the unmanned aircraft data. The available data also support that, like manned aircraft, skill-based crew errors will continue to be a major contributor in unmanned aircraft accidents. Further research is needed in this area as some of the studies considered here indicated that, for some unmanned aircraft beyond those considered in this study (the MQ-1 and MQ-9) judgment/decision errors were more common than skill-based errors. Additionally, the continued progression toward more autonomous aircraft control, how the pilot/operator interfaces with the aircraft (HMI), and crew training are interrelated and will play a significant role in future unmanned aircraft accidents.

The chi-square goodness-of-fit analysis indicated that, during approximately the first one million flight hours, the MQ-9 did not have a significant different Class A mishap rate progression than the comparison aircraft. The F-16 was the only exception and had a higher rate. These findings were validated with Mann-Whitney tests.

One of the criticisms of these results is that the method compares the MQ-9 to old data, and that the overall U.S. Air Force mishap rate has decreased over time. This is a valid point for the F-16, F-15, and A-10 comparisons—each of these aircraft achieved one million flight hours in the mid-1980s. This criticism does not hold up for the F-22. The F-22 passed the 100,000-hour milestone in FY11; the MQ-9 did in FY10. Also, the fact that the U-2 and B-1 had not yet achieved one million flight hours as of FY17 also somewhat disputes the "old data" criticism—not all the data in the U-2 and B-1 comparisons are old. Looking at the last comparison point for these aircraft (506,250 hours), the MQ-9 achieved this benchmark in FY14; the B-1 passed the same mark in FY07; the U-2 did so in FY09.

Another criticism of these results is that the aircraft fly different types of missions. As mentioned, one objective of this study was to provide a better understanding of comparisons already being made. Further, the MQ-9 flies armed ISR missions and can carry Hellfire missiles, Paveway laser-guided bombs, and Joint Direct Attack Munitions. In some respects, MQ-9 missions are not that different from those flown by the comparison aircraft.

Despite these possible criticisms, the results show that an unmanned aircraft like the MQ-9 can be operated with a similar level of safety to that of a very sophisticated manned aircraft like the F-22. These results can be applied more broadly as unmanned aircraft development continues, indicating that the unmanned

safety record will someday match and may eventually surpass that of manned aircraft.

Recommendations

The first recommendation involves the need to continually update unmanned aircraft crew training. This recommendation stems from the finding that, like manned aircraft, the unmanned aircraft considered in this study usually had skill-based crewmember errors—and a primary way to mitigate skill-based crew errors is through proper training. Moreover, as demonstrated by one of the case studies analyzed in this report, aircrew training must be tailored to the unique challenges of operating an aircraft remotely. In addition, more autonomous control will likely be required to remain competitive in military and commercial aviation. Improvements to automation will also improve safety. But more autonomous control—and improvements to the HMI/control station design—will also require updated crew training.

The second general recommendation from this study involves looking objectively at the available safety data. A comprehensive approach is needed to understand mishap statistics. Comparing mishap rate progressions based on aircraft maturity is just one part of the overall safety picture. Nevertheless, it shows that the MQ-9, during the first one million flight hours of platform maturity, had a lower Class A mishap rate progression than the F-16, and a similar progression to the F-15 and A-10; and (with the hours accumulated to date) the U-2, B-1, and F-22. These comparisons are in some ways limited due to the age of some of the data and to variation in missions. And the approach described here does not consider variables like aircraft cost, certification, and crew qualifications. But this type of comparison should be part of the discussion. A more comprehensive discussion of the available safety data will allow the continued expansion of unmanned aircraft into aviation operations.

References

Aircraft Owners and Pilots Association. (2018). *Beechraft Super King Air 200*. Retrieved from https://www.aopa.org/go-fly/aircraft-and-ownership/ aircraft-fact-sheets/beechcraft-super-king-air-200

AOPA Air Safety Institute. (2017). *26th Joseph T. Nall report, general aviation accidents in 2014*. Retrieved from https://www.aopa.org/-/media/files/ aopa/home/training-and-safety/nall-report/26thnallreport2017.pdf

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775-779. doi:10.1016/0005-1098(83)90046-8

Bearup, B. (2015, July 31). Flashback Friday: 50th anniversary of autolanding. *Airways Magazine*. Retrieved from https://airwaysmag.com/airchive/ flashback-friday-50th-anniversary-of-autolanding/

