Data-driven safety feedback as part of debrief for General Aviation pilots

Nicoletta Fala

5 Apr 2019

Ph.D. Defense

https://www.wallpaperup.com/45352/Airplane_Plane_Engine_Propeller_black_white.html



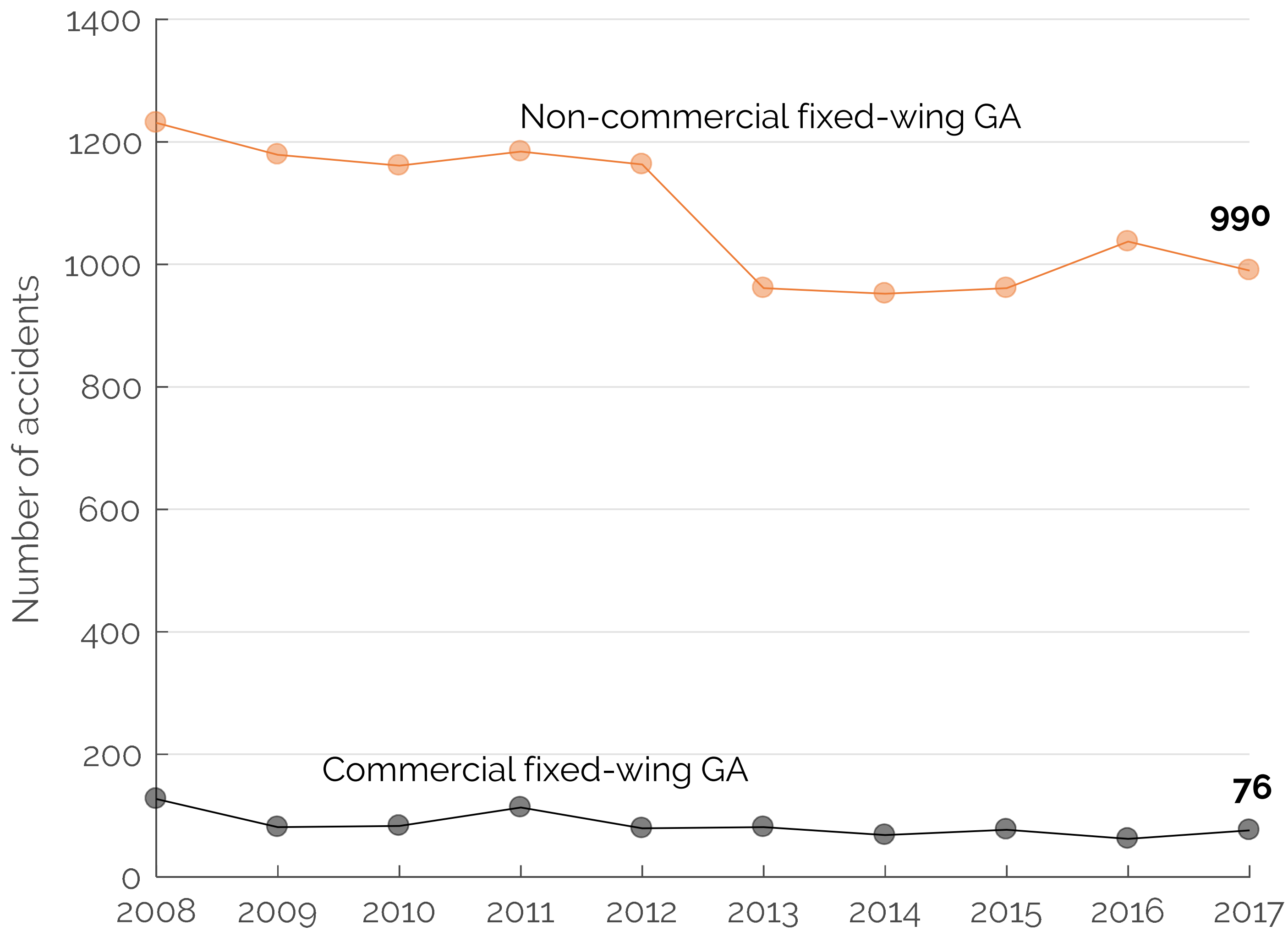


General Aviation aircraft make up 90% of the fleet

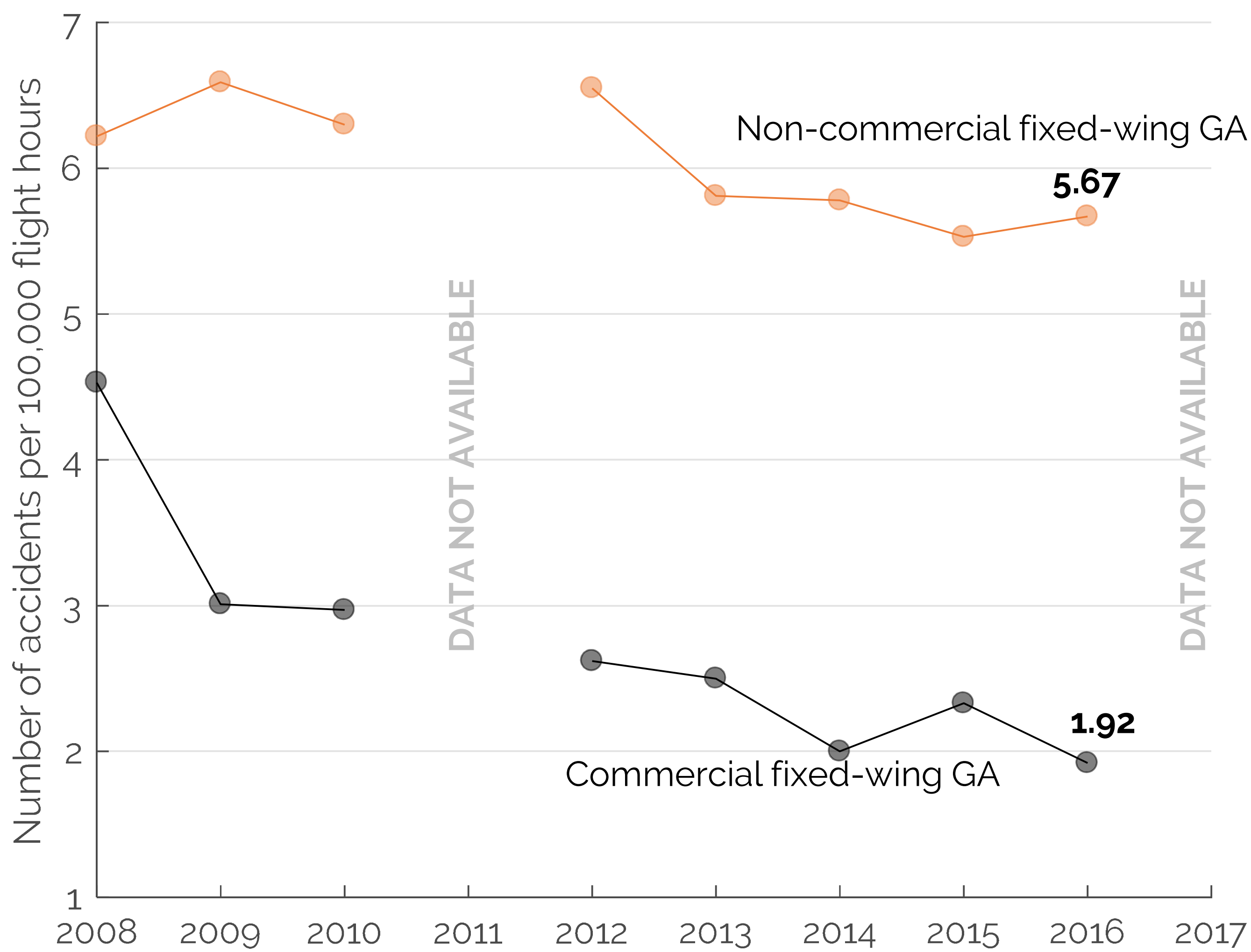
The foundation of most flying activities.

- Primary training ground
- 446,000 aircraft
- 24.8 million hours
- 162,455 FAA-licensed private pilots (airplane)

Several hundred pilots lose their lives in GA accidents each year



Several hundred pilots lose their lives in GA accidents each year



Most licensed GA pilots are flying without a flight instructor

✈ Typical flight lesson

- Lesson plan/discussion, pre-flight, flight, flight de-brief
- Practice emergency procedures, maneuvers

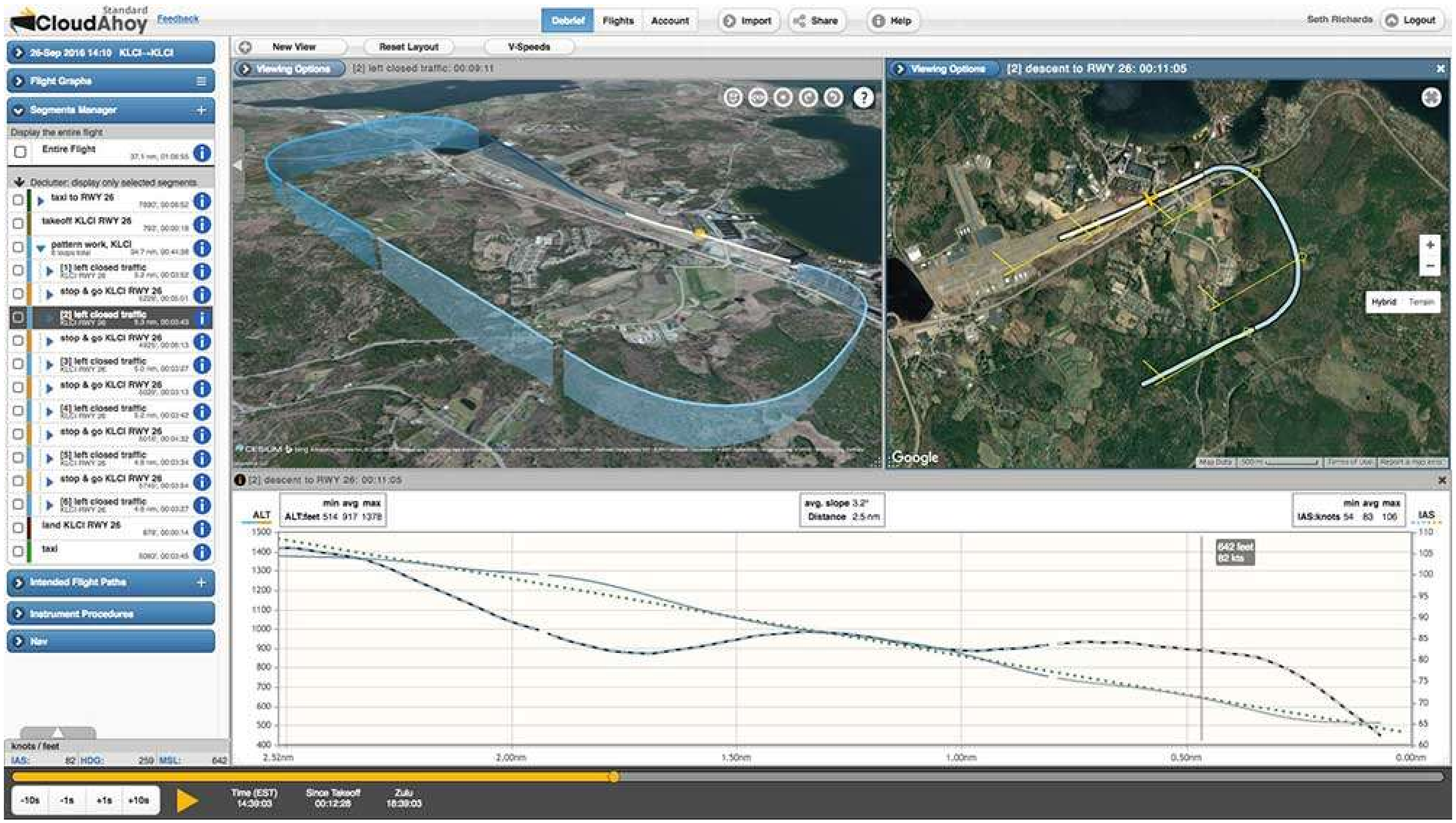
✈ Typical flight after license

- Pre-flight, flight
- Get to a destination, have fun, return
- Put the plane back in the hangar

