Toward Single Pilot Operations: The Impact of the Loss of Non-verbal Communication on the Flight Deck

Joel Lachter[1,2]
+1 650 604 0796
Joel.Lachter@nasa.gov

Vernol Battiste[1,2]
+1 650 604 3666
Vernol.Battiste@nasa.gov

Michael Matessa[3]
+1 650 604 5975
mpmatess@rockwellcollins.com

Quang V. Dao[1,2]
+1 650 604 6620
Quang.V.Dao@nasa.gov

Robert Koteskey[1,2]
+1 650 604 4230
Robert.W.Koteskey@nasa.gov

Walter W. Johnson[2]
+1 650 604 3667
Walter.Johnson@nasa.gov

[1]San Jose State University [2]NASA [3]Rockwell Collins
Mail to: {Author Name}, NASA Ames Research Center, Moffett Field, CA 94035

ABSTRACT
Since the 1950s, the crew required to fly transport category aircraft has been reduced from five to two. NASA is currently exploring the feasibility of a further reduction to one pilot. In this study we examine the effects of separating the pilots on crew interaction. The results are consistent with earlier research on decision-making between remote groups. Pilots strongly prefer face-to-face interactions; however, we could find no impact of separation on their ultimate decisions. There were a number of areas in which separation negatively affected communications. We discuss possible mitigations for these areas.

Keywords
Single pilot operations (SPO), reduced-crew operations, teleconferencing.

INTRODUCTION
There is a growing aviation community interest in potential future operations of transport category aircraft (those flown under Part 121 of the Federal Aviation Regulations, FARs) with a single pilot rather than a two-person flight crew [1]. There are historical precedents for crew reduction. In the 1950s, the flight crew was quite large, consisting of a Captain, First Officer, Flight Engineer, Navigator, and Radio Operator. By the 1980s, technological advances had allowed the crew to be reduced to two. Under current operations, workload is very low on the flight-deck for much of the en route portion of nominal flights. If an alternative method could be found for handling high workload portions airlines could save substantial money by moving to single pilot operations (SPO); in 2010 US air carriers spent over nine billion dollars on pilots [2]. Further, a move to SPO could also help carriers reduce the complexity of scheduling and positioning crews, and allow smaller cockpits resulting in lighter aircraft and consequently reduced fuel usage.

A move to SPO might be facilitated by work currently being done on unmanned aerial vehicles (UAVs). UAVs are typically flown by a ground-based pilot who performs his duties using radio telemetry to and from the aircraft. Given this, what are the obstacles to placing the first officer on the ground?

This paper discusses an initial exploratory study conducted to look for issues with this approach, that is, having ground personnel take up some or all of the workload currently done by the first officer. Specifically, the current simulation was designed to look for effects of pilots separating pilots on their ability to work together. The results from this initial study will inform design decisions for development of a ground station that supports collaborative piloting.

A HUMAN IN THE LOOP SIMULATION

Previous Research
There is a large literature on various forms of remote collaboration dating back to the early days of computer messaging. A review article by Williams [3] discusses the major findings. For cooperative tasks with a common goal and objective outcomes, there is little difference in the actual results using different communication methods (e.g, type written messages, verbal “telephone” communications, and face-to-face). There were, however, differences in the style of communication; typing resulted in a slower process and fewer overall words exchanged than varieties of verbal communication (e.g., telephone or face-to-face). In contrast, for “conflictful” tasks such as negotiations where goals are not always aligned, there were more differences between the different forms of communication and more dependence on the particulars of the task. Negotiations conducted in less personal manner tended to be decided on their merits while those conducted in a more personal manner (such as face-to-face) appeared to be settled more

on interpersonal dynamics [3]. Similarly, in a study in which participants were to discuss issues on which their opinions differed, they were more likely to change their views when the discussion was audio-only than when it was face to face [3,4]. On the other hand, in a prisoner's dilemma type game which pits group against individual interests, group interests were more highly favored in a richer media environment (e.g., audio and visual) than a less rich one (e.g., audio only [5]). There are also interesting effects of communication type on interpersonal relationships. People prefer to converse in the richer communication channels (i.e., face-to-face is preferred to audio-visual which is preferred to audio-only) and they rate people with whom they have face-to-face contact more highly than those with which conversations are mediated. Also, in studies of group dynamics, richer communications tend to result in clearer leaders while more impoverished communications lead to egalitarian relationships.

