Multimodal Information Management: Evaluation of Auditory and Haptic Cues for NextGen Communication Displays*

DURAND R. BEGAULT,[1] *AES Fellow*, **RACHEL M. BITTNER,**[2] *AES Member*, **AND**
(Durand.R.Begault@nasa.gov) (rmb456@nyu.edu)

MARK R. ANDERSON,[3] *AES Fellow*
(Mark.R.Anderson@nasa.gov)

[1] *Human Systems Integration Division, NASA Ames Research Center, Moffett Field, CA 94035 USA*
[2] *Courant Institute of Mathematical Sciences, New York University, New York, NY 10012 USA*
[3] *ASRC, NASA Ames Research Center, Moffett Field, CA 94035 USA*

Auditory communication displays within the NextGen data link system may use synthetic speech messages to replace voice communications. An interface design for selecting among multiple incoming messages can impact both user performance and preference. Two design factors were evaluated: physical pressure-sensitive switches versus flat panel "virtual switches," and auditory feedback from switch contact. Performance with stimuli using physical switches was 1.2 s faster than virtual switches (2.0 s vs. 3.2 s); auditory feedback provided a 0.6 s performance advantage (2.3 s vs. 2.9 s). There was no interaction between these variables. Preference data were highly correlated with performance.

1 INTRODUCTION

This paper reports the results of a trade study for two proposed versions of a multichannel interface for aviation communications. We begin by reviewing the motivation for the study within the context of the Next Generation Air Transport (NextGen) system currently under development in the United States. Next, we present details of a classic multi-line telephone, the design inspiration for the interfaces tested. Finally, the methodology and results of the trade study are described.

1.1 Problem Statement

The increased operational autonomy of flight crews in the NextGen environment will potentially result in higher overall workload and greater demands on visual and auditory modalities to safely interact with automation and the overall complexity of the future flight environment. The National Research Council's "Decadal Survey of Civil Aeronautics: Foundation for the Future" identified as a high priority "Interfaces that ensure effective information sharing and coordination among ground-based and airborne human and machine agents," and "Interfaces and procedures that support human operators in effective task and attention management" [1]. To mitigate overloading of the visual perceptual system, auditory displays can be enhanced beyond normal radio communications and caution-warning signals to include synthetic speech messaging. These messages can convey data such as flight status and trajectory for shared situational awareness among aircraft and from aircraft to ground control.

Currently, pilots communicate with air traffic control (ATC) and company primarily via voice radio frequency but can also communicate short text messages via a control and display unit (CDU), primarily using protocols of the Aircraft Communications and Reporting System (ACARS). The messaging capabilities and functionality of data link have continually increased since its introduction in 1978, with a corresponding increase in visually acquired text information. Compared to listening, such text-based systems have an obvious "bottle neck" effect on the rate of information that can be transmitted and acknowledged [2].

One possible scenario for an improved information display within the NextGen data link system will involve multiple synthetic speech messages replacing current ATC and company text or radio voice communications [3]. Synthesized voice data link information has been shown in one study to be preferred over text messages [4] and in

---

*This is an updated and expanded version of AES 133rd Convention paper 8733.

another to improve overall performance and reduce workload [5]. Another study examined the use of synthetic speech to provide virtual versions of "party line information," i.e., auditory communications from other surrounding aircraft (referred to as audio "twittering" in [6]). Prior research has shown that the traffic and weather information obtained from monitoring a party line radio frequency is critical for maintaining situational awareness [7]. Party line information from en-route proximate aircraft will become even more essential in air operations that require autonomous decision-making.

Proposals have been made for improving NextGen auditory displays, such as implementation of spatial auditory cues, and varying the prosody and speaking rate of stored speech [3,8]. By contrast, investigations of designs to accommodate the user's interaction with auditory displays have been for the most part neglected.[1]

In a future system, multiple speech communication messages may be sent to the flight deck from automated systems, resulting in an *audio message array*, i.e., a set of messages that cannot be audited in real time. Message arrays are currently familiar in the flight deck from the stream of text messages that can be selected by scrolling and reading from the CDU. However, there are temporal advantages to acquiring information via listening as opposed to reading, particularly when complex visual displays require constant vigilance. There is a need for a system that organizes such messages so that they may be selected for listening in terms of priority or age of the message and delayed if they interrupt a higher-priority activity. Furthermore, the user control interface for the system should be designed so as to optimize both objective measures of *performance* (response time, completion time) and subjective measures of *preference* (acceptance, perceived performance) on the part of users. Latorella [9] investigated the impact of auditory versus text-based data link messaging and concluded "Flight deck performance may improve by providing . . . a playback feature allowing pilots to confirm their interpretation of interrupting annunciations."