Cohen, T. N., Wiegmann, D. A., & Shappell, S. A. (2015). Evaluating the reliability of the Human Factors Analysis and Classification System. *Aerospace Medicine and Human Performance*, *86*(8), 728-735. doi:10.3357/amhp.4218.2015

Cooper, G. R. (2015). Attribute-based time series cross-correlation measures. *Pure and Applied Geophysics*, *173*(5), 1683-1702. doi:10.1007/s00024-015-1186-7

Frey, C. B., & Osborne, M. A. (2013). *The future of employment: How susceptible are jobs to computerisation?* Retrieved from https://www.oxfordmartin.ox .ac.uk/downloads/academic/The_Future_of_Employment.pdf

Giese, S., Carr, D., & Chahl, J. (2013). Implications for unmanned systems research of military UAV mishap statistics. *2013 IEEE Intelligent Vehicles Symposium (IV)*. doi:10.1109/ivs.2013.6629628

Glover, N. (2018, May 24). Naval Safety Center launches HFACS app. *San Diego Navy/Marine Corps Dispatch* [San Diego], pp. 7, 9.

Herz, R. P. (2008). *Assessing the influence of human factors and experience on predator mishaps* (3299218) (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (3299218)

Human Factors Analysis and Classification System. (2014). *About the HFACS system*. Retrieved May 1, 2018, from https://www.hfacs.com/index.html

International Civil Aviation Organization. (2017). *Remotely piloted aircraft system (RPAs) concept of operations (CONOPS) for international IFR operations*. Retrieved from https://www.icao.int/safety/UA/Documents/RPAS% 20CONOPS%20-%20EN.pdf

International Society of Air Safety Investigators. (2015). *Unmanned aircraft system handbook and accident/incident investigation guidelines*. Retrieved from http://www.isasi.org/Documents/UAS%20Handbook%202%20revised %20033115.pdf

Kreuzer, M. P. (2015). Nightfall and the cloud: Examining the future of unmanned combat aerial vehicles and remotely piloted aircraft. *Air & Space Power Journal*, *29*(5), 57-73.

Larson, G. C. (2012, August). The first autolanding - two pilots won a trophy in 1937 for keeping their hands off the controls. *Air & Space Smithsonian Magazine*. Retrieved from https://www.airspacemag.com/history-of-flight/the-first-autolanding-3818066/

Leedy, P. D., & Ormrod, J. E. (2016). *Practical research: Planning and design* (11th ed.). Boston, MA: Pearson.

Lee, C. H. (2011). Embracing autonomy: The key to developing a new generation of remotely piloted aircraft for operations in contested air environments. *Air & Space Power Journal*, *25*(4), 76-88. Retrieved from http://www.airuniversity.af.mil/ASPJ/

Lim, Y., Bassien-Capsa, V., Ramasamy, S., Liu, J., & Sabatini, R. (2017). Commercial airline single-pilot operations: System design and pathways to certification. *IEEE Aerospace and Electronic Systems Magazine*, *32*(7), 4-21. doi:10.1109/maes.2017.160175

Losey, S. (2018, February 16). Air Force announces official retirement date for iconic MQ-1 Predator drone. *Air Force Times*. Retrieved from https://www.airforcetimes.com/news/your-air-force/2018/02/16/air-force-announces-official-retirement-date-for-iconic-mq-1-predator/

Myers, M. (2015, April 16). SECNAV: F-35C should be Navy's last manned strike jet. *Navy Times*. Retrieved from https://www.navytimes.com/news/your-navy/2015/04/16/secnav-f-35c-should-be-navy-s-last-manned-strike-jet/

National Museum of the Air Force. (2015, July 21). *LTV A-7D Cosair II*. Retrieved from http://www.nationalmuseum.af.mil/Visit/Museum-Exhibits/Fact Sheets/Display/Article/196072/ltv-a-7d-corsair-ii/

Naval Aviation National Museum. (2016). *A-7E Corsair II*. Retrieved from http://www.navalaviationmuseum.org/attractions/aircraft-exhibits/item/?item=a-7e_corsairII

Reason, J. (2000). Human error: Models and management. *BMJ*, *320*(7237), 768-770. doi:10.1136/bmj.320.7237.7682013/03/05/us-drone-strike_n_2813857.html

Schoeni, A. (1978, March). New role for old plane. *All Hands*, (738), 4-6.