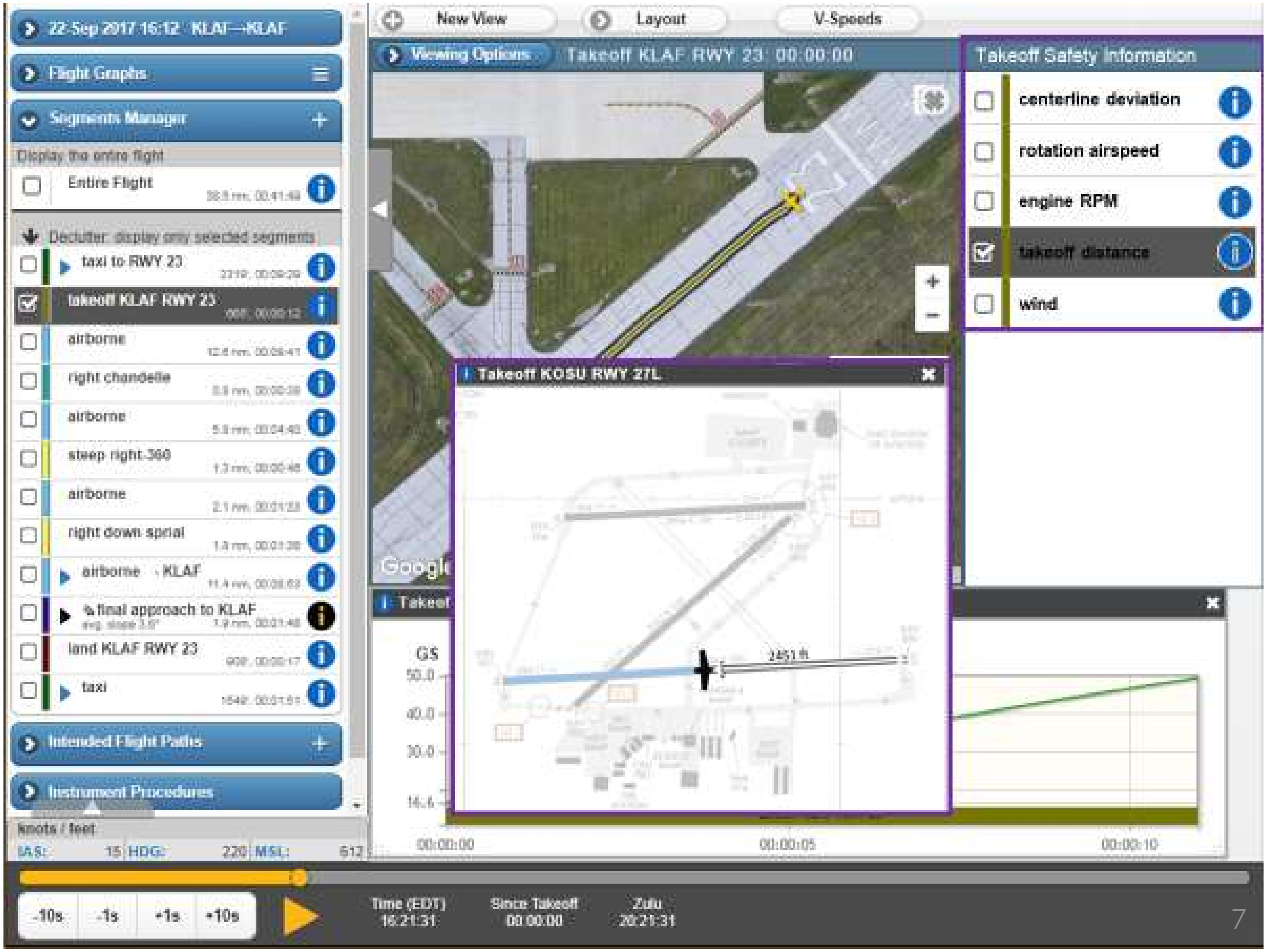
Can we use flight data to replicate the post-flight debrief?

Pilots already use commercially-available products to **visualize** their flights

CloudAhoy
debriefing for pilots



Can we provide pilots with effective risk information during their post-flight debrief?



Can we provide pilots with effective risk information during their post-flight debrief?

Accident Analysis

- What events/behaviors should we be trying to avoid and therefore look for in flight data?

Flight Data Analysis

- How do we calculate and detect these events?

Risk Communication

- How do we communicate such information to pilots so that they can improve?

Does changing how we present risk-related feedback affect its **effectiveness**?

✈ Effectiveness:

- Accuracy of risk perception
- Motivation to change unsafe behavior

Does **changing how** we present risk-related feedback affect its effectiveness?

- Framing Language
  - Risk-centric or safety-centric
- Representation method
  - Graphical or numerical
- Parameter type
  - Safety or performance

Thesis Outline

- *Flight data (G1000)*
- *Weather information*
- *Airport and runway databases*

Flight Data Analysis

Data Collection

Data Processing

Parameter Definitions

Parameter Calculations

- *Accident database*

Accident Analysis

Hazardous States/Triggers

Risk Communication

Debrief Screen Prototypes

Experiment Setup

Survey

What behaviors should we be trying to avoid and therefore look for in flight data?

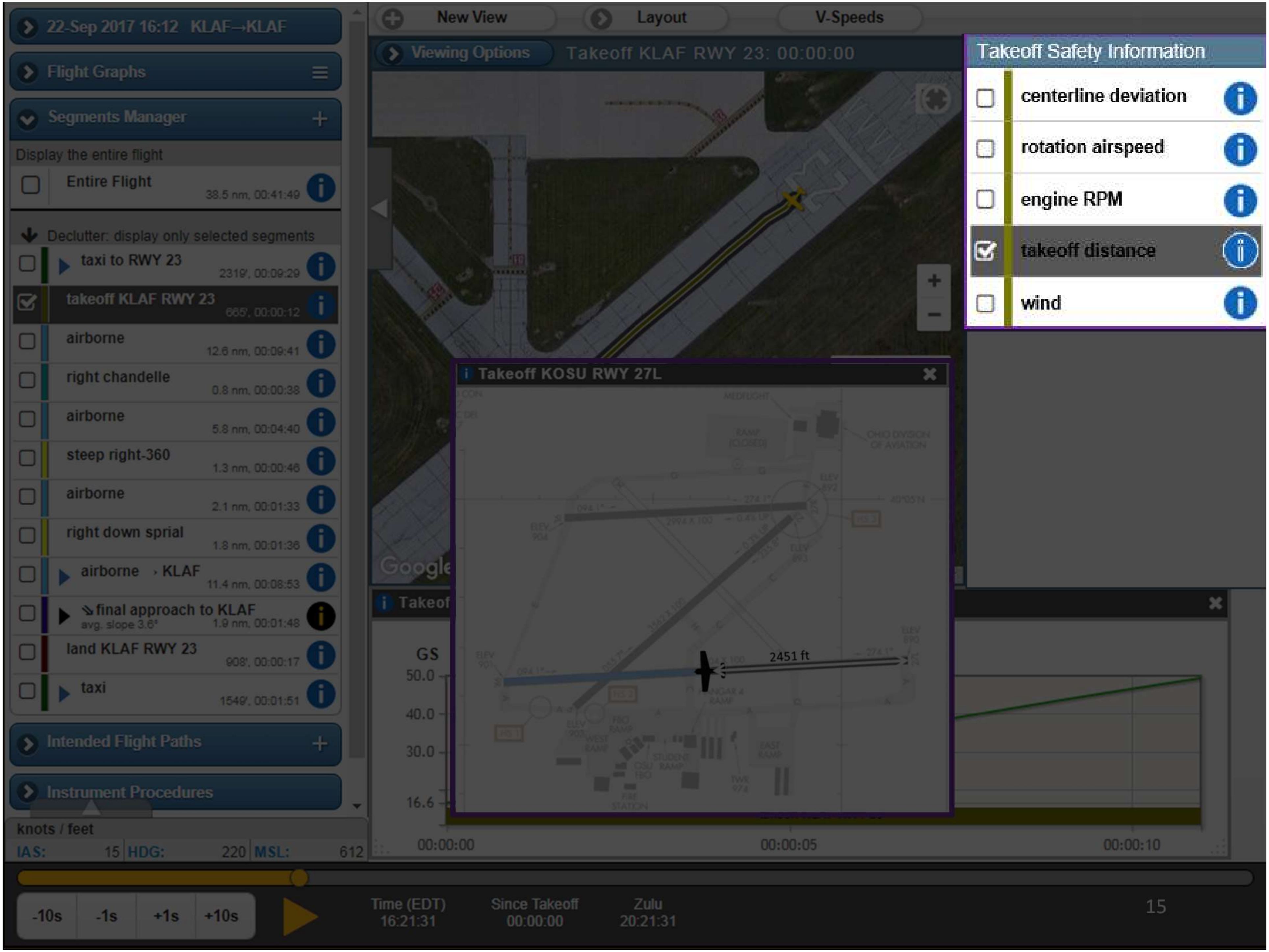
- *Flight data (G1000)*
- *Weather information*
- *Airport and runway databases*

Flight Data Analysis
Data Collection
Data Processing
Parameter Definitions
Parameter Calculations

- *Accident database*

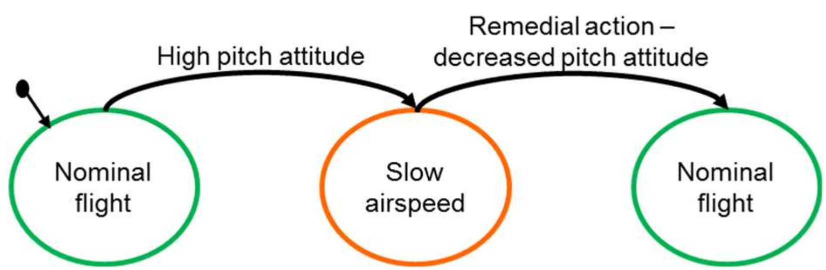
Accident Analysis
Hazardous States/Triggers

Risk Communication
Debrief Screen Prototypes
Experiment Setup
Survey

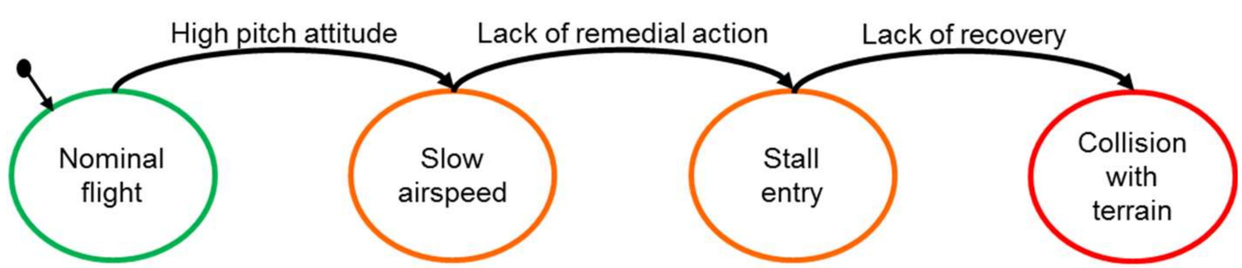

22-Sep 2017 16:12 KLAF→KLAF
Flight Graphs
Segments Manager
Display the entire flight
Entire Flight 38.5 nm, 00:41:49
Declutter: display only selected segments
taxi to RWY 23 2319', 00:09:29
takeoff KLAF RWY 23 665', 00:00:12
airborne 12.6 nm, 00:09:41
right chandelle 0.8 nm, 00:00:38
airborne 5.8 nm, 00:04:40
steep right-360 1.3 nm, 00:00:46
airborne 2.1 nm, 00:01:33
right down sprial 1.8 nm, 00:01:36
airborne › KLAF 11.4 nm, 00:08:53
final approach to KLAF avg. slope 3.6° 1.9 nm, 00:01:48
land KLAF RWY 23 908', 00:00:17
taxi 1549', 00:01:51
Intended Flight Paths
Instrument Procedures
knots / feet
IAS: 15 HDG: 220 MSL: 612
New View
Layout
V-Speeds
Viewing Options
Takeoff KLAF RWY 23: 00:00:00
Takeoff Safety Information
centerline deviation
rotation airspeed
engine RPM
takeoff distance
wind
Takeoff KOSU RWY 27L
2451 ft
GS
50.0
40.0
30.0
16.6
00:00:00
00:00:05
00:00:10
-10s
-1s
+1s
+10s
Time (EDT) 16:21:31
Since Takeoff 00:00:00
Zulu 20:21:31
15

Used a state-based model to define unsafe flight events

- **State** –period of time where the system (aircraft and pilot) exhibits a particular behavior
  - **Nominal State** – safe flight state
  - **Hazardous State** – unsafe flight state that may result in an accident
- **Trigger** – event that causes transition between two states