Recently, flat-panel touchscreens have been increasingly used in place of traditional manual controls, charts, and displays for reasons including ease of reconfiguration and expense [10, 11]. These include simulated "virtual" buttons and switches modeled upon their analog equivalents that are significantly different in that finger action is interpreted by the software using duration and area of contact by the finger, as opposed to physical displacement of the control. Another feature of many physical buttons and switches is that they provide auditory feedback cues as to their status (on/off) and location as a consequence of their mechanics, while touch screens must synthesize audio cues if they are to exist. Merchel et al. have experimented with synthesized audio-driven tactile feedback on "groovebox" touch screens for audio loop discrimination [12]. Altinsoy has provided a recent overview of the perceived quality of auditory-tactile environments, citing the importance of multimodal "confirmation" of user expectation within a virtual environment; presumably, a silent "switch" would be a "disappointment" of user expectation [13].

Fig. 1. Top: Western Electric Model 565 telephone. Bottom: switch assembly

1.2 Prior Development of Multi-Line Communication Interfaces for Message Arrays

A ubiquitous desktop telephone design, seen in models produced by the Western Electric Company (a subsidiary of AT&T) during the 1940s–1970s, used a multi-button interface design for managing incoming calls across different "lines" (red hold button, transparent line selection buttons). Fig. 1 shows an exemplar telephone with a six-button interface (Western Electric model 565, circa 1970s). The basic design for line selection remains influential for modern office desktop telephones. Its simplicity exemplified a successful human factors approach to the problem of managing multiple lines of communication. Its raised mechanical buttons provide robust tactile, haptic, and auditory cues, unlike the touchscreen or membrane buttons used in its modern equivalents.

The buttons are mechanical two-state switches that provide several forms of haptic, auditory, and visual cues as to their status. The spacing and layout of the buttons conforms to the hand and fingers for ease of operation. Due to their nearly 0.5-inch raised profile, tactile exploration of the button set without "consequence" (accidentally engaging the button) is possible using only tactile and auditory cues.

Each of the protruding plastic buttons is housed in a plastic sleeve that contacts the resonant plastic body of telephone. As a result, moving the button laterally or pushing it lightly (not fully inwards, e.g., to explore its state using only tactile feedback), results in distinctive "click" or "rattle" sounds depending on the manner of touch. The buttons face an array of spring loaded, single pole, single throw line switches mounted to a metal telephone assembly (Fig. 1, bottom). If a line selection button is currently engaged, pushing a second line button fully inwards makes a distinct three-part sequence of "engagement" sounds: (1)

---

[1] Although speech recognition is one means by which interaction might be accommodated, high error rates and temporal lag will make manual interaction with controls inevitable.

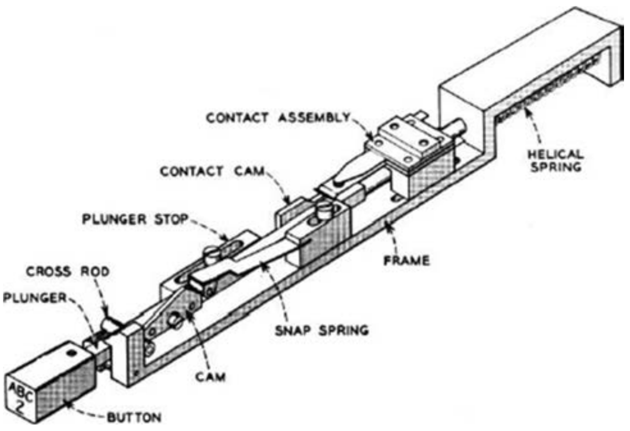


Fig. 2. The adjustable "universal push-button switch" (Deininger, [14]).

the release or "pop out" of the previously engaged line button, (2) the movement of the internal switch locking bar to engage the new switch at the end of the throw, and (3) the "click" sound caused by the release of pressure from the finger on the plastic button. The "hold" button does not lock when pushed inward; it has its own sound characteristic comprised of (1) a spring-like sound caused by contact with the release mechanism of the locking bar at the end of its throw and (2) a secondary switch release sound that occurs simultaneously with the sound of the currently engaged line button being released.