Shappell, S., & Wiegmann, D. (2004). HFACS analysis of military and civilian aviation accidents: A North American comparison. In *Proceedings of the ISASI 35th Annual International Seminar* (pp. 135–140). Sterling, VA: International Society of Air Safety Investigators.

Straub, J. (2014). Unmanned aerial systems: Consideration of the use of force for law enforcement applications. *Technology in Society*, *39*, 100-109. doi:10.1016/j.techsoc.2013.12.004

Thompson, P. (2014, June 14). The same man invented autopilot and the mile high club 100 years ago. *Jalopnik*. Retrieved from https://jalopnik.com/lawrence-sperry-inventor-of-the-autopilot-and-the-m-1592623110

Tvaryanas, A. P., Thompson, W. T., & Constable, S. H. (2006). Human factors in remotely piloted aircraft operations: HFACS analysis of 221 mishaps over 10 years. *Aviation, Space, and Environmental Medicine*, 77(7), 724-732. Retrieved from https://www.asma.org/journal

Tvaryanas, A. P., & Thompson, W. T. (2008). Recurrent error pathways in HFACS data: Analysis of 95 mishaps with remotely piloted aircraft. *Aviation, Space, and Environmental Medicine*, 79(5), 525-532. doi:10.3357/asem.2002.2008

United States Air Force. (2013). *United States Air Force aircraft accident investigation board report, MQ-1B, T/N 05-3154*. Retrieved from https://efoia.milcloud.mil/App/AddAttachment.aspx?docid=5295&ispaldoc=F

United States Air Force. (2015a). *MQ-1B Predator fact sheet*. Retrieved from http://www.af.mil/About-Us/Fact-Sheets/Display/Article/104469/mq-1b-predator/

United States Air Force. (2015b). *MQ-9 Reaper fact sheet*. Retrieved from http://www.af.mil/About-Us/Fact-Sheets/Display/Article/104470/mq-9-reaper/

United States Air Force. (2016). *United States Air Force aircraft accident investigation board report, MQ-9A, T/N 10-4113*. Retrieved from http://www.airforcemag.com/AircraftAccidentReports/Documents/2017/060716_MQ9A_Creech.pdf

United States Air Force. (2017a). *MC-12 fact sheet*. Retrieved from http://www.af.mil/About-Us/Fact-Sheets/Display/Article/104497/mc-12/

United States Air Force. (2017b). *Safety investigations and reports* (AFI 91-204). Retrieved from http://www.e-publishing.af.mil

United States Air Force Safety Center. (2018). *Aviation statistics*. Retrieved April 7, 2018, from http://www.safety.af.mil/Divisions/Aviation-Safety-Division/Aviation-Statistics/

United States Department of Defense. (2005a). *Department of Defense Human Factors Analysis and Classification System (DOD HFACS) Version 7.0*. Retrieved from http://www.safety.marines.mil/Portals/92/Mishap%20Investigation%20and%20Reporting%20Toolbox/DoD%20HFACS%207.0%20for%20USMC.pdf

United States Department of Defense. (2005b). *FY2005-2030 Unmanned systems roadmap*. Retrieved from https://fas.org/irp/program/collect/uav_roadmap2005.pdf

United States Department of Defense. (2007). *FY2007-2032 Unmanned systems integrated roadmap*. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/

a475002.pdf

United States Department of Defense. (2009). *FY2009-2034 Unmanned systems integrated roadmap*. Retrieved from https://www.globalsecurity.org/intell/ library/reports/2009/dod-unmanned-systems-roadmap_2009-2034.pdf

United States Department of Defense. (2013). *FY2013-2038 Unmanned systems integrated roadmap* (14-S-0553). Retrieved from http://www.dtic.mil/dtic/ tr/fulltext/u2/a592015.pdf

United States Department of Transportation. (2013). *Unmanned aircraft system (UAS) service demand 2015 - 2035: Literature review & projections of future usage* (DOT-VNTSC-DoD-13-01). Retrieved from https://fas.org/ irp/program/collect/service.pdf

United States Department of Transportation. (2016a). *Airplane flying handbook* (FAA-H-8083-3B). Retrieved from https://www.faa.gov/ regulations_ policies/handbooks_manuals/aviation/airplane_handbook/media/airplane_ flying_handbook.pdf

United States Department of Transportation. (2016b). *Pilot's handbook of aeronautical knowledge* (FAA-H-8083-25B). Retrieved from https:// www.faa.gov/regulations_policies/handbooks_manuals/aviation/phak/ media/pilot_handbook.pdf