Here we extend this research to an aviation environment. This environment differs from those previously studied in several important respects. First, pilots are given extensive training on procedures that assume they are collocated. Second, pilots have particular roles. The captain is ultimately responsible for the flight; however either he or the first officer can take on the role of pilot flying or pilot monitoring. Finally, flying an aircraft is more complicated than those tasks for which face-to-face and mediated interactions have previously been studied. One consequence of this complexity is that flying is not easily categorized as cooperative or conflictful. For example, while pilots share basic goals of safety and efficiency, they may disagree about priorization when abnormal conditions occur such as poor weather conditions or passenger illness.

Method

To examine the ability of non-collocated pilots to collaborate, we created a desktop simulation environment that allowed us to run pilots in two conditions, seated side-by-side as they are today, or seated in separate rooms with similar displays and controls but only an open microphone for communication. We felt "building a wall" between the pilots while maintaining current day operating procedures would expose aspects of current procedures and coordination that would have to be addressed in the development of equipment and procedures for coordinating a single pilot with a ground operator. It was expected that performance on some tasks would be dependent on visual cues and thus that performance on those tasks would suffer when the pilots were separated.

Participants

Ten two-pilot crews participated in this experiment; one crew per day. Pilots recruited had a minimum of 300 hours Boeing commercial glass cockpit experience in the past six months (or last six months flying before retiring). All pilots were active or had retired within the last six months.

Experimental design

The design of this experiment is extremely simple. There is one within subjects fixed factor: Together vs Separate and one random factor: Crew.

In the Together condition participant crews sat together at a low fidelity, desktop based flight simulator (see Figure 1). In the Separate condition, the FO was relocated to a second room, where the right hand seat of the dual cockpit configuration was reproduced (Figure 2). The captain remained at the left seat of the dual pilot flight deck, however, the right side monitors with the FO's PFD, Nav Display and a tablet that presented workload probes to the FO were turned off. An "open mike" communication system (functionally similar to speakerphones) was placed between the Captain and FO's stations so that all verbal communications could continue as normal. However, no body language or other non-verbal communication (e.g., sharing notes taken on ATIS or pointing to approach plates) could occur.

Concept of operations

Operations in this study were as close to current day as practical. In particular, the separation of the two pilots was not meant to introduce a new concept of remote operation. Both pilots were assumed to be in the same aircraft. However, because, we were testing the effect of non-physical co-presence on decision-making and communication, we chose to make the pilots invisible to one another. The primary differences from current day operations resulted from the limited fidelity of the simulator

Figure 1: Two person desktop simulator.

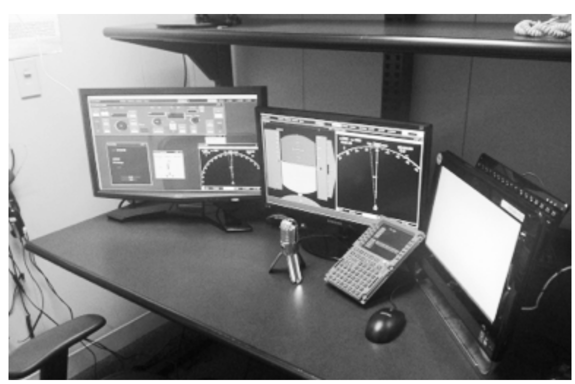

Figure 2: First officer's desktop simulator in the separate configuration.

that, for example, precluded an out-the-window view so pilots were always "head down".

Scenarios

Each crew participated in six 15-minute scenarios. In each scenario pilots were required to divert to an airport other than their original destination. Diversions were for weather, equipment failures, medical emergencies and/or airport closures. Additional events constrained the choice of airport. For example, failure of the antiskid system required pilots to calculate the stopping distance of the aircraft to determine what runways they could land on. In all cases, crews had access to a (confederate) air traffic controller.

Experimental aircraft were either arrivals flying into or departures flying out of one of six airports in the Rocky Mountain region: DEN (Denver), COS (Colorado Springs), PUB (Pueblo), CYS (Cheyenne), EGE (Eagle; Vale) and GJT (Grand Junction). This airspace was chosen because typical weather patterns can shut down many of these airports and there are no other major airports for a considerable distance surrounding this area. We felt this would give us a degree of control in what airports crews would (or at least should) ultimately decide to divert to.

Before beginning each scenario, each pilot was given a Flight Plan, Weather Briefing, Maintenance Briefing, sectionals and approach plates for airports in the region, materials similar to what they would be given on a real flight. Pilots went through this material, seated together in the Together condition and separately (with an open microphone) in the Separate condition.

Flight crews in all conditions were expected to safely and efficiently manage their flight toward a safe landing. Additionally, arriving flights were asked to land with approximately 8000 pounds of fuel.