While no published research relating to the development of the line selection buttons has been found by the authors, Deininger [14] describes research into the relationship between force-displacement of telephone number keys and user speed, accuracy, and preference. Fig. 2 shows an adjustable "universal push-button switch" used at Bell Laboratories to determine optimal settings of the interface for user performance and preference. We virtually replicated the auditory cues that result from adjustments of the force-feedback components of this experimental switch via software control (ref. Appendix A).

The multi-line telephone is the user's interface to the communication system, allowing the selection and organization of multiple incoming calls. The selection among various input lines in a chosen time sequence can be quickly accomplished by the user without regard to the details of the communication system itself, such as the volume adjustment or frequency selection necessary on the flight deck radio. In auditory display parlance, these can be considered as "information streams"; and as applied to NextGen, as an array of stored synthetic speech data link messages.

In addition to the auditory, haptic, and visual cues for line selection, the telephone incorporates two different audio cues for incoming calls. Most often, the familiar classic telephone ring is used to direct visual attention to the telephone receiver so that the line associated with an incoming call can be answered. The audio alert is unspecific as to the incoming line (counting from left to right, buttons 2 – 5 in Fig. 1, top), requiring a subsequent action for first visually identifying the appropriate flashing button (button 2, 3, 4, or 5), and then engaging it by pushing the button inward (a "visually-guided haptic target"). A second "buzzer" sound indicates that a call has come internally via the intercom and is answered using the rightmost "priority line" button (button 6 in Fig. 1, top).

The user handles multiple incoming calls by placing some on "hold" while answering others. If a second incoming call arrives while the first line is engaged, the user presses the leftmost, red colored "hold" button (button 1 in Fig. 1, top), which then illuminates. A flashing light indicates an incoming call; a "winking" light (slower rate than flashing) indicates a held line.

1.3 Experimental Objectives

The first objective of this study was to compare performance and preference in the use of physical versus simulated touch panel controls within the context of a user-control button interface for avionic message selection. The second objective was to evaluate the contribution of auditory feedback from interaction with the controls. Performance was defined as the time required to activate a message in response to an aural alert ("ring"), and the duration a message was active. Preference was defined in terms of Likert scale ratings of "perceived performance" and "overall preference" for the interface type. A final objective was to examine the relationship between objective performance and perceived performance.

We hypothesized that there would be a significant effect on objective performance between physical and virtual switches. Physical switches can be tactilely located without use of the vision and provide richer tactile feedback when depressed, compared to a virtual switch. The "closed loop" feedback of the physical system was predicted to allow superior performance. Furthermore we hypothesized that the use of auditory feedback for both the physical and virtual switch would improve performance, since an additional perceptual modality is provided. Finally, we hypothesized that the objective measures of performance would correlate to subjective measures of preference.

2 METHOD

2.1 Experimental Design

A repeated measures design (within-participant comparisons for each condition tested) was used to evaluate two levels of switch type (physical pressure–sensitive switches with tactile feedback and flat panel virtual switches without tactile feedback) and two levels of auditory feedback from the switches (none, or auditory cues for touching and engaging the switch). Combinations of these levels resulted in four experimental conditions: physical switches with and without auditory feedback and touch panel switches with and without auditory feedback.

Each experimental block had 32 "normal" trials (calls on regular lines) and eight "priority" trials (calls on the priority line). There were three experimental blocks per condition, and each of the 10 participants was tested under all four conditions.