United States Department of Transportation. (2017). *Private pilot - airplane; airman certification standards* (FAA-S-ACS-6A). Retrieved from https://www.faa.gov/training_testing/testing/acs/media/private_airplane_a cs_6A.pdf

Urnes, J. M., & Hess, R. K. (1985). Integrated flight control systems development - The F/A-18A Automatic Carrier Landing System. *Journal of Guidance, Control, and Dynamics*, *8*(3), 289 - 295. doi:10.2514/3.19978

Weiss, N. A., & Weiss, C. A. (2017). *Introductory statistics* (10th ed.). Boston, MA: Pearson.

Wiegmann, D. A., & Shappell, S. A. (2003). *A human error approach to aviation accident analysis: The human factors analysis and classification system*. Aldershot, Great Britain: Ashgate.

Williams, K. W. (2004). *A summary of unmanned aircraft accident/incident data: Human factors implications* (DOT/FAA/AM-04/24). Washington, DC: U.S. Dept. of Transportation, Federal Aviation Administration, Office of Aerospace Medicine.

Williams, K. W. (2006). *Human factors implications of unmanned aircraft accidents: flight-control problems* (DOT/FAA/AM-06/8). Washington, DC: U.S. Dept. of Transportation, Federal Aviation Administration, Office of Aerospace Medicine.

Appendix

This Appendix lists the data from the U.S. Safety Center website depicted in the figures and shows the results of the calculations to obtain the cumulative Class A mishap rates. It also includes the interpolated values (in boldface) for the comparison points used in the analysis.

Table A1

MQ-9 Class A Mishap Data and Comparison Points

Year*	Class A Mishaps*	Cumulative Class A Mishaps**	Hours*	Cumulative Hours*	Cumulative Class A Mishap Rate**
FY09	4	10	25,391	52,394	19.1
				100,000	**11.5**
FY10	1	11	56,103	108,497	10.1
				150,000	**8.5**
FY11	2	13	86,526	195,023	6.7
				225,000	**6.4**
FY12	4	17	118,039	313,062	5.4
				337,500	**5.2**
FY13	3	20	155,802	468,864	4.3
				506,250	**4.1**
FY14	4	24	179,560	648,424	3.7
				759,375	**3.9**
FY15	10	34	199,967	848,391	4.0
FY16	5	39	238,674	1,087,065	3.6
				1,139,063	**3.5**
FY17	4	43	291,927	1,378,992	3.1

Note: Adapted http://www.safety.af.mil/Portals/71/documents/Aviation/Aircraft%20Statistics/Q-9.pdf. Retrieved April 7, 2018. Interpolated values for comparison points are boldface.
*Data from U.S. Air Force Safety Center website. **Calculated data.

Table A2

F-16 Class A Mishap Data and Comparison Points

Year*	Class A Mishaps*	Cumulative Class A Mishaps**	Hours*	Cumulative Hours*	Cumulative Class A Mishap Rate**
CY81	5	14	56,423	92,398	15.2
				100,000	**15.2**
				150,000	**15.3**
CY82	17	31	107,389	199,787	15.5
				225,000	**14.9**
				337,500	**12.3**
CY83	11	42	150,728	350,515	12.0
				506,250	**10.0**
CY84	10	52	199,761	550,276	9.4
				759,375	**8.1**
CY85	10	62	219,647	769,923	8.1
CY86	11	73	254,491	1,024,414	7.1
				1,139,063	**6.8**
FY87	8	81	233,560	1,257,974	6.4

Note: Adapted from
http://www.safety.af.mil/Portals/71/documents/Aviation/Aircraft%20Statistics/F-16.pdf.
Retrieved April 7, 2018. Interpolated values for comparison points are boldface.
*Data from U.S. Air Force Safety Center website. **Calculated data.