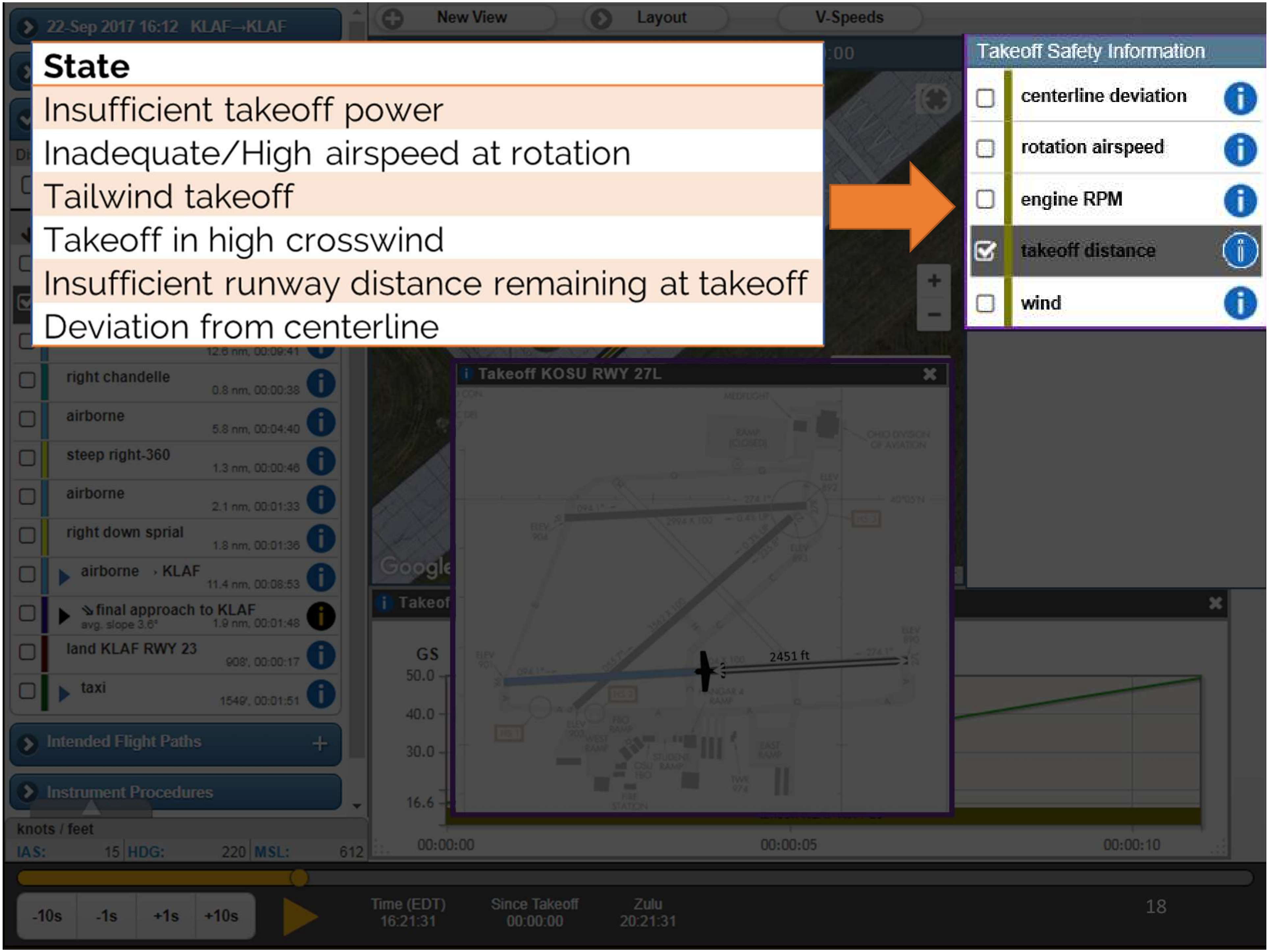
A flight that did not result in an accident was not necessarily safe



Generated a list of hazardous states from the NTSB accident database

Hazardous State or Trigger
Insufficient takeoff distance remaining
Insufficient takeoff power
Tailwind takeoff
Takeoff in high crosswind
Deviation from centerline
Inappropriate runway selection
Inadequate airspeed at rotation
High airspeed at rotation
Takeoff from inappropriately short runway

Task	***Task A. Normal Takeoff and Climb***
References	FAA-H-8083-2, FAA-H-8083-3, FAA-H-8083-23; POH/AFM
Objective	To determine that the applicant exhibits satisfactory knowledge, risk management, and skills associated with a normal takeoff, climb operations, and rejected takeoff procedures. ***Note:*** *If a crosswind condition does not exist, the applicant's knowledge of crosswind elements must be evaluated through oral testing.*
Knowledge	The applicant demonstrates understanding of:
PA.IV.A.K1	1. Takeoff distance.
PA.IV.A.K2	2. Takeoff power.
PA.IV.A.K3	3. Atmospheric conditions.
PA.IV.A.K4	4. Wind conditions and effects.
PA.IV.A.K5	5. The application of V_X or V_Y and variations with altitude.
PA.IV.A.K6	6. The manufacturer's recommended emergency procedures for relating to the takeoff sequence.
Risk Management	The applicant demonstrates the ability to identify, assess and mitigate risks, encompassing:
PA.IV.A.R1	1. Selection of runway based on wind, pilot capability, and aircraft limitations.
PA.IV.A.R2	2. The demonstrated crosswind component for the aircraft.
PA.IV.A.R3	3. Windshear.
PA.IV.A.R4	4. Tailwind.
PA.IV.A.R5	5. Wake turbulence.
PA.IV.A.R6	6. Go/no-go decision-making.
PA.IV.A.R7	7. Task management.
PA.IV.A.R8	8. Low altitude maneuvering.
PA.IV.A.R9	9. Wire strikes.
PA.IV.A.R10	10. Obstacles on the departure path.
PA.IV.A.R11	11. A rejected takeoff and predetermining takeoff abort criteria.
PA.IV.A.R12	12. Handling engine failure during takeoff and climb.
PA.IV.A.R13	13. Criticality of takeoff distance available.
PA.IV.A.R14	14. Plans for engine failure after takeoff.
PA.IV.A.R15	15. Sterile cockpit environment.
Skills	The applicant demonstrates the ability to:
PA.IV.A.S1	1. Verify ATC clearance and no aircraft is on final before crossing the hold line.
PA.IV.A.S2	2. Verify aircraft is on the assigned/correct runway.
PA.IV.A.S3	3. Ascertain wind direction with or without visible wind direction indicators.
PA.IV.A.S4	4. Determine if the crosswind component is beyond the pilot's ability or aircraft manufacturer maximum demonstrated value.
PA.IV.A.S5	5. Position the flight controls for the existing wind conditions.
PA.IV.A.S6	6. Clear the area; taxi into the takeoff position and align the airplane on the runway centerline/takeoff path.
PA.IV.A.S7	7. Confirm takeoff power; and proper engine and flight instrument indications prior to rotation:
PA.IV.A.S7a	a. Retracts the water rudders, as appropriate, confirm takeoff power and proper engine instrument indications prior to rotation, establishes and maintains the most efficient planning/lift-off attitude, and corrects for porpoising and skipping (ASES, AMES)
PA.IV.A.S8	8. Rotate and lift-off at the recommended airspeed and accelerate to V_Y (or other speed as appropriate for aircraft).
Task	***Task A. Normal Takeoff and Climb***
PA.IV.A.S9	9. Establish a pitch attitude that will maintain V_Y +10/-5 knots (or other airspeed as appropriate for aircraft).
PA.IV.A.S10	10. Retract the landing gear and flaps in accordance with manufacturer's guidance.
PA.IV.A.S11	11. Maintain takeoff power and V_Y +10/-5 knots or to a safe maneuvering altitude.
PA.IV.A.S12	12. Maintain directional control and proper wind drift correction throughout the takeoff and climb.
PA.IV.A.S13	13. Comply with responsible environmental practices, including noise abatement and published departure procedures.
PA.IV.A.S14	14. Complete the appropriate checklist.
PA.IV.A.S15	15. Comply with manufacturer's recommended emergency procedures related to the takeoff sequence.


22-Sep 2017 16:12 KLAF→KLAF
New View
Layout
V-Speeds
State
Insufficient takeoff power
Inadequate/High airspeed at rotation
Tailwind takeoff
Takeoff in high crosswind
Insufficient runway distance remaining at takeoff
Deviation from centerline
Takeoff Safety Information
centerline deviation
rotation airspeed
engine RPM
takeoff distance
wind
right chandelle
airborne
steep right-360
airborne
right down sprial
airborne › KLAF
final approach to KLAF
land KLAF RWY 23
taxi
Intended Flight Paths
Instrument Procedures
knots / feet
IAS: 15
HDG: 220
MSL: 612
Takeoff KOSU RWY 27L
2451 ft
GS
50.0
40.0
30.0
16.6
00:00:00
00:00:05
00:00:10
-10s
-1s
+1s
+10s
Time (EDT) 16:21:31
Since Takeoff 00:00:00
Zulu 20:21:31
18

How do we calculate and detect these behaviors?

- *Flight data (G1000)*
- *Weather information*
- *Airport and runway databases*

Flight Data Analysis

Data Collection

Data Processing

Parameter Definitions

Parameter Calculations

Risk Communication

Debrief Screen Prototypes

Experiment Setup

Survey

- *Accident database*

Accident Analysis

Hazardous States/Triggers

Flight data can come in various forms

✈ Flight data recorders (FDR)

- Location
- AHRS
- Engine
- Comm/Nav

✈ ADS-B devices

- Location

✈ Smartphone/Tablet

- Location
- AHRS

We can process flight data to make it more complete and uniform across the board

Different formats

- G1000 vs Avidyne
- ADS-B, Smartphone

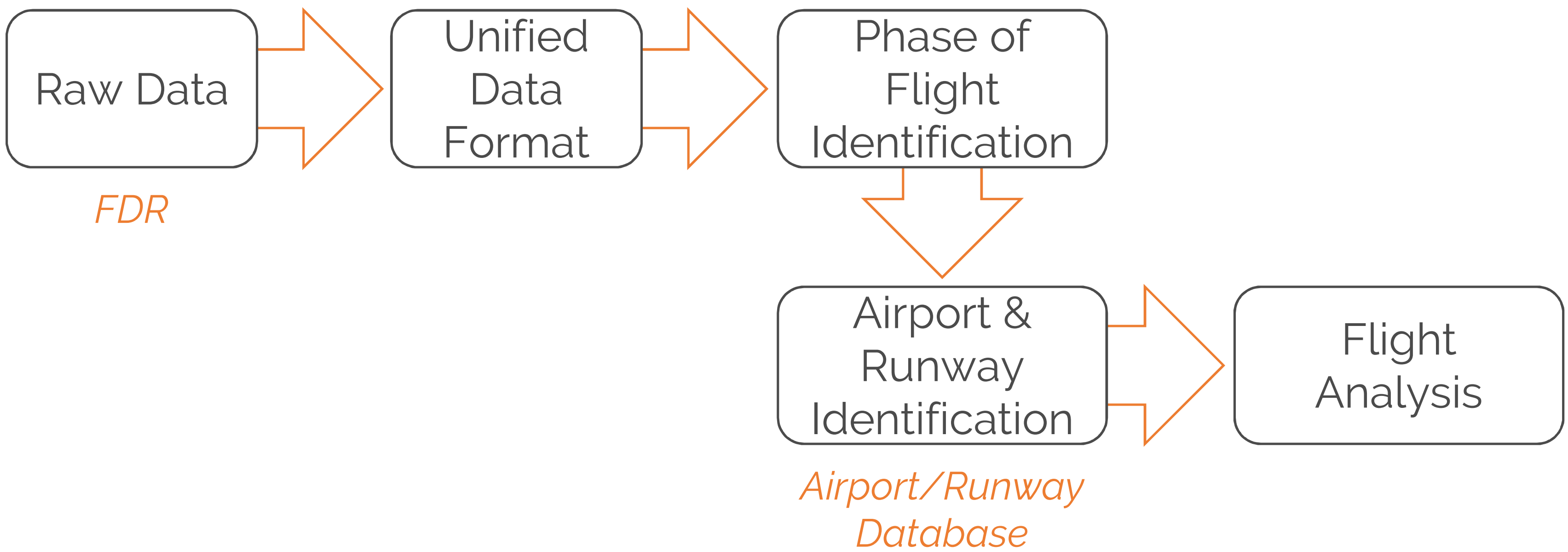
Missing information

- Airport in vicinity
- Departure/Arrival Runway
- Weather at the surface

Missing data points

- Recording frequency

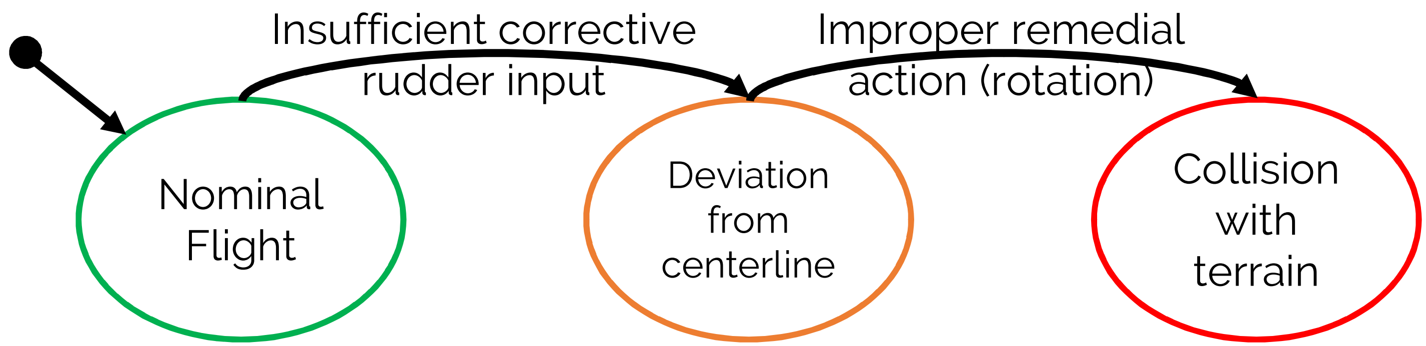
A series of automated algorithms processes the raw flight data



Created algorithms to detect each state in the post-processed flight data

State
Insufficient takeoff power
Inadequate/High airspeed at rotation
Takeoff in tailwind/high crosswind
Insufficient runway distance remaining at takeoff
Deviation from centerline

Deviation from the centerline is usually the result of insufficient rudder control