Schedule

Each day consisted of approximately two hours of training, six 15 minute experimental runs, and a one to two hour debriefing.

Dependent variables

We collected a variety of objective and subjective measures. For subjective measures we collected both real time and post-trial workload ratings. Real time workload was collected using a probe procedure. Every three minutes the pilots heard a tone and a response panel appeared on a display to the left (for the captain) or right (for the first officer) of the flight displays (see Figures 1 and 2), from which the pilot selected one of nine boxes marked 1-Low to 9-High. The post-trial workload was assessed on a paper questionnaire with the same scale. Two workload questions were asked on this questionnaire; one asked pilots to rate their overall workload for the scenario and the other asked them to rate their peak workload. This post-trial questionnaire also asked for ratings and comments about other aspects of the scenario. There was also a more extensive questionnaire given at the end of the experimental session on which pilot participants were asked to directly compare communication and decision processes in the Together condition to those in the Separate condition. The questions presented in these questionnaires were developed using a "cognitive walkthrough" procedure in which crews were run through a subset of the scenarios that were presented in the main experiment [6].

In addition to these subjective measures we looked at objective measures of the quality of the decisions being made as well as the process by which they were made. The scenarios were designed with the goal of making one particular divert decision a clearly better choice. Whether the crews ended up at the correct airport provided an objective measure of decision quality. Similarly the amount of fuel remaining provided a measure of whether their decisions were made and executed in a sufficiently timely manner.

All interactions were videotaped to allow for analysis of particular crew actions. We were particularly interested in examining Crew Resource Management (CRM) behaviors. CRM is aimed at reducing errors by improving teamwork among the crew. It is credited with reducing the accident rate over the last 40 years [7]. CRM procedures insure that decisions are cross checked and that the reasoning behind them is explicit and understood by both crew members.

Results

A simulation such as this generates a large amount of data in many categories. To clarify the presentation, we have organized the results according to the conclusions we draw from them. We will begin by discussing the surprising (at least to us) number of measures where separation showed no adverse impact. We then turn to the measures where separation did show an adverse impact, and attempt to characterize how these measures differ from those where no differences were seen. Finally we turn to data from the post-simulation questionnaire on which pilots were asked to directly compare the Together and Separate conditions.

Measures where separation showed no adverse impact

The literature comparing audio-only to face-to-face communications suggests that, for objective tasks, performance is roughly equivalent [3]. However, given that pilots have extensive experience with procedures for when they are sitting together (today's nominal operation), one might expect superior performance in the Together condition. On many measures this was not the case. In this section we will discuss a number of measures on which we thought it likely we would find differences but found none.

Workload ratings. Pilots rated their workload in three separate assessments: Real Time ratings of workload assessed by probes during the trial, Overall ratings of average workload for a trial, assessed on a questionnaire after the trial, and Peak workload for the trial, assessed on the same post-trial questionnaire. Figure 3 shows means for all three. No differences between the Separate and Together conditions were significant ($F_{1,9} = 1.4$ for Real Time; Fs < 1.0 for Overall and Peak). This does not seem to be due to