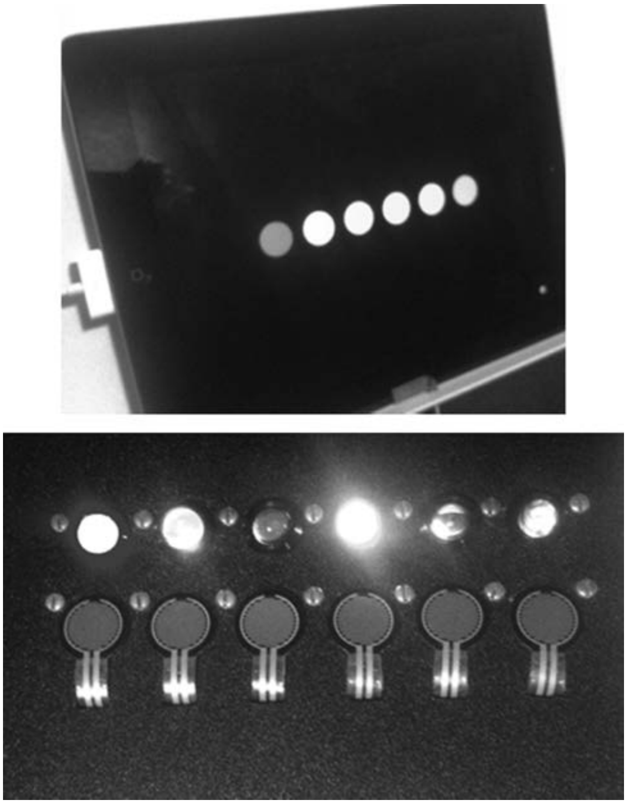


Fig. 3. Top: flat panel virtual buttons (iPad). Bottom: pressure sensitive strain gauges (Phidgets, Inc.).

2.2 Participants

Participants were recruited via the San Jose State University Research Foundation office at NASA Ames. Informed consent protocols were executed in line with the NASA Human Research Institutional Review Board. Participant ages ranged from 19 – 27 ($M = 22.5$, $SD = 2.1$). They were paid a nominal amount for their participation. All participants were screened for normal hearing via a questionnaire. None were employed as professional pilots or air traffic controllers.

2.3 Apparatus: Auditory Display Interface

The design principles of the multi-line telephone described above influenced the design of two equivalent interfaces used in this study. Each was intended as a rough prototype of a NextGen five-channel data link message playback system for on-demand audio. For the virtual button conditions, a flat panel display (Apple iPad II) was used with virtual 5/8-in. diameter circular buttons that responded to time and placement of the finger (Fig. 3, top). The flat panel communicated wirelessly with the experimental host computer via TCP/IP. For the physical button conditions, low-profile raised pressure sensors (5/8-in. diameter circular strain gauges, produced by Phidgets, Inc.) were mounted to a 6.5 × 4.5 × 1.5 in project box (Fig. 3, bottom). The strain gauge voltage output was digitally converted and transmitted via USB to the experimental host computer. The spacing of the buttons corresponded to that of the multi-line telephone. The flat panel display used illuminated buttons (red for hold; amber for line activated) while the physical buttons used LED lamps mounted 1 in. above the center. The shape, color, size, and spacing of the physical and virtual buttons were comparable.

Audio cues, developed from recordings of the Western Electric telephone switches, were activated using digital audio software (MAX/MSP 5.0; see Appendix A). Three distinct types of auditory cues were used in the aural feedback conditions: (1) a relatively low-level plastic clicking sound when touching (but not engaging) the buttons; (2) a two-part sound sequence corresponding to engaging the button switch; and (3) a two-part sound sequence corresponding to disengagement. The first part of the sound sequence was activated when exceeding a digital value caused by a moderate level of finger pressure (duration on the flat panel), and the second part responded to the button release action of the finger.

The edge of the raised profile of the strain gauges and their surface texture allowed the physical buttons to be found tactilely without visual cues. In the case of the virtual buttons, haptic feedback resulted only from the relatively stiff impedance of the entire screen surface and not from the edge or texture of any individual button; visual cues had to be used to locate a particular button. The answering logic of the model analog telephone was implemented into custom experimental control software named Button Audio Manager (BAM), ref. Appendix A. Logic patches of the software included means for assigning and buffering incoming messages, depending on the current state of the device.

2.4. Stimuli

Ten participants ran twelve blocks (as described in Sec. 2.1) lasting approximately 5 minutes each in a simplified flight simulation. They were instructed to accomplish the task of listening and responding to a continuous stream of incoming messages via prescribed procedures as quickly and as accurately as possible. “Normal” data link synthesized voice messages (DLM) “rang” on buttons 2 – 5, and “priority” messages “rang” on button 6. If a message arrived marked as priority, the participant was obligated to put other messages on hold (by selecting button 1) to audition it. In all experimental conditions, “Normal” messages were signaled by a brief buzzer alert sound and “priority” messages were signaled by a bell sound.