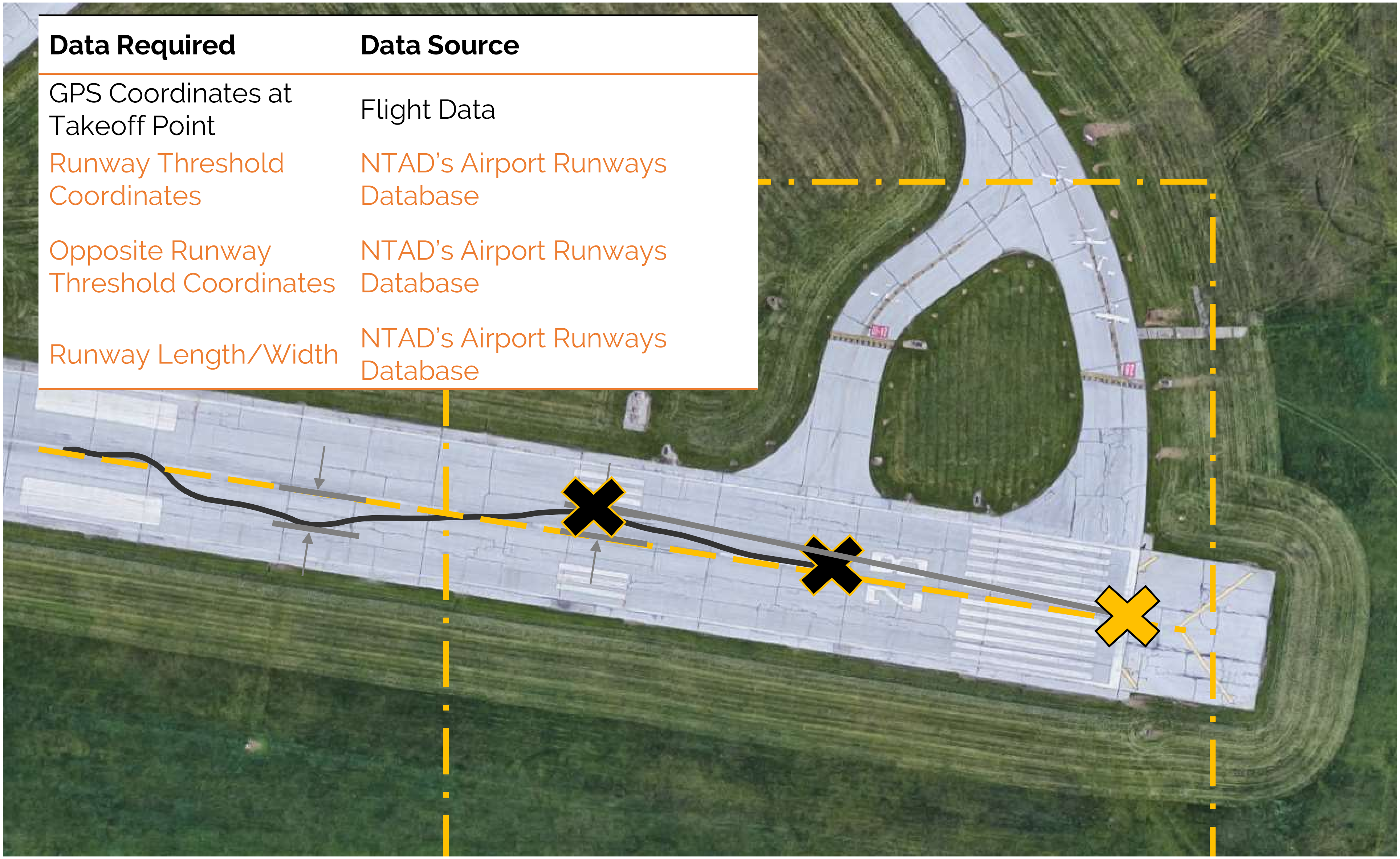


✈ Accident cause: “the pilot’s loss of directional control during takeoff, resulting in a decision to rotate early, and a collision with a hangar and subsequent fire.”

GAA16CA284

Calculate deviation from the centerline from the flight data

Data Required	Data Source
GPS Coordinates at Takeoff Point	Flight Data
Runway Threshold Coordinates	NTAD's Airport Runways Database
Opposite Runway Threshold Coordinates	NTAD's Airport Runways Database
Runway Length/Width	NTAD's Airport Runways Database



How do we communicate such information to pilots so that they can improve?

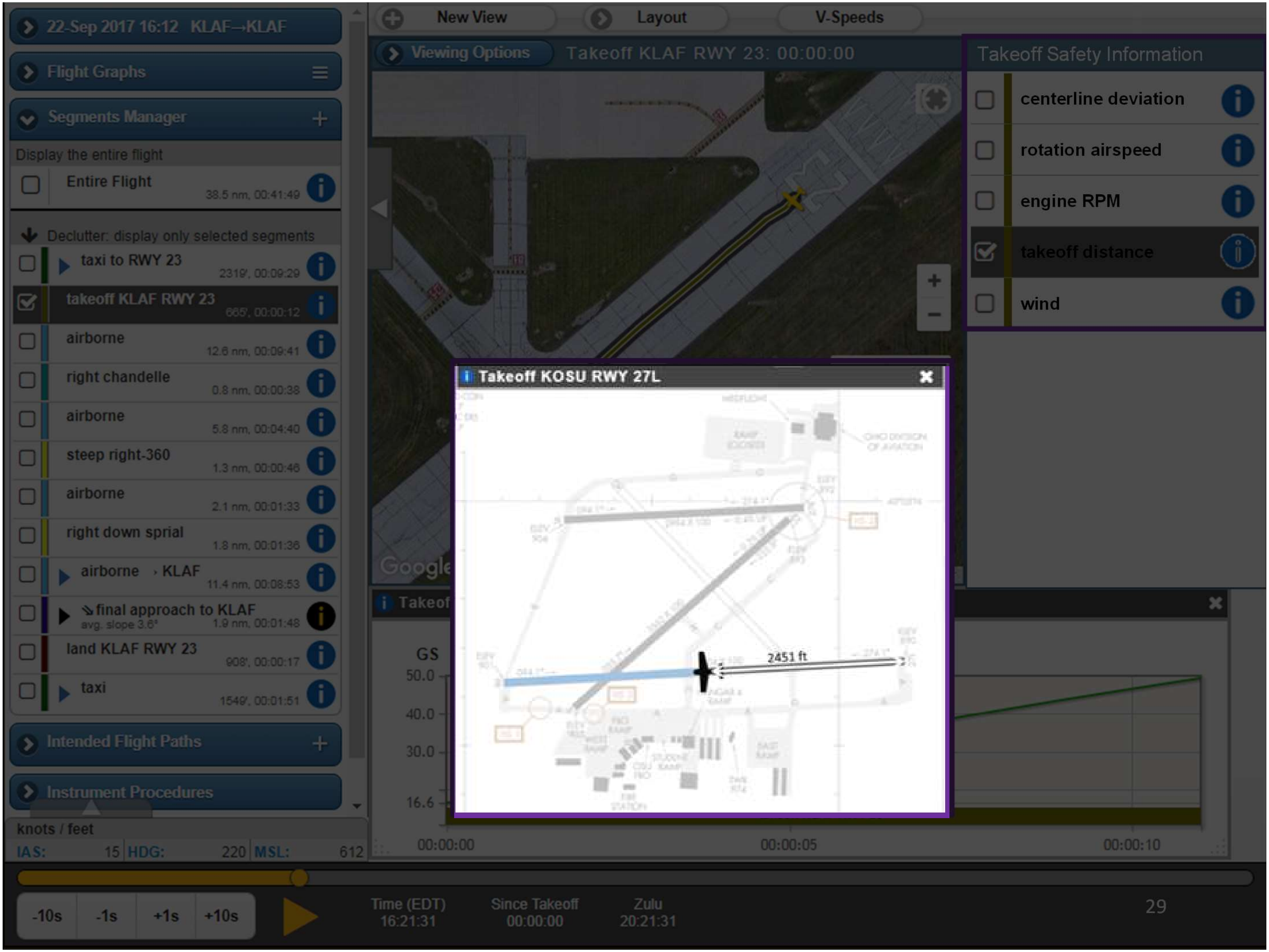
- *Flight data (G1000)*
- *Weather information*
- *Airport and runway databases*

Flight Data Analysis
Data Collection
Data Processing
Parameter Definitions
Parameter Calculations

- *Accident database*

Accident Analysis
Hazardous States/Triggers

Risk Communication
Debrief Screen Prototypes
Experiment Setup
Survey


22-Sep 2017 16:12 KLAF→KLAF
Flight Graphs
Segments Manager
Display the entire flight
Entire Flight
38.5 nm, 00:41:49
Declutter: display only selected segments
taxi to RWY 23
2319', 00:09:29
takeoff KLAF RWY 23
665', 00:00:12
airborne
12.6 nm, 00:09:41
right chandelle
0.8 nm, 00:00:38
airborne
5.8 nm, 00:04:40
steep right-360
1.3 nm, 00:00:46
airborne
2.1 nm, 00:01:33
right down sprial
1.8 nm, 00:01:36
airborne › KLAF
11.4 nm, 00:08:53
final approach to KLAF
avg. slope 3.6°
1.9 nm, 00:01:48
land KLAF RWY 23
908', 00:00:17
taxi
1549', 00:01:51
Intended Flight Paths
Instrument Procedures
knots / feet
IAS: 15 HDG: 220 MSL: 612
New View
Layout
V-Speeds
Viewing Options
Takeoff KLAF RWY 23: 00:00:00
Takeoff Safety Information
centerline deviation
rotation airspeed
engine RPM
takeoff distance
wind
Takeoff KOSU RWY 27L
2451 ft
GS
50.0
40.0
30.0
16.6
00:00:00
00:00:05
00:00:10
-10s
-1s
+1s
+10s
Time (EDT)
16:21:31
Since Takeoff
00:00:00
Zulu
20:21:31

Does changing how we present risk-related feedback affect its effectiveness?

✈ List factors to investigate

Does changing how we present risk-related feedback affect its effectiveness?

- Research on cognitive biases
  - Tversky & Kahneman, 1974
- Research on risk communication
  - Medicine
  - Education
  - Sports coaching
- We don't know how to communicate risk to pilots
  - Different population

We can present parameters in different ways

Does **changing how** we present risk-related feedback affect its effectiveness?

- Framing Language
  - Risk-centric or safety-centric
- Representation method
  - Graphical or numerical
- Parameter type
  - Safety or performance

10

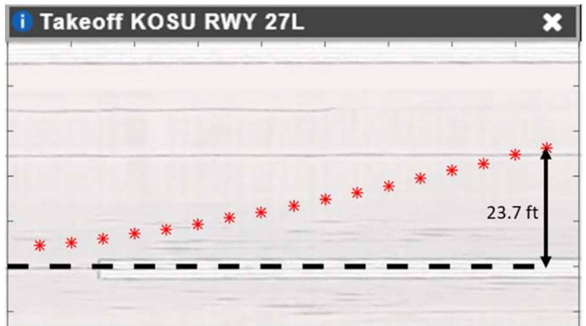
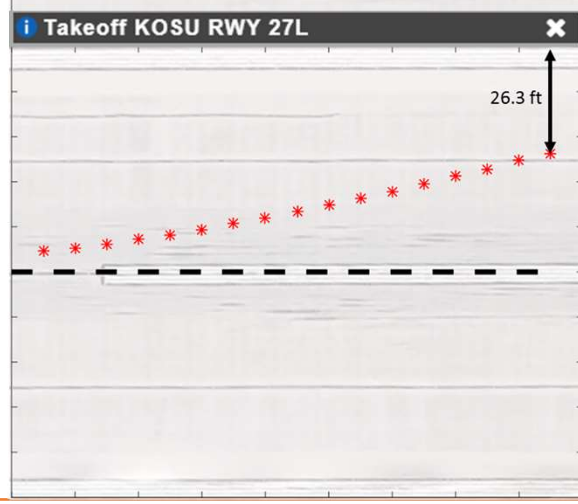
Does changing how we present risk-related feedback affect its effectiveness?