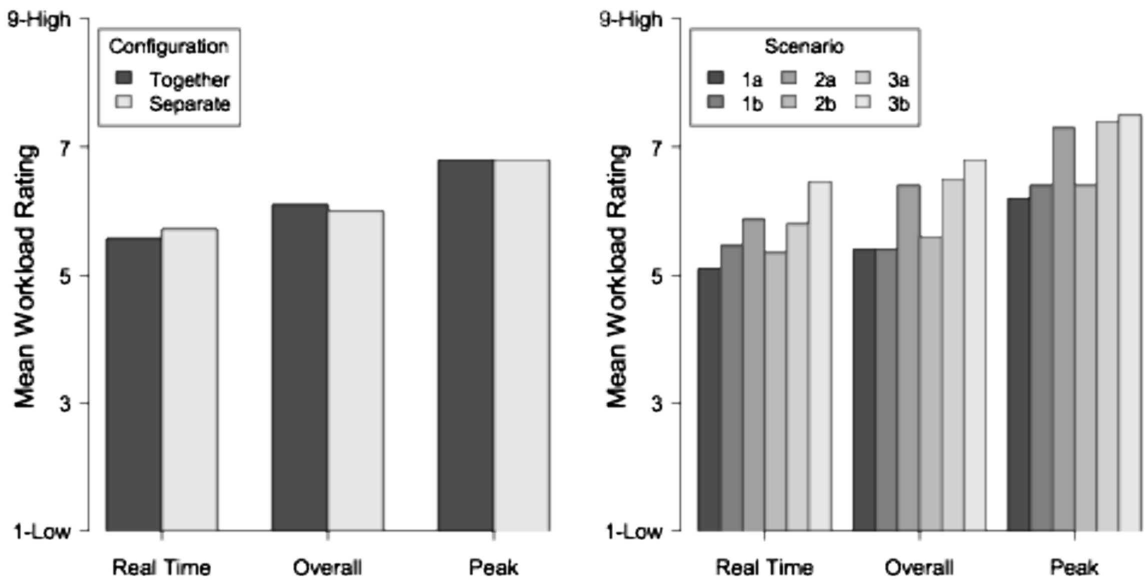

Figure 3. Pilot estimates of their workload. Left panel shows ratings broken down by configuration; right panel shows ratings broken down by scenario.

lack of sensitivity since on all three measures there were highly significant differences between the scenarios (see Figure 3; $F_{5,45} = 6.0$, $p < .001$ for Real Time; $F_{5,45} = 6.6$, $p < .001$ for Overall; $F_{5,45} = 7.2$, $p < .0001$ for Peak).

Other post-trial ratings. In addition to workload, the post-trial questionnaire asked pilots to rate ten aspects of their experience on the scenario. These questions and the pilot ratings are summarized in Figure 4. As can be seen, there is little effect for six of the ten ratings. We will discuss those

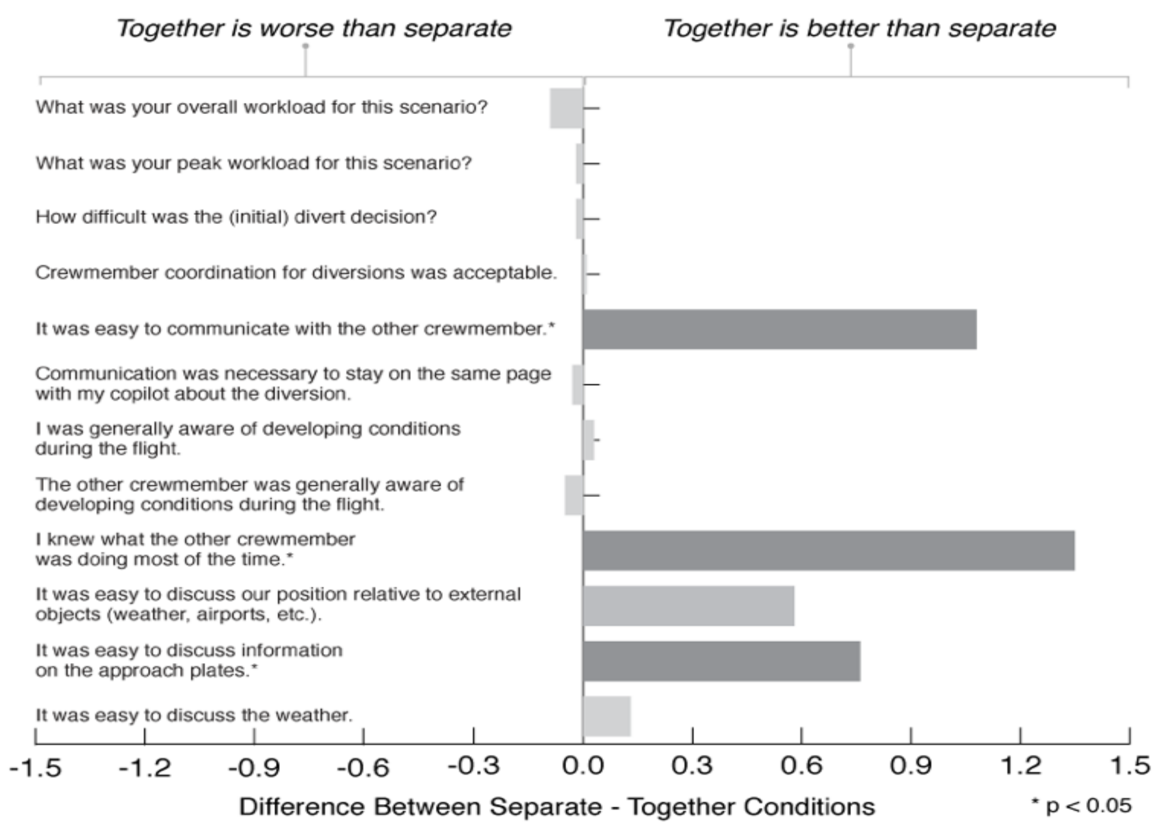

Figure 4. Differences between Together and Separate in response to post-trial questionnaire. Dark shaded bars indicated significant differences.

measures with differences under “Measures adversely impacted by separation” below. However, we found it notable that no difference was found across a variety of questions given that pilots are very practiced in procedures that assume they are side-by-side. Difficulty of the divert decision, crew coordination, awareness of developing conditions and ability to discuss the weather were all rated very similarly in the Together and Separate conditions.

