Epistemological Chicken: What do we learn from aircraft 'bird-ingestion' tests?

John Downer
ESRC Research Officer at CARR
January 2009

Photo taken by Flickr user grego!/ Greg Lam Pak Ng

It seems increasingly likely that birds were to blame for the ditching of US Airways flight 1549 into the Hudson river shortly after its takeoff from LaGuardia. The dramatic water 'landing' and its cause echo a much less happy incident in 1960, when a similarly sized airliner, Eastern Airlines Flight 375, struck a flock of starlings as it left Boston's Logan International Airport. In both cases the birds damaged the aircraft's engines and caused it to ditch in the water. There the similarity ends, however. For unlike Flight 1549's controlled and casualty-free touchdown on the icy Hudson, Flight 375 yawed hard to the left at 300 feet and plunged headlong into the shallow green water of Winthrop Bay with the loss of sixty-two lives. Nobody would suggest that New York's escape could not have been Boston's tragedy; even aviation insiders are astonished by the fortuitous escape. Having brushed so closely with disaster, therefore, it is worth looking at the measures taken to protect the flying public from errant avians.

Birds are a well-recognized aviation problem. The Federal Aviation Administration (FAA) estimates that "bird-strikes" (as they are known) cost the US aviation industry $600 million in damage every year. The U.S. Air Force - whose lower-flying aircraft share more airspace with birds - reports that birds kill two aircrew members every three to five years and down two of their aircraft annually. Even the Space Shuttle has a recorded bird collision, from July 26, 2005, as Discovery left the launch pad.

Civil aircraft generally fly beyond the reach of birds, so most collisions happen near airports. Groundsmen employ elaborate and guileful ruses to keep birds away from runways -- ranging from plastic hawks and rubber snakes to distress calls played over loudspeakers -- but their effects are limited. Natural selection has yet to endow birds with an aversion to airports and most become inured to even quite sophisticated attempts at intimidation: scarecrows are claimed as nesting places and noise-generators double as popular perches.

The onus, therefore, is on the aircraft and especially its engines: they must be bird-resilient. This is no mean challenge. Protective grills over engine mouths are an engineering dead-end. Any grill strong and dense enough to withstand birds at high speeds would occlude airflow to the turbines and would pose its own risk of being smashed into the blades. Instead, engines are designed to be tough. Incredibly tough. "Hard as nails", as we used to say; although that hardly begins to do them justice. Modern engines would smash nails like matchsticks. Aircraft engine turbine blades cost tens of thousands of dollars each and represent the very forefront of materials science. Their metal elements are 'grown' as a single crystal: a delicate and esoteric art.

To demonstrate that new engine designs can safely 'take a bird' they must demonstrate their bird-gulping powers in a series of standardized 'type-certification' tests drafted and overseen by the FAA. On first inspection these tests look straightforward: they emulate a 'bird-strike' and measure the engine's ability to cope with it. Although we might forgive a casual observer for thinking them a prank devised by a bored, mischievous and weirdly resourceful teenager.

The procedure is simple. First the engineers firmly mount the enormous engine on an outdoor test stand. Then they gradually open the throttle, urging it to maximum climbing speed where it bellows like a wounded Kraken and blows like an uncorked hurricane. The giant fan-blades spin faster and faster until their tips are moving at close to the speed of sound and the engine's bowels are hot enough to melt steel. And then, into the mouth of this Brobdingnagian blender-furnace, this magnificent technological jewel, the engineers hold their breath, cross their fingers, and launch an unplucked four-pound chicken.

The 'chicken test', as it is widely known, (although any four-pound bird will do) is but one of a series of bird-ingestion tests any new engine must pass before the FAA will certify it. Along with the single four-pound bird (recently raised to eight-pounds for very large engines), the engine must also swallow a volley of eight one-and-half pound birds, fired in quick succession, and a further volley of sixteen smaller birds of three ounces each.

If the turbine disintegrates or catches fire when the chick hits the fan, if the 'pilot' cannot shut it down afterwards or if it releases blade fragments through the engine housing, then the engine fails the test and slinks-off back to the drawing-board. The tests with smaller birds are more demanding: engines also fail these tests if their output is reduced by more than twenty-five percent, or if they fail within five minutes.

These criteria look straightforward in principle but in practice the tests involve complex and ambiguous judgments. Engines that look as if they failed, may pass (if, for instance, two birds struck the same fan blade -- an occurrence deemed 'unrepresentative' by the FAA), and engines that look as if they passed may still fail (if, for instance, the birds did not strike the right point in the engine).

More important than uncertainties in the pass/fail criteria are ambiguities about what passing the tests (even unambiguously) actually implies. We might imagine that successful bird-strike tests indicate that an engine will safely ingest birds in practice, but experience warns otherwise: engines sometimes fail in when they should not. By nature, tests are artificial: they simulate real situations in a controlled way so that we can learn. To control is to simplify, however, and to simplify is to lessen. No simulation is perfect and lab tests never fully replicate the world they are designed to recreate. What we gain in intelligibility, therefore, we risk losing in applicability. A very simplified test would produce many excellent measurements that meant almost nothing.