- List factors to investigate
- Design ways to present feedback

We can present parameters in different ways using a 2^3 full-factorial design

Treatment Group	Framing Language	Representation Method	Parameter Type
1	safety-centric	graphical	performance
2	risk-centric	graphical	performance
3	safety-centric	numerical	performance
4	risk-centric	numerical	performance
5	safety-centric	graphical	safety
6	risk-centric	graphical	safety
7	safety-centric	numerical	safety
8	risk-centric	numerical	safety

Tested the risk representation method and parameter type factors through the debrief

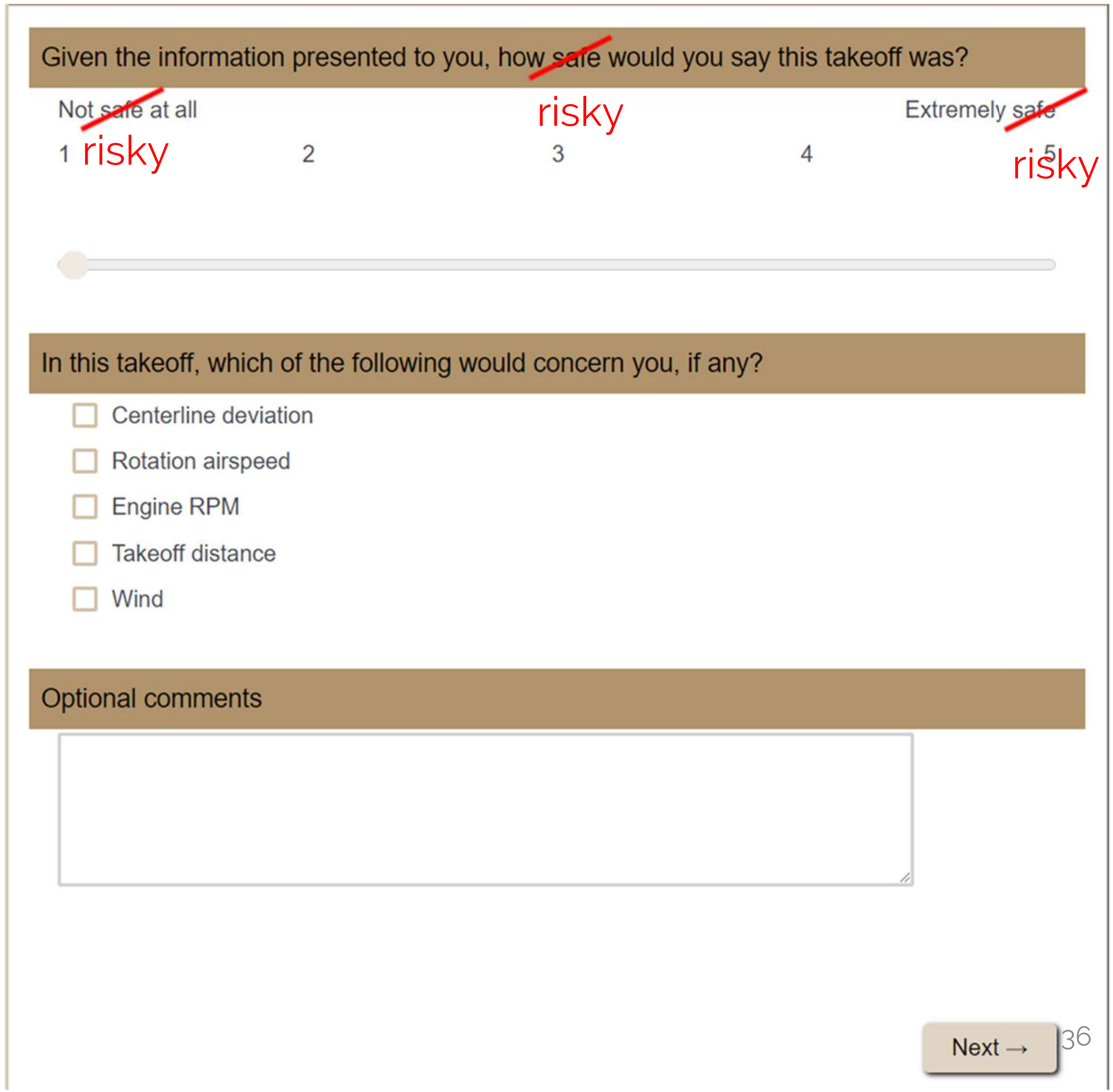
Parameter Type	Representation Method: Graphical	Representation Method: Numerical
Safety		Takeoff KOSU RWY 27L Distance from runway edge 26.3 ft Total runway width 100 ft
Performance		Takeoff KOSU RWY 27L Deviation from centerline 23.7 ft Total runway width 100 ft

Tested the framing language factors through the questions

Framing language

Safety-centric

Risk-centric



Does changing how we present risk-related feedback affect its effectiveness?

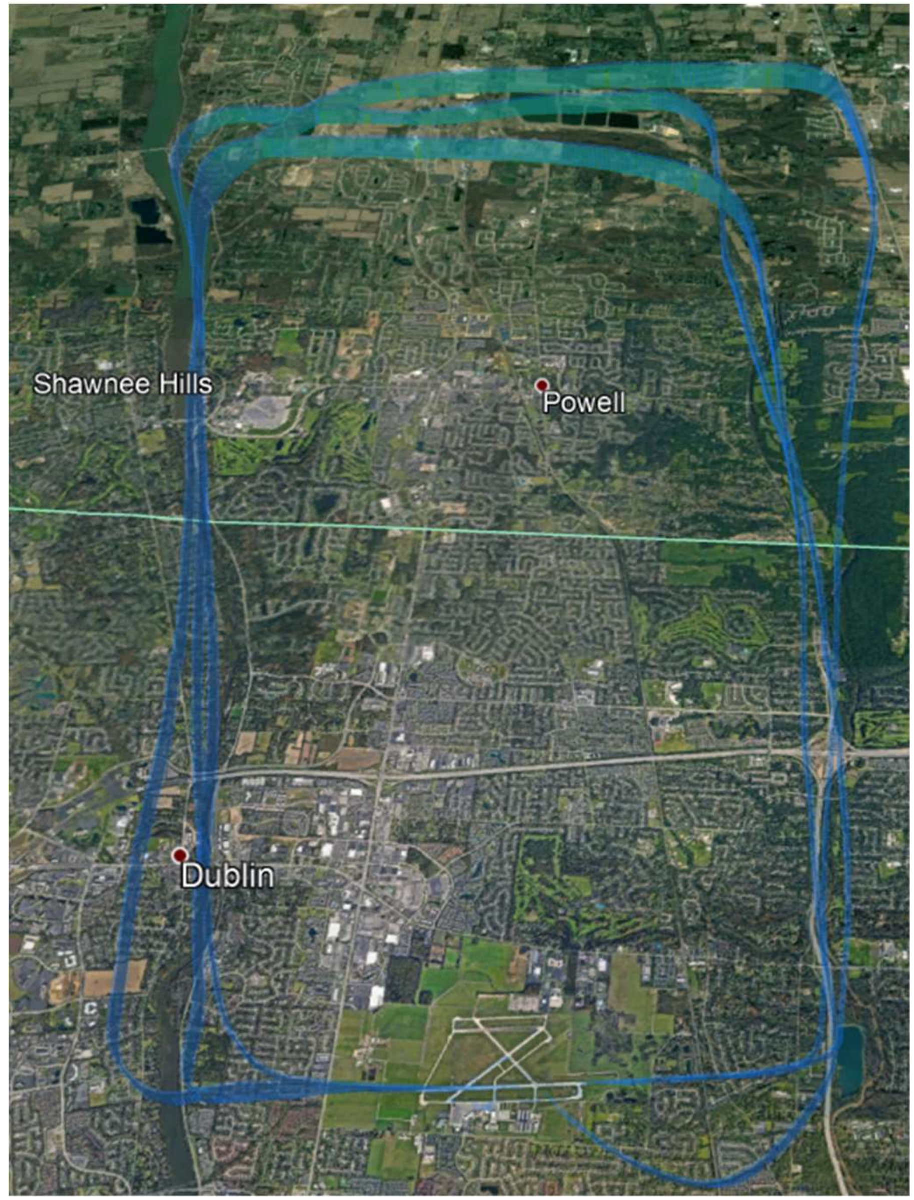
- List factors to investigate
- Design ways to present feedback
- Apply to sample flights

Used data from three flights to create debrief prototypes

- Flight A, Flight B, Flight C
- C172 at KOSU
- Different risk in each takeoff
  - B (safest), C, A (riskiest)

Hazardous State	Flight A	Flight B	Flight C
Centerline deviation	X		X
Rotation airspeed	X	X	X
Engine RPM		X	
Takeoff distance	X		X
Wind	X		

All three flights were very basic, with one takeoff and landing



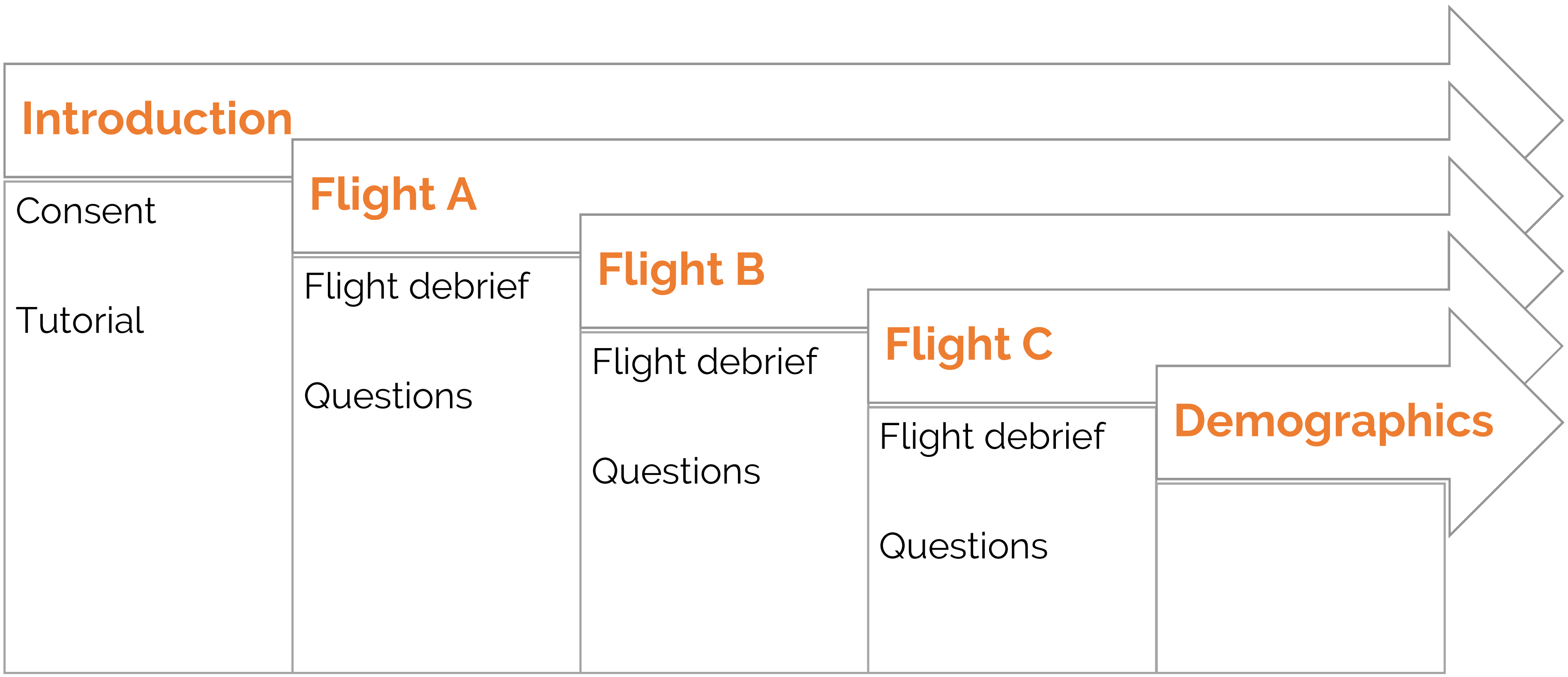
Does changing how we present risk-related feedback affect its effectiveness?

- List factors to investigate
- Design ways to present feedback
- Apply to sample flights
- Design interactive debrief prototype
    - **nicolettafala.com/debriefexample**

Does changing how we present risk-related feedback affect its effectiveness?

- List factors to investigate
- Design ways to present feedback
- Apply to sample flights
- Design interactive debrief prototype
- Survey pilots to evaluate feedback effectiveness

Surveyed pilots to see if different factors impact risk-perception or motivation to change



Questions that address the two parts of “feedback effectiveness”

- How risky do you think this takeoff was?
- Which behaviors concern you?
- What would you do to fix those behaviors?

nicolettafala.com/survey

PURDUE UNIVERSITY

RESEARCH PARTICIPANT CONSENT FORM

Data-driven safety feedback as part of debrief for General Aviation pilots

Principal Investigator: Associate Professor Karen Marais

School of Aeronautics and Astronautics

Purdue University

IRB Protocol # 1804020499

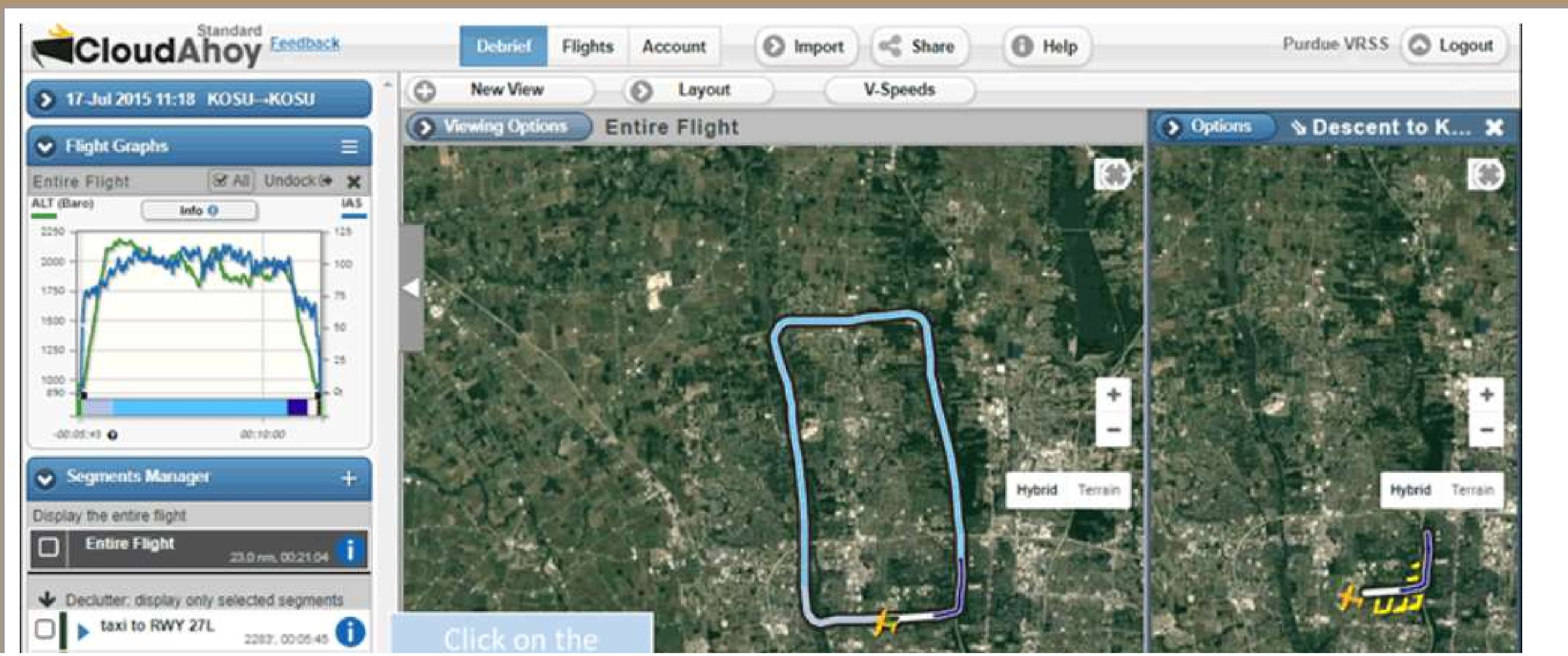
What is the purpose of this study?