Objective measures. To test for differences in outcomes between the Together and Separate conditions, we examined the final destination chosen by the crew and how much fuel they landed with. Each scenario was designed so that one of the six airports was the “optimal” airport to land at, and, in fact, a plurality of crews landed at this airport for each of the scenarios, arguing that it was, indeed, the best choice. However, on roughly one quarter of the trials, crews selected a different airport. There was no significant difference between the number of times the crew chose the optimal airport in the Together and Separate conditions, $t(9)=1.62$, $p = 0.14$. In fact, contrary to our expectations, the trend was for more flights to land at the “optimal” airport in the Separate condition than in the Together condition (80% vs. 63%).

In arrival scenarios pilots were asked to land with 8000 pounds of fuel. Based on pre-study interviews this was seen as a common practice for the aircraft simulated in this study. Fuel was not an issue in the departure scenarios but in the arrival scenarios the majority of aircraft landed with less. On average arrival aircraft landed with 7,100 pounds of fuel in the Together condition and 6,900 pounds of fuel in the Separate condition, a difference that was not significant ($F < 1$). (Note that, in our scenarios aircraft never actually landed; landing fuel is estimated by the FMS at the end of the scenario.)

Crew resource management. As noted above, CRM is considered by many to be a key factor in the increase in safety over the last 40 years. To that end, we examined one key component of CRM, the degree to which pilots double check and acknowledge input by the other crewmember. We felt this measure might reflect differences because some acknowledgements are non-verbal and these may have been disrupted when the pilots could not see each other. To get a clear count of the number of acknowledgements we looked at every time the execute button on the CDU was pressed and determined whether the changes had been acknowledged. Acknowledgement rates were about 50% in both conditions and did not differ significantly ($F < 1$).

Measures adversely impacted by separation

While there was surprising similarity in the Together and Separate conditions on many measures, there were a number of measures on which they were quite different. These measures point to specific deficits occurring when the pilots are separated. We discuss these here.

Confusions. Analysis of the video was done to find instances where the pilots appeared confused (e.g., asked what the other pilot was doing or misinterpreting the approach plates). There were 70 confusions identified, the majority of which occurred when the pilots were separated. Fifty-two confusions occurred in the Separate condition while only eighteen occurred in the Together condition. This is a significant difference ($\chi^2 > 15$, $p < .001$). We categorized the confusions according to what the pilots were confused about and whether they were in the Separate or Together conditions (see Table 1). Most confusions fell into three categories: “What is the other pilot doing?”, “Where is information in the briefing material?”, and “Where is information on the approach plates?”. For each of these categories many more confusions occurred in the Separate condition than in the Together condition, although for other confusions this pattern was less obvious.

One should be careful in interpreting these observed confusions. It is likely that some confusions occurred that were not observed and they may have been easier to observe in the Together condition. For example, if a pilot was not sure what his co-pilot was doing in the Together condition he might be able to find out by watching him, while in the Separate condition he might have to ask. The latter would almost certainly be more easily observed than the former. However, these confusion data closely match the ratings data reported below where observability is not an issue. Thus, we feel it is very likely that the differences seen here were differences experienced by the pilots.

Post-trial ratings. As noted above, many questions from the post-trial questionnaire were rated almost identically in the Together and Separate conditions. However, significant differences were found when pilots were asked to rate their agreement on several questions (see Figure 4). Specifically, “I knew what the other crewmember was doing most of the time,” (Together 2.28, Separate 3.63; $F_{1,9} = 18.75$, $p < .01$); “It was easy to discuss information on the approach plates,” (Together 3.16, Separate 3.92; $F_{1,9} = 10.6$, $p < .01$); and “It was easy to communicate with the other crewmember,” (Together 2.30, Separate 3.38; $F_{1,9} = 12.2$, $p < .01$). Further, a marginally significant difference was

Table 1: Number of confusions

	Separate	Together
What is the other pilot doing?	14	0
Where is information on the approach plates?	8	3
Where is information in the briefing material?	6	0
Other	24	15
Total	52	18

found when they were asked to rate their agreement with: "It was easy to discuss our position relative to external objects (weather, airports, etc.)," (Together 2.63, Separate 3.63; $F_{1,9} = 3.8$, $p = .08$). Notice that the first two questions correspond to the two most common sources of confusion found in the confusion analysis: what the other pilot was doing and the approach plates. The third is a generic question about communication. These three all appear to deal with matters where one might expect a greater reliance on non-verbal communications; the actions of the other pilot are usually observed and approach plates and briefing materials can be passed back and forth and pointed to. Similar issues also arise in discussing the positions of external objects, especially in this experiment where there was no out-the-window view. Location information was only found on paper (the approach plates for airport locations; ATIS slips for weather) and on the Navigation Display. These are information dense displays that pilots are probably used to sharing and pointing at and thus it is logical that using them should be judged more difficult in the Separate condition.