Table A3

F-15 Class A Mishap Data and Comparison Points

Year*	Class A Mishaps*	Cumulative Class A Mishaps**	Hours*	Cumulative Hours*	Cumulative Class A Mishap Rate**
CY77	6	7	42,369	67,674	10.3
				100,000	**10.6**
CY78	8	15	69,023	136,697	11.0
				150,000	**10.6**
				225,000	**8.8**
CY79	5	20	96,959	233,656	8.6
				337,500	**7.4**
CY80	5	25	109,309	342,965	7.3
CY81	5	30	132,291	475,256	6.3
				506,250	**6.1**
CY82	3	33	153,369	628,625	5.2
				759,375	**4.8**
CY83	4	37	169,438	798,063	4.6
CY84	3	40	175,515	973,578	4.1
				1,139,063	**3.9**
CY85	5	45	185,324	1,158,902	3.9

Note: Adapted from
http://www.safety.af.mil/Portals/71/documents/Aviation/Aircraft%20Statistics/F-15.pdf.
Retrieved December 2, 2017. Interpolated values for comparison points are boldface.
*Data from U.S. Air Force Safety Center website. **Calculated data.

Table A4

A-10 Class A Mishap Data and Comparison Points

Year*	Class A Mishaps*	Cumulative Class A Mishaps**	Hours*	Cumulative Hours*	Cumulative Class A Mishap Rate**
CY78	7	9	44,538	66,433	13.5
				100,000	**12.6**
				150,000	**11.2**
CY79	8	17	86,544	152,977	11.1
				225,000	**9.3**
CY80	5	22	130,159	283,136	7.8
				337,500	**7.2**
CY81	5	27	174,924	458,060	5.9
				506,250	**5.6**
CY82	4	31	219,349	677,409	4.6
				759,375	**4.4**
CY83	7	38	226,129	903,538	4.2
CY84	6	44	224,058	1,127,596	3.9
				1,139,063	**3.9**
CY85	4	48	224,133	1,351,729	3.6

Note: Adapted from
http://www.safety.af.mil/Portals/71/documents/Aviation/Aircraft%20Statistics/A-10.pdf.
Retrieved December 5, 2017. Interpolated values for comparison points are boldface.
*Data from U.S. Air Force Safety Center website. **Calculated data.

Table A5

U-2 Class A Mishap Data and Comparison Points

Year*	Class A Mishaps*	Cumulative Class A Mishaps**	Hours*	Cumulative Hours*	Cumulative Class A Mishap Rate**
FY80	2	13	10,080	96,574	13.5
				100,000	**13.3**
FY81	1	14	10,211	106,785	13.1
FY82	0	14	10,131	116,916	12.0
FY83	0	14	12,555	129,471	10.8
FY84	2	16	13,257	142,728	11.2
				150,000	**11.1**
FY85	1	17	11,788	154,516	11.0
FY86	0	17	13,954	168,470	10.1
FY87	0	17	16,786	185,256	9.2
FY88	0	17	16,730	201,986	8.4
FY89	0	17	17,620	219,606	7.7
				225,000	**7.7**
FY90	1	18	18,001	237,607	7.6
FY91	0	18	19,820	257,427	7.0
FY92	1	19	16,597	274,024	6.9
FY93	0	19	18,085	292,109	6.5
FY94	2	21	15,643	307,752	6.8
FY95	1	22	17,726	325,478	6.8
				337,500	**6.9**
FY96	2	24	16,518	341,996	7.0
FY97	1	25	11,601	353,597	7.1
FY98	0	25	11,431	365,028	6.8
FY99	2	27	11,436	376,464	7.2
FY00	0	27	10,965	387,429	7.0
FY01	0	27	10,285	397,714	6.8
FY02	1	28	14,581	412,295	6.8
FY03	1	29	13,325	425,620	6.8
FY04	0	29	13,294	438,914	6.6
FY05	1	30	13,183	452,097	6.6
FY06	0	30	14,511	466,608	6.4
FY07	1	31	15,600	482,208	6.4
FY08	0	31	15,469	497,677	6.2
				506,250	**6.1**
FY09	0	31	15,810	513,487	6.0
FY10	0	31	15,778	529,265	5.9
FY11	0	31	16,496	545,761	5.7
FY12	0	31	16,782	562,543	5.5
FY13	0	31	16,713	579,256	5.4
FY14	0	31	16,713	595,969	5.2
FY15	0	31	16,159	612,128	5.1
FY16	1	32	16,502	628,630	5.1

| FY17 | 0 | 32 | 16,440 | 645,070 | 5.0 |

Note: Adapted from
http://www.safety.af.mil/Portals/71/documents/Aviation/Aircraft%20Statistics/U-2.pdf.
Retrieved April 7, 2018. Interpolated values for comparison points are boldface.
*Data from U.S. Air Force Safety Center website. **Calculated data.