This study seeks to evaluate whether data-driven post-flight debrief can be used to impact how pilots react to safety information. As a pilot, you can help us answer our research questions by evaluating the risk of hypothetical flights that you will have the chance to review. Through this research, we hope to come up with recommendations on how to communicate risk to pilots in a flight debrief format.

Review the following takeoff phase of flight as presented in these debrief screens, taking as much time as you need. The aircraft involved is a Cessna 172.

The debrief screens are semi-interactive: Under "Segments Manager," click on "takeoff KOSU RWY 27L" to choose the takeoff segment. Then click on each event you want to further investigate from the "takeoff safety information" list on the right.

When you are ready to answer questions about this takeoff, proceed to the next screen. Note that you will not be able to return to the debrief after clicking "Next."



Given the information presented to you, how ~~safe~~ risky would you say this takeoff was?

Not ~~safe~~ risky at all | | | | Extremely ~~safe~~ risky

1 2 3 4 5

In this takeoff, which of the following would concern you, if any?

- [ ] Centerline deviation
- [ ] Rotation airspeed
- [ ] Engine RPM
- [ ] Takeoff distance
- [ ] Wind

Optional comments

Next →

PURDUE UNIVERSITY

What changes (up to 5) do you think you could make to an upcoming flight after the information presented here, if any?

Change 1	
Change 2	
Change 3	
Change 4	
Change 5	

Next →

How likely are you to make each of these changes to an upcoming flight?

Not likely at all — Extremely likely

1 2 3 4 5

Change A

Change B

How important do you think each of these changes is to improving safety on takeoff?

Not important at all — Extremely important

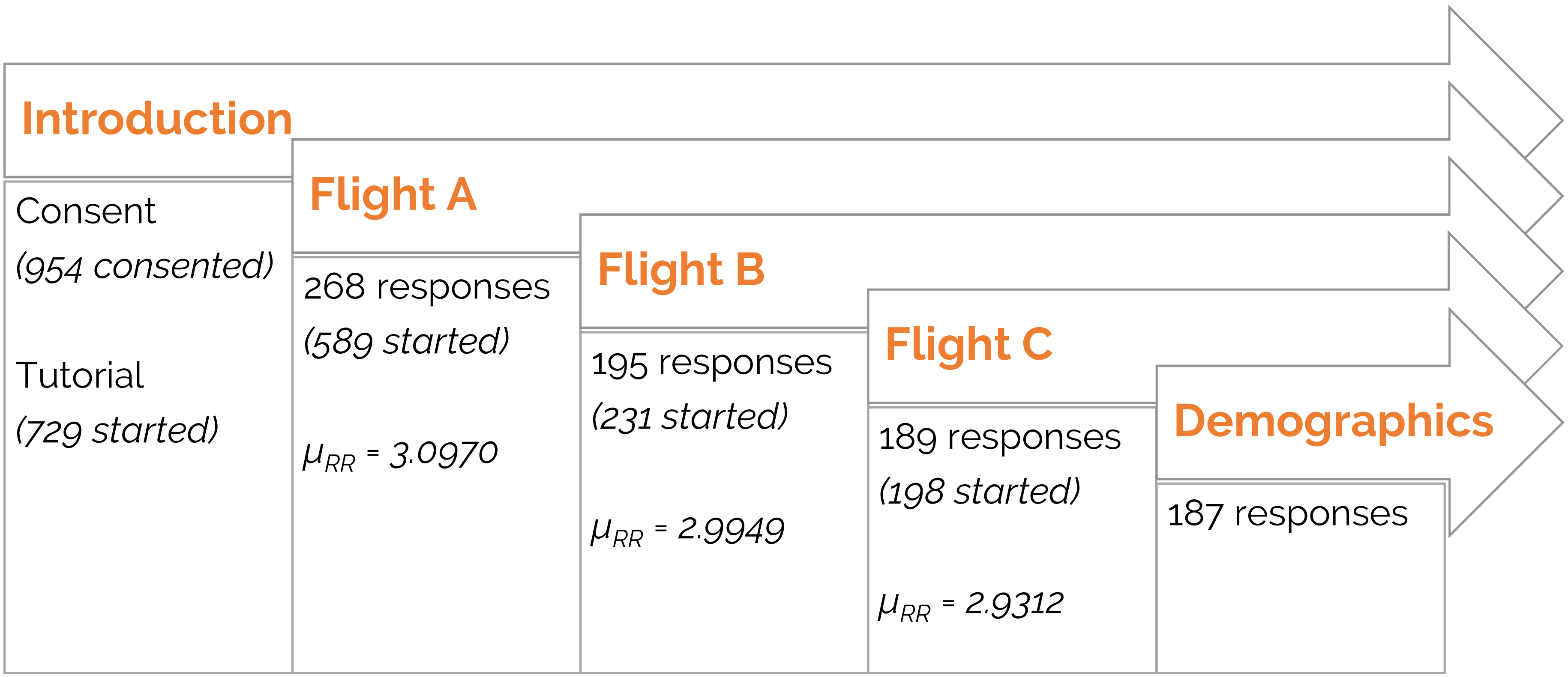
1 2 3 4 5

Change A

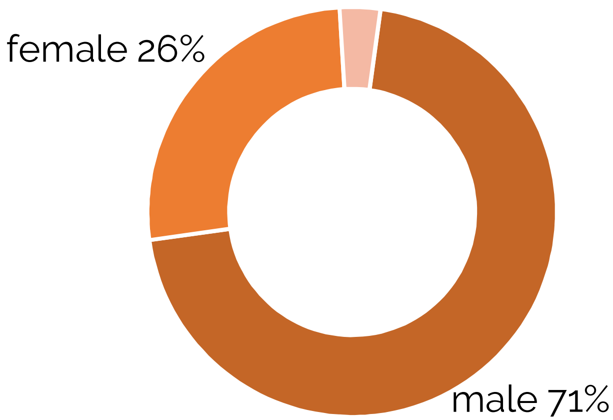
Change B

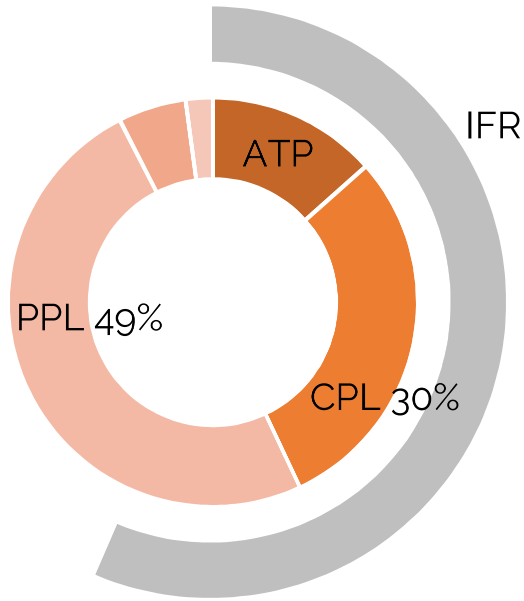
Next →

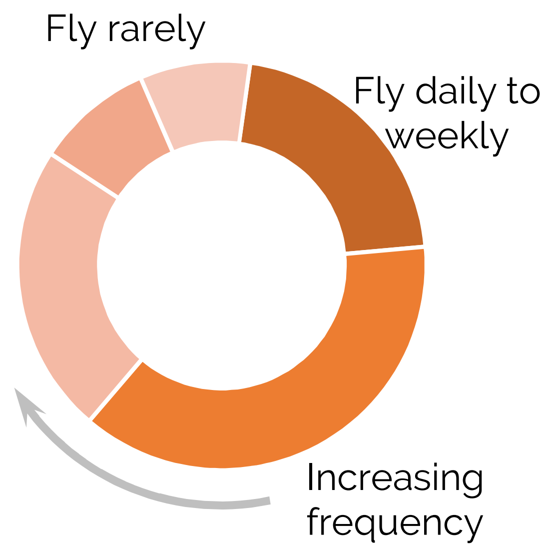
187 responses were complete—
268 were usable

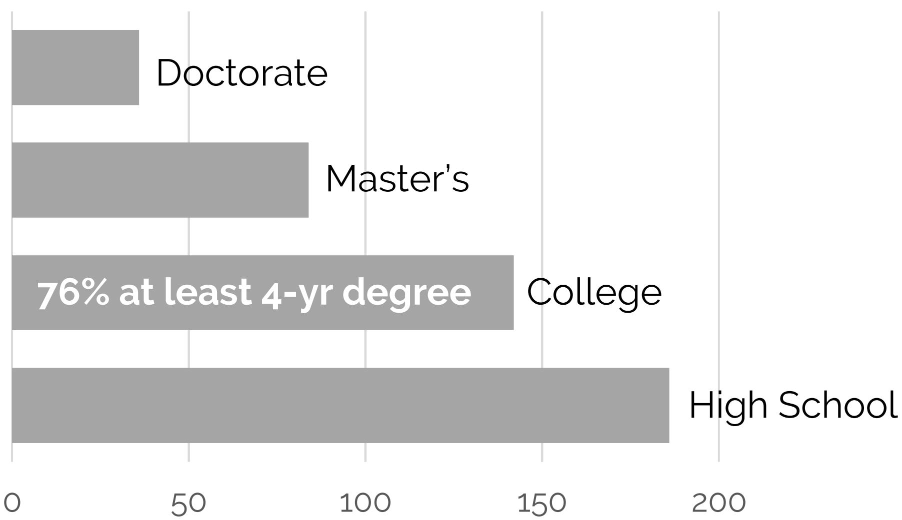


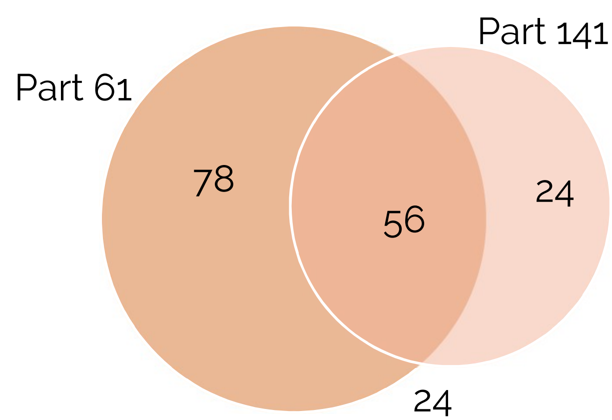
Out of the 187 complete responses... (~70% of total sample)











Survey analysis overview

- Main effects: one factor and one flight at a time
  - Representation method; parameter type; framing language
  - Risk rating; number of changes
  - Observations:
    - histograms; descriptive statistics
  - Mann-Whitney U
  - Repeat for each factor

- Interaction effects:
  - Scheirer-Ray-Hare (SRH) Test
  - ANOVA

- Repeat for each flight

Pilots were more likely to quit when reviewing the graphical representation method

	Number of completed responses					
Flight	**Graphical**		**Numerical**		**Total**	
A	123	33%	145	39%	268	36%
B	91	64%	104	73%	195	68%
C	83	77%	106	98%	189	88%