Responding to probes. In analyzing the probe data (see Figure 3, above) we came across an interesting phenomenon. For two of our 20 crews, the FO largely ignored the probes in the Separate condition but answered quite regularly in the Together condition. In one crew the FO responded to none of the 15 probes in the Separate condition, but 12 of the 13 probes in the Together condition ($\chi^2 = 20.6$, $p < .0001$). In the other crew, the FO responded to 3 of the 14 probes in the Separate condition, but 10 of the 14 probes in the Together condition ($\chi^2 = 5.2$, $p < .05$). The exact reason for this is unclear. Seeing the other pilot react to the probe could serve as an added cue to respond. Alternatively, the presence of another pilot might provide “social pressure” to respond. Both of these point to potential issues for SPO. A second pilot may provide an additional cue to perform tasks, and may provide social pressure to perform tasks (like checklists) that are not productive 99% of the time but are occasionally critical.

Post-simulation questionnaire comparisons

Despite similarities in their overall performance, pilots generally disliked the Separate condition. This can be seen most clearly in their responses to the post-simulation questionnaire and their comments in the debriefing. A particularly interesting contrast is between the ratings of actual workload during the trial (as seen in the workload probes and post-trial questionnaire discussed above) and ratings of workload on the post-simulation questionnaire.

Workload comparisons. On each trial we obtained ratings of workload during the trial (Real Time) or just after the trial (Overall and Peak Workload ratings). Then in a post-simulation questionnaire we asked pilots to rate whether their workload had been higher when they were Together or Separate on a nine point scale (with the center marked as No Difference). Frequency histograms for pilot ratings on all four measures are shown in Figure 5. For the probe and post-trial ratings (panels A- C), the mean Together rating was subtracted from the mean Separate rating for each pilot, while panel D uses the direct comparisons asked for the in post simulation questionnaire.

Panel D shows that fifteen pilots responded that their workload was higher when they were Separate and four responded that they were equal and only one responded that workload was higher when Together. This contrasts with the results found for trial specific ratings (panels A-C), where no significant differences were found between the Separate and Together conditions. Clearly many pilots who rated the Together trials and the Separate trials the same during or immediately after the trial, later rated workload as being higher in the Separate condition. Why?

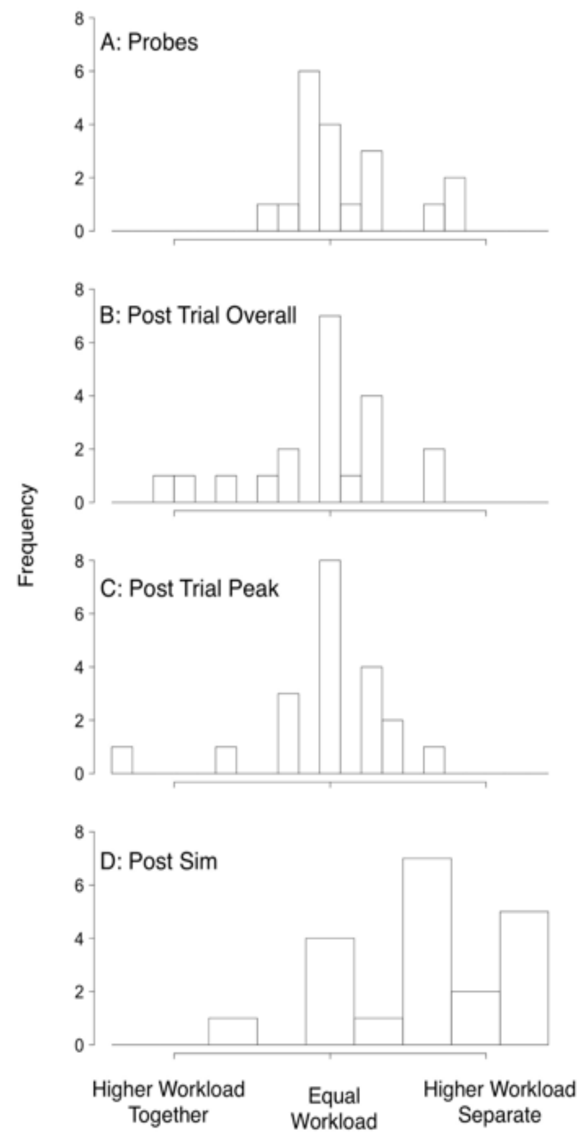

Figure 5: Differential workload by pilot. A-C) The mean of all workload ratings for each pilot in the Together condition subtracted from the mean in the Separate condition. D) Ratings of whether workload had been higher in the Together or Separate condition.