Bird-strike tests can be read as delicate bargain between representativeness and practicality. Take, for example, the birds. In many respects, the FAA choose their birds with a careful eye for realism. They prefer freshly killed birds, for instance, because frozen ones might contain dense ice particles if incompletely thawed, or be dehydrated if over-thawed. This concern for authenticity only extends so far, however. They stipulate the masses of the birds, for example, but not the species. Chickens are cheap and available test subjects but rarely fall awry of jet engines outside the laboratory; in the tests they stand-in for other species of a similar mass, such as ducks and gulls. Yet different species with the same mass still have varying shapes, volumes and densities and this might be consequential. When an American Airlines flight swallowed a great-crested cormorant in 2004, for instance, American's spokesperson justified the extensive damage by explaining that "...a cormorant is chunkier, meatier and has more bones than a looser, watery bird". Their claim had some foundation. Engineers have found that small changes in bird volume or density can affect what they vividly refer to as its "slice mass", which, in turn, can significantly alter its digestibility.

Another question-mark hangs over the volleys of eight 'medium' and sixteen 'small' birds. They are supposed to represent flocks but critics debate the similarity. They debate whether a series of birds, one after another in neat succession, is very representative of the red, feathery maelstrom of an engine inhaling a large group of birds. They also question the numbers. Many small birds flock in very large numbers, so when an aircraft hits a flock it can ingest many more than the sixteen stipulated in the test, as did the MD-80 transport aircraft that left 430 dead starlings on a runway at Dallas, or the US Airways B-737 that left over 200 gulls on the tarmac at Daytona Beach. The tests do not test 'large' birds in volleys at all, although geese, swans and storks all flock in large numbers, especially around migration time: as the passengers who found themselves in the Hudson will testify.

The speed of the birds is also contentious. The term 'bird-strike' is misleading as the aircraft does most of the striking, but since test-sites do not lend themselves to moving engines the engineers accelerate the birds instead. Hence the cannon. It uses compressed-air to accelerate the birds to two-hundred knots: the approximate speed of plane during takeoff and landing. This speed is contentious, however, with critics arguing that the maximum allowed airspeed below 10,000 feet is 250 knots (and sometimes higher) and that this, therefore, should be the bird velocity. Compressed air cannons are certainly capable of this. The U.S. Air Force, whose aircraft travel faster at low altitudes, has a sixty-foot cannon that will fire a four-pound feathered bird, beak first, at over a thousand miles per hour. They call it the 'rooster booster'.

These doubts are illustrative but hardly comprehensive. Even a cursory look at the dialogue around these tests reveals a plenitude of further uncertainties. Critics debate whether test engines should be attached to automatic 'surge recovery' systems; whether birds should strike the engine housing; whether the bird size categories are truly representative; and whether manicured test-engines are fair proxies for engines in service.

Few of these questions are straightforward. In most cases the FAA actively defend the choices that others challenge. Regarding bird speeds, for instance, they claim the tests become less rather than more severe at higher speeds ('slice-mass' again), and that 200 knots is more likely to "...result in the highest blade stresses at the blade's critical location". The Airline Pilots Association, in turn, call this an unproven assumption and following the debate from here leads down a rabbit-hole of claims and counterclaims that invariably arrives at a place where both sides are ultimately standing on untested (and untestable) assumptions.

In the face of all this uncertainty, we should understand that aircraft routinely inhale birds without barely a hiccup, and that the safety of modern air travel is indisputable. A countable number of flights take-off every year and an almost identical number land safely without incident. We should also understand that no engine could be immune to flocks of large birds like canada geese, especially while retaining anything like the efficiency of modern engines. A pant-wetting splash in the Hudson once a decade is probably an acceptable trade for cheap and fuel-efficient air travel. We do not respond to car accidents by demanding everyone drive a tank.

Also worth understanding, however, is that bird-strike tests are far from unique in their ambiguity. The disconcerting fact is that they are perfectly typical: all technological tests have this quality, and everything we know about the functioning of any complex machine is tainted by it to some degree. Tests and experiments get us closer to understanding how a technology will perform in the world, but not all the way. There exists an unbridgeable 'epistemic gap' between the laboratory and the world it represents. Epistemologists (that is to say, philosophers of knowledge) have long known this, sometimes referring to it as 'the problem of relevance': an overly academic term for a such a material issue.

It should frame our understanding of the technological world in which we move, and counsel us against the idea that we can 'know' complex systems before we have protracted experience with them. Regulators may speak reassuringly of rigorous tests and objective analyses, but this is a brave face. The truth is that we do not know our machines as well as we think we do, and this, at minimum, is something we should know.