Flight B was more affected by representation method than Flights A/C

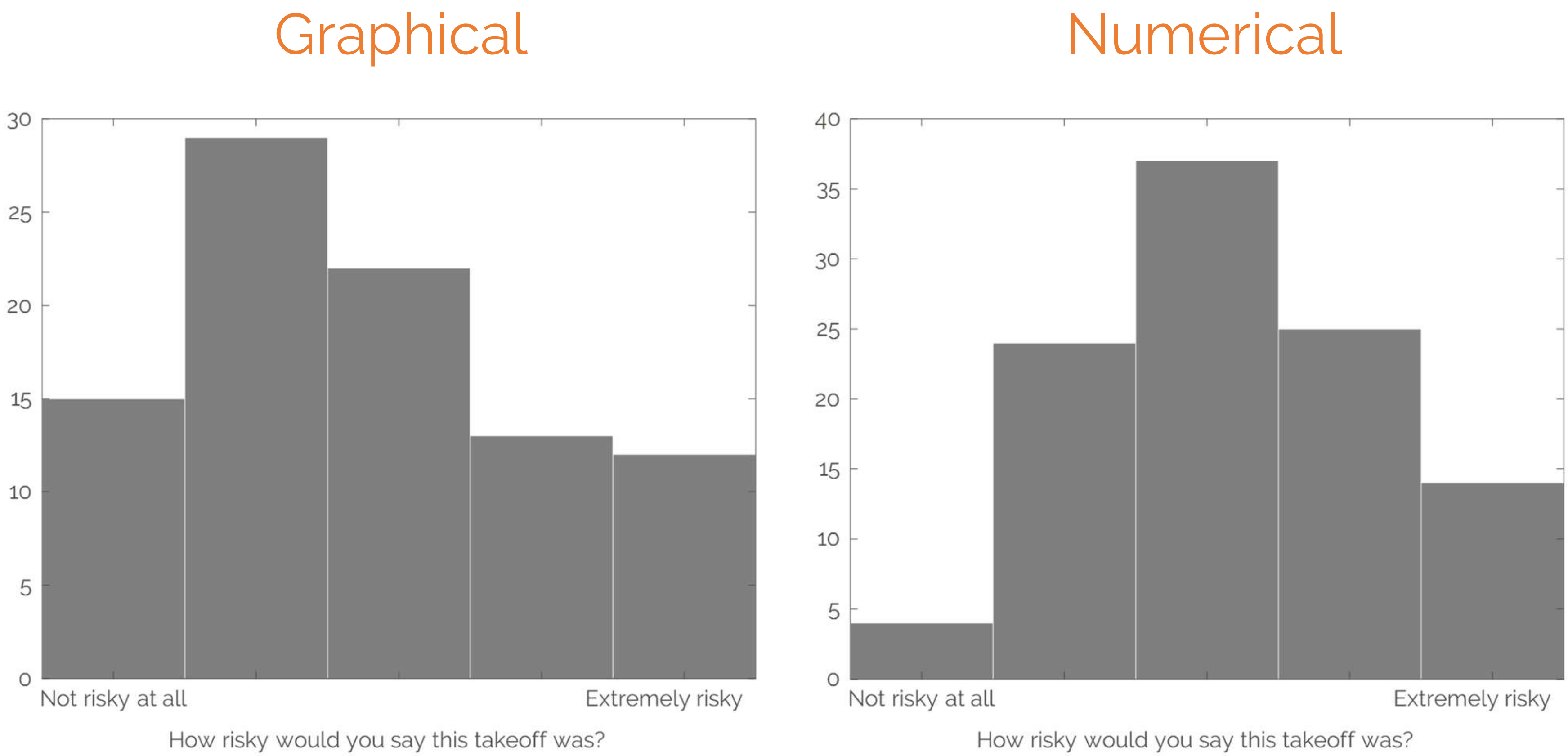
Risk rating (5-pt Likert scale)

	Graphical				Numerical			
Flight	**Mean**	**Standard Deviation**	**Median**	**IQR**	**Mean**	**Standard Deviation**	**Median**	**IQR**
A	3.1951	0.9889	3	2	3.0138	1.0340	3	2
B	2.7582	1.2679	3	2	3.2019	1.0647	3	2
C	2.9277	0.9342	3	2	2.9340	1.0353	3	2

Number of changes (0-5)

	Graphical				Numerical			
Flight	**Mean**	**Standard Deviation**	**Median**	**IQR**	**Mean**	**Standard Deviation**	**Median**	**IQR**
A	1.3984	1.3474	1	2	1.5724	1.3629	2	3
B	1.0000	0.9661	1	2	1.3654	0.8251	1	1
C	1.3133	1.1575	1	2	1.5377	1.0882	2	1

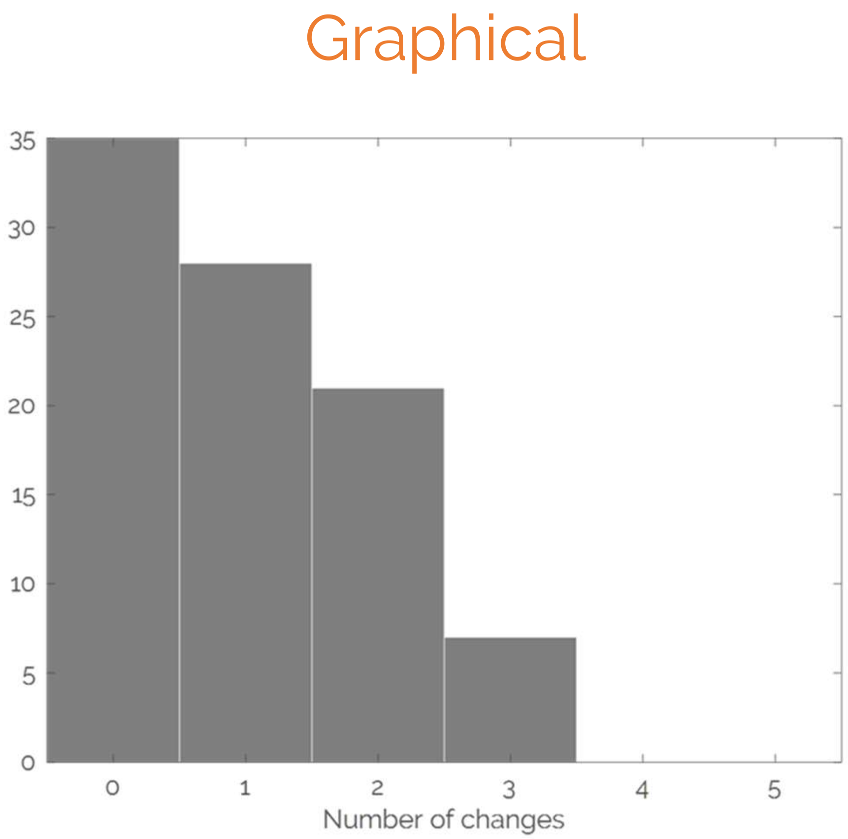
The graphical representation method makes pilots rate their risk lower

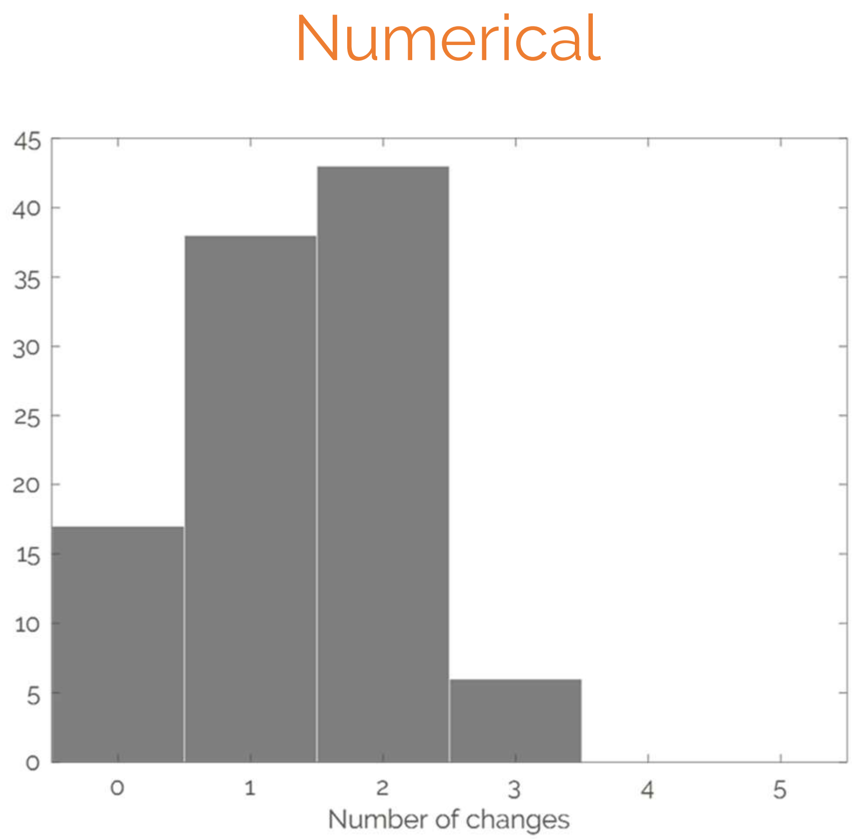


z-value	Rank sum	p-value
-2.7339	7845	0.0063

Flight B

The graphical representation method makes pilots provide fewer changes

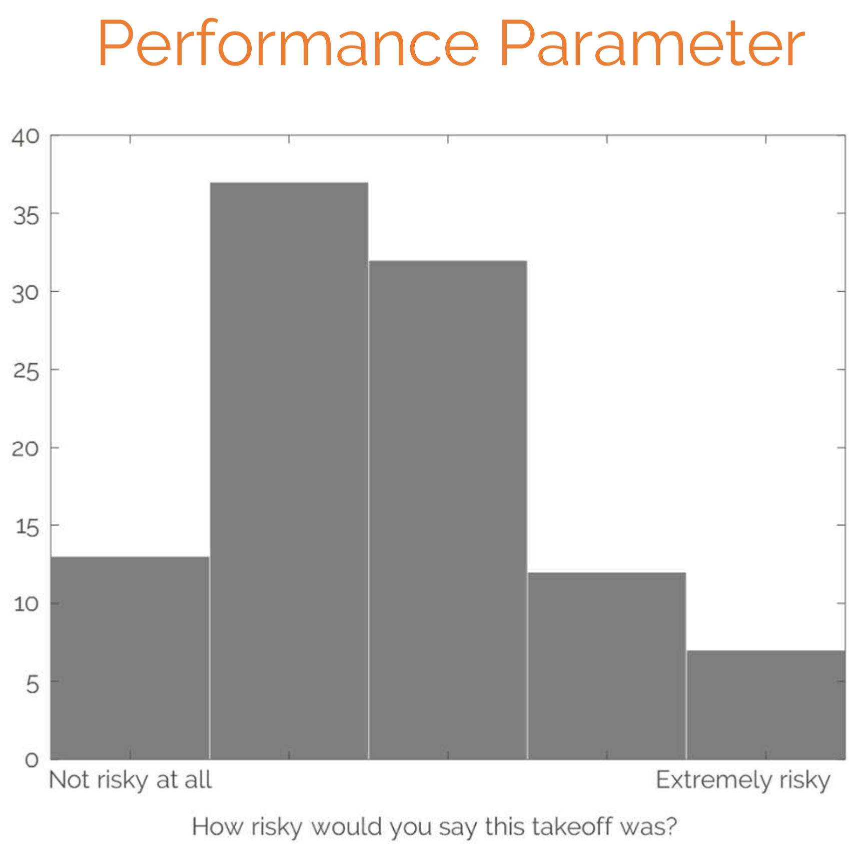




z-value	Rank sum	p-value
-2.9717	7805	0.0030

Flight B

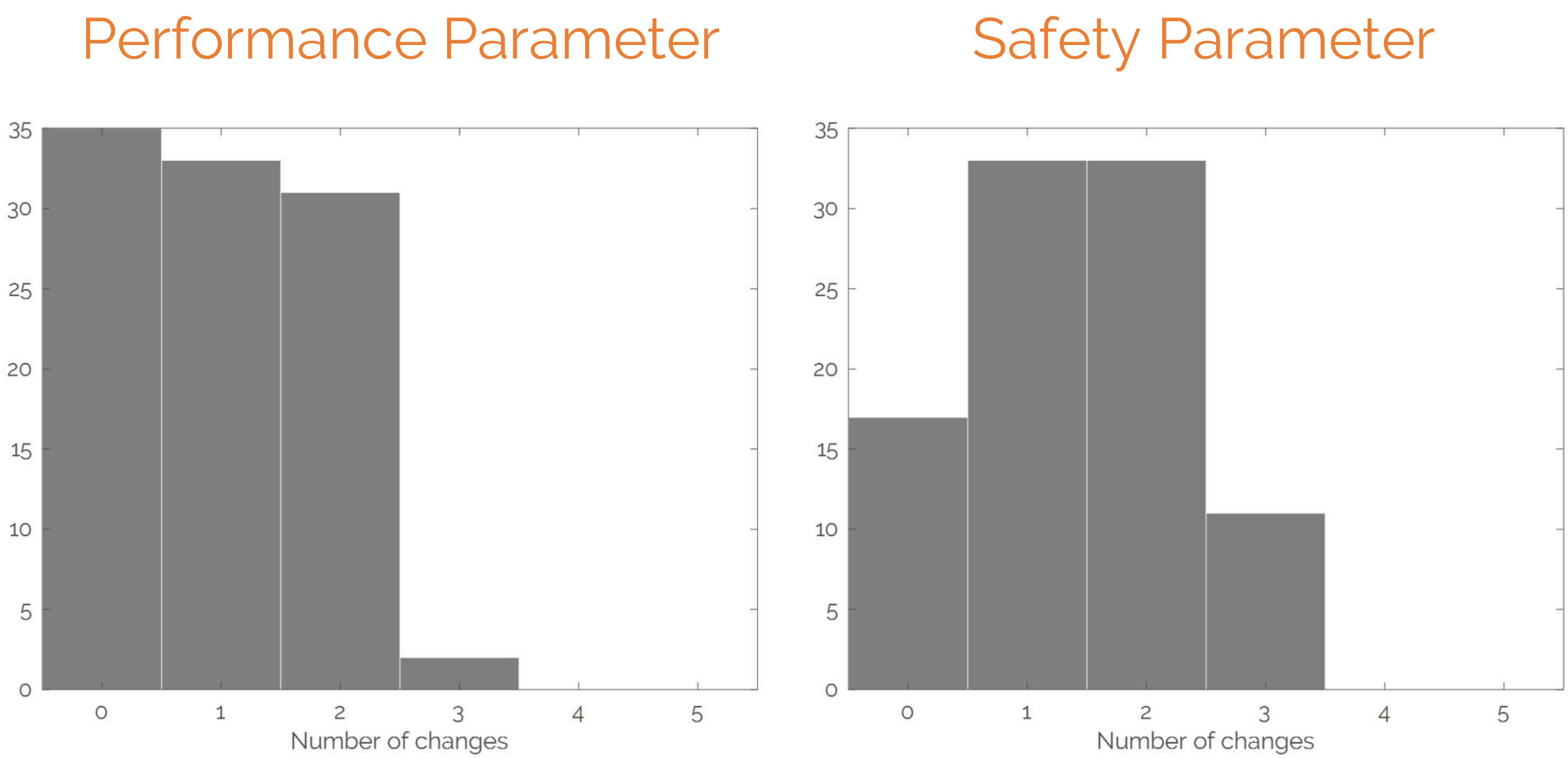
The performance parameter type makes pilots rate their risk lower