We thought of several related reasons as to why the data might show a different relative workload for the Together and Separate conditions as a function of when the ratings were taken. The first is the way the questions are worded. The probes and post-trial questionnaires ask pilots to rate their workload. The post-simulation questionnaire, on the other hand, asks for a comparison or choice ("My workload was higher when…"). It is possible that such a "two alternative forced choice" task was a more sensitive measure of differences in workload. An alternative is that the post-simulation question asks the pilots to average over a variety of situations encountered across the six scenarios. Thinking back and comparing their experiences in the Together and Separate conditions may induce pilots to think more generally about the concept.

Another alternative explanation of the difference between the direct ratings of workload on the probes and post-trial questionnaires versus comparison on the post-simulation questionnaire is that on the post-simulation questionnaire pilots rate something easier to call to mind than workload. Kahneman has argued that such substitutions are quite common [8]. Evidence for this comes from comments made by the pilots in a comment section following the workload question. Ten pilots cited increased need to talk as a reason for their higher workload ratings, and six cited a lack of awareness as to what the other pilot was doing. Interestingly, these issues (ease of communication and knowing what the other pilot was doing) were the two issues where the Separate condition was rated the most inferior to the Together condition on the post-trial questionnaire where no effect of workload was found.

Comparisons between the Separate and Together conditions. The post-simulation questionnaire included ten questions where pilots were asked to rate whether the Separate or Together condition better exemplified safe and efficient flight operations across all flight regimes. These included questions on safety, workload (see above), ease of communication, coordination, and decision-making. With three exceptions, no pilot rated the Separate condition as better exemplifying a positive quality or the Together condition as better exemplifying a negative one. Thus, the overwhelming majority of responses indicated a preference for the Together condition.

Opinions on possible mitigations

The goal of this initial study was to surface potential issues with pilot coordination and decision making caused by separation. Cognitive walkthroughs conducted in the run-up to the experiment suggested that not knowing what the other pilot was doing and a disruption of the standard procedures for acknowledgements were two likely issues. As a result, we asked the pilots on the post-simulation questionnaire to rate two possible mitigations generated in the cognitive walkthroughs and asked an open-ended question regarding other possible mitigations.

Video. To mitigate the lack of awareness about the other pilot's actions, one suggestion was to install video so that the pilots could see each other. Pilots were asked to rate their agreement or disagreement (on a nine-point scale, with 1- positive and 9 negative) with the statement "I would like to have some way to see my crewmember when we are separated (e.g., a video link)." The mean rating was 4.9, (5 was neutral). While some were positive toward the idea (one commented "Good idea!" and another "Just to have emotional connection w/ co-worker"), many seemed to feel that the video "Would be mostly a distraction."

CRM indicators. Another idea was to give pilots a visual signal as a method of acknowledging input (an idea we have subsequently referred to as CRM indicators). This idea also had many negative responses, but it had many positive ones also. Asked to rate their agreement with the statement "I would like to have a visual signal that the other crewmember concurred with my input rather than counting on verbal communication in the Separate condition (e.g., the altitude window turning green when the other crewmember agreed it was set correctly)," the average rating was 3.7 (on a 1 to 9 scale; moderate agreement).

Other suggestions. In response to an open-ended question ("What other tools or procedures would have helped you in the separate condition?"), four pilots suggested more explicit SOPs and two commented that they liked the CRM indicator idea. More explicit SOPs (like the CRM indicators) might serve to help pilots better understand what their partner was doing. In addition, one suggested "A 'common display' (i.e. an electronic flight bag, or EFB) that would automatically display the same data to both pilots such as approach plates, or ATIS". This idea might help to solve issues with the approach plates and briefing materials.