The twelve blocks were randomized between the four experimental conditions: virtual buttons, aural feedback; virtual buttons, no aural feedback; physical buttons, no aural feedback; physical buttons, aural feedback. There were a total of 96 “normal” and 24 “priority” messages per condition, each of which required a prescribed response. Message onset times were randomly varied to occur within an interval of 3 – 8 s. The duration of each message was approximately 5 s, in the form of <call sign> <flight number> <instruction>; for example “American 2 9 4, climb to 10,000 feet.” These “Normal” messages were meant to emulate current data link communications to the three aircraft

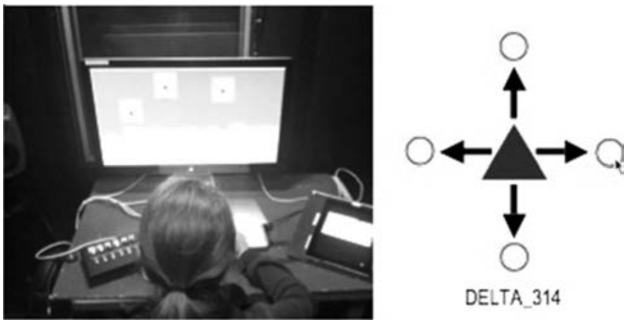


Fig. 4. *Left*: Participant running an experimental block in a soundproof experimental booth. The physical switch interface is visible on the left and was replaced with a touchscreen for the virtual switch experimental blocks. The right touchscreen presented a five-alternative forced choice question after a priority message was received. *Right*: close-up of “radar” symbol seen on the monitor. The participant clicked the circle in response to messages to “climb,” “descend,” “turn left” or “turn right.”

nearest to ownship (audio “twittering” [6]). “Priority” messages were communications for ownship.

The experiment was designed to motivate participants to listen to the content of the message while performing various tasks with a computer screen and a mouse, thereby distracting them from attending solely to the communication device. To do this, a set of procedures was given during a training session which explained how to respond to different messages and how to interact with a simplified “radar display,” as shown in Fig. 4. Participants were required to find the aircraft on the “radar” screen that corresponded to the aural DLM and click the circle corresponding to the instruction for that aircraft: climb, descend, turn left or turn right. The aircraft on the “radar” would change call sign and flight number intermittently to force the participant to continue looking up, rather than at the communication device.

Each time a priority message occurred, a question about its content was displayed on a second touchscreen and required a response within 6 s (five-alternative forced choice) to encourage the participants to answer the priority calls quickly. The responses to these content questions were not analyzed.

The dependent variables were chosen to investigate interaction with the proposed interface in terms of performance and subjective evaluation. The dependent variables involving performance (shown in Fig. 5) included the time to respond to the aural DLM alert (*t1*), time to complete audition of a single DLM (*t2*), time to respond to priority messages (*t3*), and time to complete audition of a single priority DLM (*t4*). Subjective data were gathered at the conclusion of the experiment using a questionnaire, which asked for comparative perceived self-performance under each of the four conditions, and for overall preference (hedonic rating). Data were gathered using both ratings (7 point scale) and rankings of the four conditions.

The dependent variable, “message response time,” was the interval between the time an auditory cue for an incoming message began to the time the switch was fully engaged to initiate playback of the message. The dependent variable “message completion time” was the time interval between when playback of a message began and when the participant released the message. Data for normal and priority messages were evaluated separately using a two-way analysis of variance (ANOVA) with display type and presence of an auditory cue as the independent variables.

3 RESULTS

3.1 Objective Performance

The analysis of the raw message response time data for “normal” messages indicated a significant effect for both audio cues ($F_{(1,9)} = 5.574$, $p = .043$) and for interface type ($F_{(1,9)} = 7.374$, $p = .024$). Stimuli using the physical switches were overall 1.5 s faster (2.2 s versus 3.7 s). Stimuli with supplemental audio cues were overall 0.5 s faster (2.7 s versus 3.2 s). There was no interaction between these variables. These data were analyzed in terms of the average value per participant and per condition. See Fig. 6.

Message response time data included a number of outliers (>3 SD), which were removed for each condition in subsequent analyses. The percentage of outliers was 0.94 – 1.77% between the four conditions, out of the total