z-value	Rank sum	p-value
-4.4961	8179	6.9×10^{-6}

Flight B

The safety parameter type makes pilots provide more changes

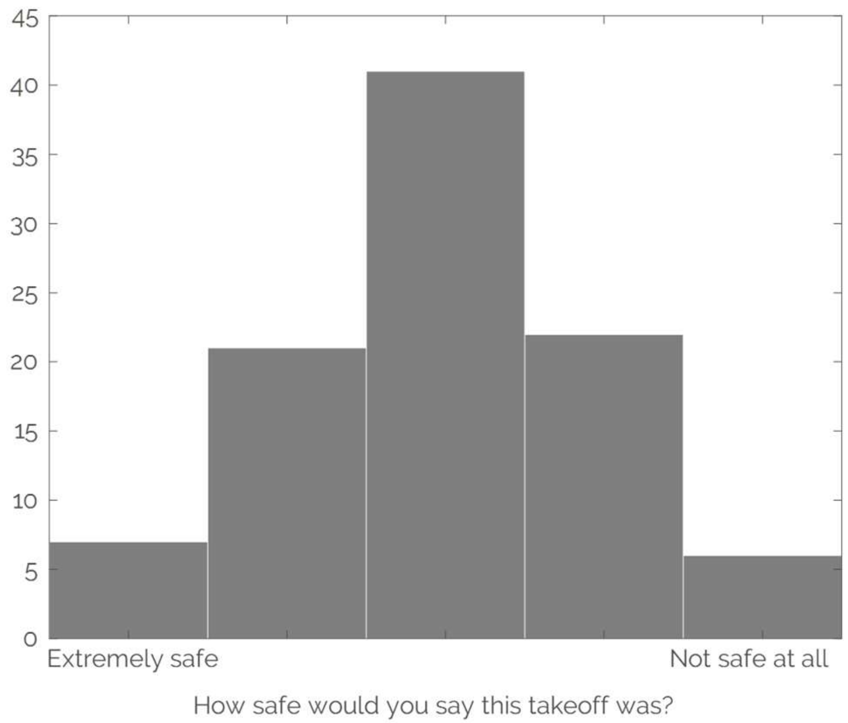


z-value	Rank sum	p-value
-2.9470	8793	0.0032

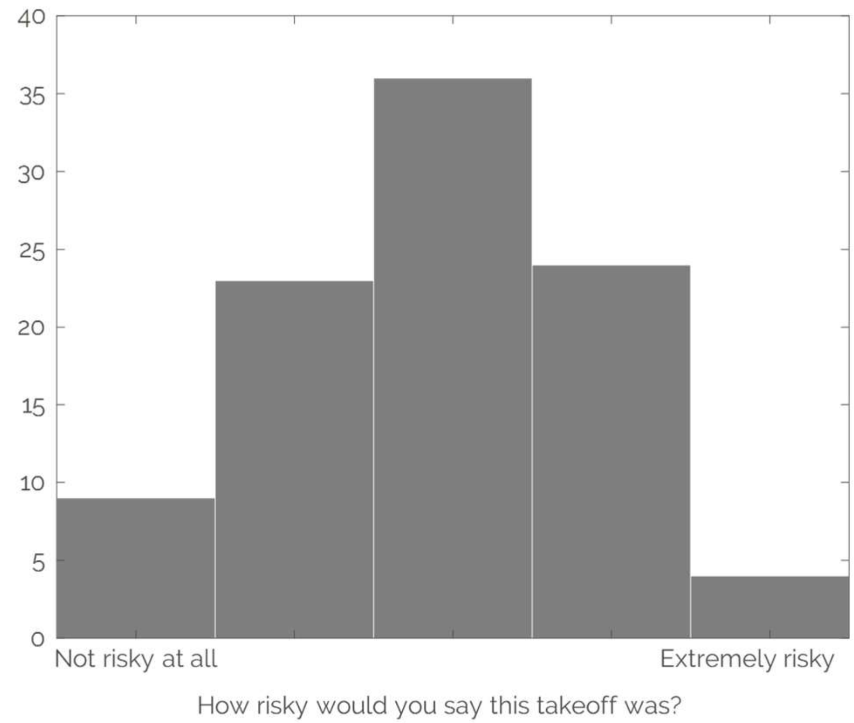
Flight B

The framing language did not impact how pilots rated the flights

Safety-centric



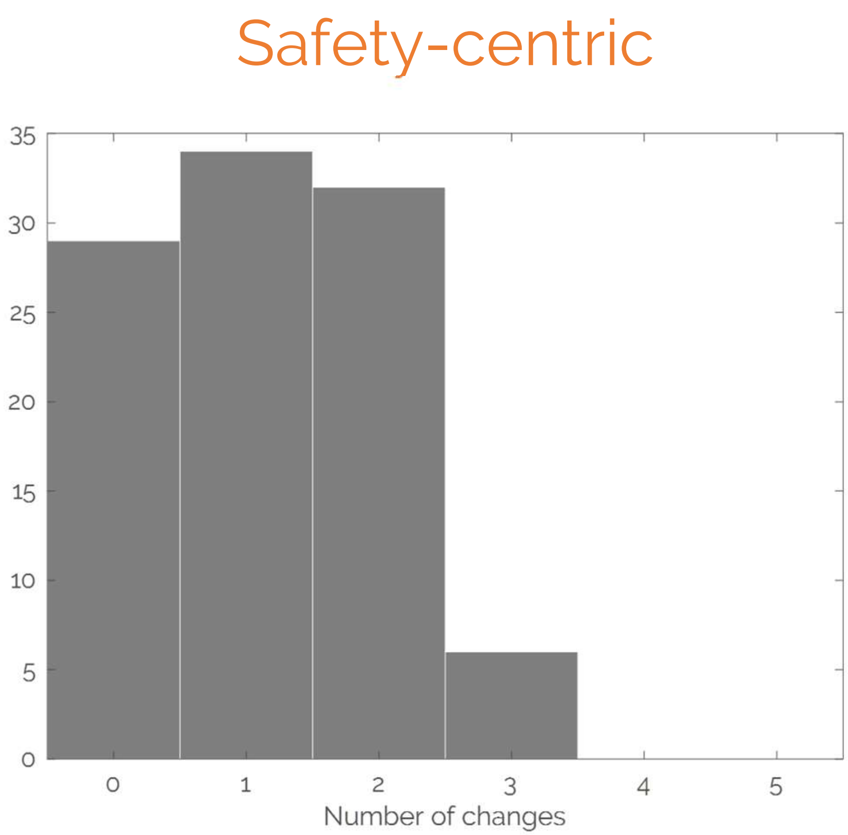
Risk-centric

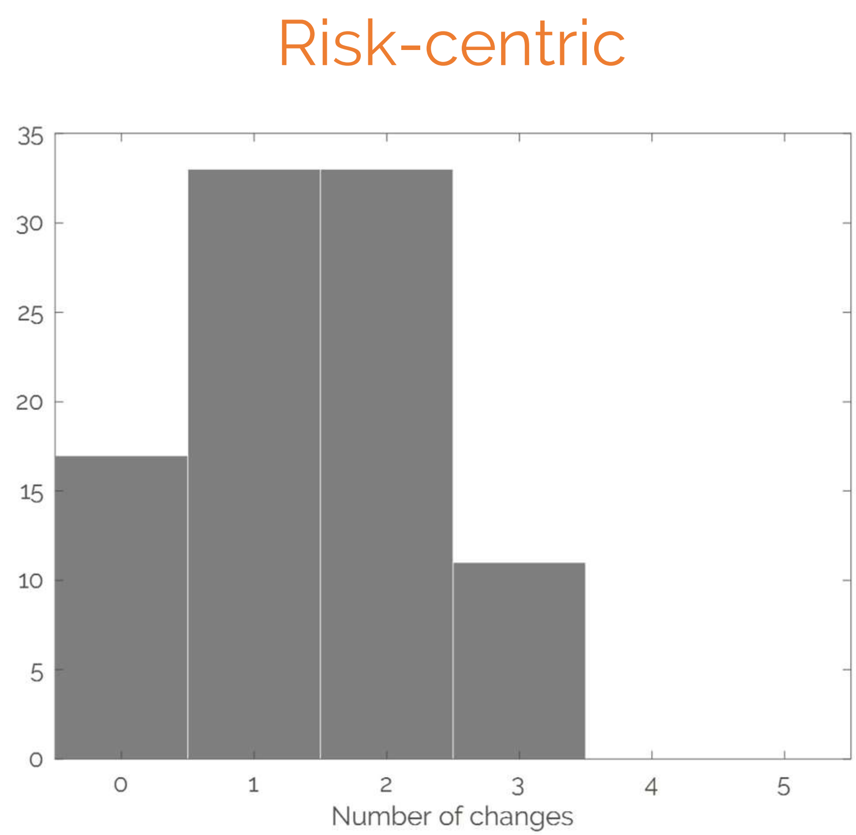


z-value	Rank sum	p-value
-0.0288	9887	0.9770

Flight B

The framing language did not impact how many changes pilots provided





z-value	Rank sum	p-value
-0.7254	9626	0.4682

Flight B

Survey analysis overview

- Main effects: one factor and one flight at a time
  - Representation method; parameter type; framing language
  - Risk rating; number of changes
  - Observations:
    - histograms; descriptive statistics
  - Mann-Whitney U
  - Repeat for each factor

- Interaction effects:
  - Scheirer-Ray-Hare (SRH) Test
  - ANOVA

- Repeat for each flight

The ANOVA for Flight B showed a slight interaction effect between representation method and parameter type

Risk rating (5-pt Likert scale)

Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
Repres	8.456	1	8.4559	6.92	0.0092
Param	28.121	1	28.1212	23.01	0.0000
Lang	0.08	1	0.0805	0.07	0.7978
Repres*Param	4.719	1	4.7194	3.86	0.0509
Repres*Lang	0.047	1	0.0472	0.04	0.8445
Param*Lang	0.045	1	0.0452	0.04	0.8476
Error	229.772	188	1.2222		
Total	270.995	194			

Number of changes (0-5)

Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
Repres	6.016	1	6.0157	7.86	0.0056
Param	8	1	8.00033	10.45	0.0014
Lang	0.167	1	0.1666	0.22	0.6414
Repres*Param	1.363	1	1.363	1.78	0.1836
Repres*Lang	0.243	1	0.24281	0.32	0.5739
Param*Lang	0.866	1	0.86562	1.13	0.2889
Error	143.882	188	0.76533		
Total	160.595	194			

The results differed for each flight, but all the tests were in agreement

Flight		Rep	Par	Lang	Rep:Par	Rep:Lang	Param:Lang
A	RR			✓	✓		
	#		✓				
B	RR	✓	✓		✓		
	#	✓	✓				
C	RR						
	#						

Overall, how we present risk information to pilots does matter…

- The flight ended up being a potential factor
  - Flight B vs Flight C
  - Different factors more prevalent in different flights

- Framing language did not change the responses as much as risk representation and parameter type

- Pilots do not like graphical representations
  - Contrary to health risk communication

Can we provide pilots with effective risk information during their post-flight debrief?

Accident Analysis

- What events/behaviors should we be trying to avoid and therefore look for in flight data?

Flight Data Analysis

- How do we calculate and detect these events?

Risk Communication

- How do we communicate such information to pilots so that they can improve?

There are limitations and opportunities for future research in these results

- How are pilots responding to the survey?
  - Commitment
  - Survey biases
- The flight as a bias
- Other cognitive biases
- Scenario-based survey
- Different sub-populations
- Smartphone data is more ambiguous

Research Contributions

- Identified unsafe events during the takeoff phase and generated list of hazardous states and triggers

- Mapped hazardous states to measurable parameters and developed algorithms to calculate and detect them

Research Contributions

- Designed debrief representations to communicate information in graphical/numerical representation methods in terms of safety and performance parameters

- Created and disseminated a survey to pilots to evaluate the effectiveness of different risk representations