CONCLUSIONS

It is clear that pilots generally did not like the Separate condition. This was expected, given their training and long experience flying side-by-side. It is then, perhaps, more surprising that greater differences were not found on our performance metrics or ratings taken during or immediately after each scenario. Our data suggest that the effects of separating pilots on their ability to collaborate are primarily focused on a relatively narrow range of issues. Primary among these is an ability to know what the other pilot is doing. This showed up in both ratings data and in the observed confusions. There were also issues around certain exchanges between the pilots where, when together, they could point or exchange paper (e.g., charts). This showed up in the number of confusions around the approach plates and briefings as well as the superiority of the Together condition in ratings of the ease of discussing the approach plates, and, to a lesser extent, when discussing external objects which were generally represented graphically on the Navigation Display and charts. These differences may explain the overall lower rating of the Separate condition for communication; although this may also be attributed to a common preference for face-to-face communication.

These findings match those found in academic laboratory settings. In that literature, the quality of decisions is generally found to be as good if not better when participants are separated, despite the fact that participants generally prefer to communicate face-to-face. In that literature there are exceptions; as here, people have more difficulty with interpersonal tasks that require knowing the state of other participants. This study incorporated a number of materials crucial to aviation about which it was also easier to communicate face-to-face, in particular the approach plates and briefing materials. The most likely reason these are easier to discuss face-to-face is that communication is simplified by the ability to point and pass these objects back and forth. The fact that our findings line up with those found in laboratory settings is encouraging because it suggests they are robust to incidental changes such as scenarios, equipment and, to some degree, the concept of operations.

Two ideas for mitigating the problems observed in this study received some support from the post-simulation questionnaire. Having some sort of visual indicator to allow for non-verbal acknowledgement of crewmember actions was generally supported (although a quarter of the pilots did not see such a scheme as being useful). Having a video feed so that the pilots could see each other got fewer positive and more neutral responses than the indicators, but was still rated positively by nine of the 20 crew members. With a proper implementation it might be useful to a large proportion of the pilots. One pilot also suggested a shared EFB that would keep pilots “on the same page”. In a follow-on study presented in this volume, we implemented these suggestions in a prototype ground station. In that study pilots rated each of these tools positively; however crews still showed a strong preference for flying in the traditional side-by-side configuration [9]. Finally, several pilots named more explicit SOPs as a possible mitigation to the problems of Separation. Improved procedures and training may reduce the confusions and ease the pilot discomfort found here. This study serves primarily to point to areas where training and elaboration of procedures is necessary.

ACKNOWLEDGMENTS

We would like to thank Summer Brandt and Sarah Ligda for their help setting up and running the simulation and Shu-Chieh Wu for editing help. Funding for this work was provided by the Concepts and Technology Development Project of NASA’s Airspace Systems Program.

REFERENCES

1. Comerford, D., Brandt, S., Lachter, J., Wu, S.-C., Mogford, R., Battiste, V., Johnson, W. W. (2012) *NASA’s single pilot operations technical interchange meeting: Proceedings and findings.* (Report no. NASA-CP-2013-216513). Moffett Field, CA: NASA Ames Research Center.
2. Airline Data Project. (2012). Retrieved from http://web.mit.edu/airlinedata/www/2011%2012%20Month%20Documents/Expense%20Related/Cockpit/Total%20Cockpit%20Cost%20-%20ALL%20AIRCRAFT.htm
3. Williams, E. (1977). Experimental comparisons of face-to-face and mediated communication: A review. *Psychological Bulletin 84:* 963-976.
4. Chen, F. S., Minson, J. A., Schöne, M., & Heinrichs, M. (2013). In the eye of the beholder: Eye contact increases resistance to persuasion. *Psychological Science: 24, 2254-2261*.
5. Wichman, H. Effects of isolation and communication on cooperation in a two-person game. *Journal of Personality and Social Psychology,* 1970, *16,* 114- 120.
6. O'Connor, R., Roberts, Z., Ziccardi, J., Koteskey, R., Lachter, J., Dao, Q., Johnson, W., Battiste, V., Vu, K.-P. L., & Strybel, T. (2013). Pre-study walkthrough with a commercial pilot for a preliminary single pilot operations experiment. In S. Yamamoto (Ed.): *HIMI/HCII 2013, Part II, Lecture Notes in Computer Science, 8017*, 136-142.
7. Diehl, A. (1991). The effectiveness of training programs for preventing aircrew “error”. In R. S. Jensen (Ed.), Proceedings of the 6th International Symposium on Aviation Psychology (pp. 640-655). OH: The Ohio State University.
8. Kahneman, D. (2011). Thinking, Fast and Slow. New York, NY: Farrar, Straus, and Giroux.
9. Lachter, J., Brandt, S.L., Battiste, V., Ligda, S.V., Matessa, M., Johnson, W.W. Toward single pilots operations: Developing a Ground Station. HCI-Aero 2014.