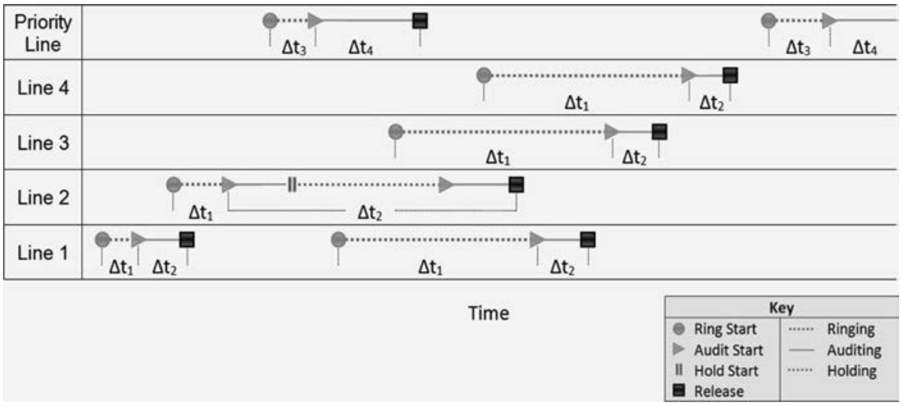


Fig. 5. Sample data link message “score” for an experimental run. Δt_1 = message response time; Δt_2 = message completion time. Δt_3 = priority message response time; Δt_4 = priority message completion time.

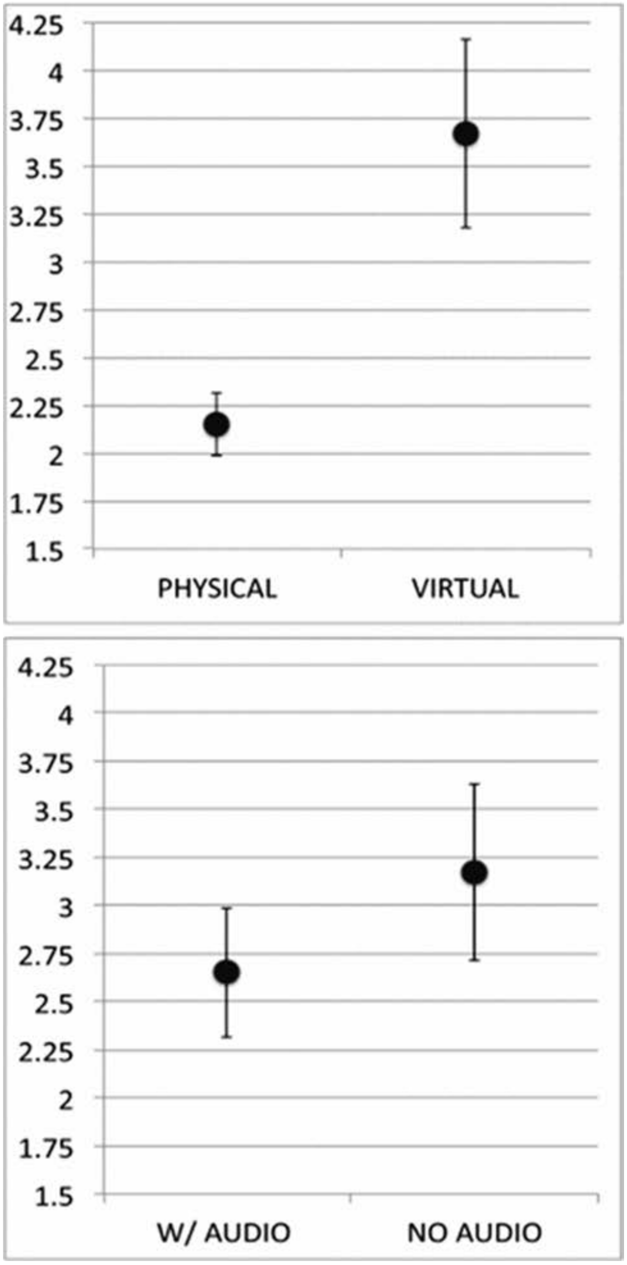


Fig. 6. Raw Timing data (s) for "normal" messages (mean and standard error).