Table A6

B-1 Class A Mishap Data and Comparison Points

Year*	Class A Mishaps*	Cumulative Class A Mishaps**	Hours*	Cumulative Hours*	Cumulative Class A Mishap Rate**
FY90	1	4	26,705	84,279	4.7
				100,000	**5.3**
FY91	2	6	23,355	107,634	5.6
FY92	3	9	26,970	134,604	6.7
FY93	1	10	30,179	164,783	6.1
				150,000	**6.5**
FY94	0	10	29,382	194,165	5.2
FY95	0	10	27,778	221,943	4.5
				225,000	**4.5**
FY96	0	10	26,370	248,313	4.0
FY97	1	11	24,803	273,116	4.0
FY98	1	12	23,744	296,860	4.0
FY99	0	12	22,884	319,744	3.8
				337,500	**3.6**
FY00	0	12	24,703	344,447	3.5
FY01	0	12	24,627	369,074	3.3
FY02	1	13	26,130	395,204	3.3
FY03	1	14	20,993	416,197	3.4
FY04	2	16	27,773	443,970	3.6
FY05	4	20	21,482	465,452	4.3
FY06	1	21	19,632	485,084	4.3
				506,250	**4.3**
FY07	1	22	24,083	509,167	4.3
FY08	3	25	22,406	531,573	4.7
FY09	0	25	22,426	553,999	4.5
FY10	0	25	22,760	576,759	4.3
FY11	0	25	23,535	600,294	4.2
FY12	2	27	28,824	629,118	4.3
FY13	1	28	24,725	653,843	4.3
FY14	0	28	21,692	675,535	4.1
FY15	0	28	22,534	698,069	4.0
FY16	0	28	13,241	711,310	3.9
FY17	0	28	10,950	722,260	3.9

Note: Adapted from
http://www.safety.af.mil/Portals/71/documents/Aviation/Aircraft%20Statistics/B-1.pdf.
Retrieved April 11, 2018. Interpolated values for comparison points are boldface.
*Data from U.S. Air Force Safety Center website. **Calculated data.

Table A7

F-22 Class A Mishap Data and Comparison Points

Year*	Class A Mishaps*	Cumulative Class A Mishaps**	Hours*	Cumulative Hours*	Cumulative Class A Mishap Rate**
FY10	0	6	24,675	94,401	6.4
				100,000	**6.4**
FY11	1	7	15,289	109,690	6.4
FY12	3	10	26,506	136,196	7.3
				150,000	**7.0**
FY13	1	11	26,184	162,380	6.8
FY14	1	12	29,939	192,319	6.2
FY15	1	13	31,993	224,312	5.8
				225,000	**5.8**
FY16	1	14	30,889	255,201	5.5
FY17	1	15	33,834	289,035	5.2

Note: Adapted from
http://www.safety.af.mil/Portals/71/documents/Aviation/Aircraft%20Statistics/F-22.pdf.
Retrieved April 7, 2018. Interpolated values for comparison points are boldface.
*Data from U.S. Air Force Safety Center website. **Calculated data.

Table A8

A-7 Class A Mishap Data and Comparison Points

Year*	Class A Mishaps*	Cumulative Class A Mishaps**	Hours*	Cumulative Hours*	Cumulative Class A Mishap Rate**
CY71	3	4	37,094	48,896	8.2
				100,000	**6.6**
CY72	3	7	62,810	111,706	6.3
				150,000	**7.0**
CY73	9	16	88,297	200,003	8.0
				225,000	**7.8**
CY74	5	21	89,547	289,550	7.3
				337,500	**8.0**
CY75	12	33	89,495	379,045	8.7
CY76	7	40	99,284	478,329	8.4
				506,250	**8.3**
CY77	7	47	108,681	587,010	8.0
CY78	9	56	100,882	687,892	8.1
				759,375	**8.2**
CY79	8	64	92,410	780,302	8.2
CY80	3	67	91,478	871,780	7.7
CY81	4	71	80,848	952,628	7.5
CY82	2	73	84,315	1,036,943	7.0
CY83	2	75	83,947	1,120,890	6.7
				1,139,063	**6.7**
CY84	6	81	85,643	1,206,533	6.7

Note: Adapted http://www.safety.af.mil/Portals/71/documents/Aviation/Aircraft%20Statistics/
A-7.pdf. Retrieved April 7, 2018. Interpolated values for comparison points are boldface.
*Data from U.S. Air Force Safety Center website. **Calculated data.