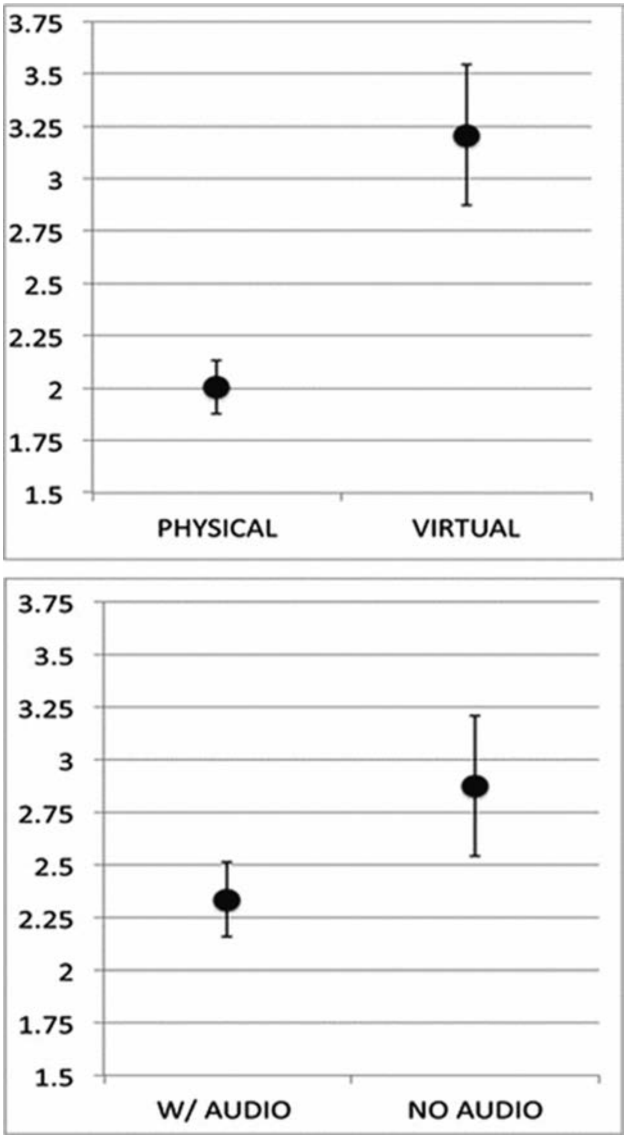


Fig. 7. Timing data (s) with outliers removed for "normal" messages (mean and standard error).

960 trials/condition. Analysis of the response timing data for "normal" messages with outliers removed indicated a significant effect for both audio cues ($F_{(1,9)} = 9.125$, $p = .014$) and for interface type ($F_{(1,9)} = 12.68$, $p = .006$). Stimuli using the physical switches were overall 1.2 s faster than the virtual switches (2.0 s versus 3.2 s). Stimuli with supplemental audio cues were overall 0.6 s faster than without supplemental audio cues (2.3 s versus 2.9 s). There was no interaction between these variables. See Fig. 7.

Response time data are typically log transformed as $log_{10}(1+X)$ to normalize the distribution of response data (correct for positive skew). An ANOVA of the transformed data confirmed the significant results of the untransformed data analysis. Analysis of the log-transformed data indicated a significant effect for both audio cues ($F_{(1,9)} = 6.429$, $p = .032$) and for interface type ($F_{(1,9)} = 16.795$, $p = .003$).

The raw values for "priority" message response time data indicated a total of 1.25 – 2.92% outliers (>3 standard deviations) within each of the four conditions, of the total of 240 trials per condition. Analysis of the timing data for "priority" messages with outliers removed indicated a significant effect for interface type ($F_{(1,9)} = 58.164$, $p < .001$). Stimuli using the physical switches were overall 0.6 s faster (2.1 s versus 2.7 s). There was no significant effect of the audio cue or interaction. An ANOVA of the log-transformed data confirmed the significant results of the untransformed data analysis, with a significant effect for interface type ($F_{(1,9)} = 29.213$, $p < .001$).

Message completion time data had a modest effect for interface type, with the physical switches 0.3 s faster (4.9 s versus 5.2 s): ($F_{(1,9)} = 11.318$, $p = .008$). There was no significant effect of the audio cue or interaction.

Table 1. Wilcoxon pairwise comparison for perceived performance ratings.

PERCEIVED PERFORMANCE RATING

	virtual-no audio	physical-no audio	virtual-w/audio	physical-w/audio	
	15	23	24.5	37.5	
virtual-no audio	15	8	9.5	**22.5**	**significant**
physical-no audio	23		4	14.5	
virtual-w/audio	24.5			13	
physical-audio	37.5				

PERCEIVED PERFORMANCE RANKING

	virtual-no audio	physical-no audio	virtual-w/audio	physical-w/audio	
	14	23	27	36	
virtual-no audio	14	9	13	**22**	**significant**
physical-no audio	23	0	1.5	13	
virtual-w/audio	27			9	
physical-audio	36				

The effect, while modest, could be cumulative and therefore more meaningful in the case of multiple messages.

3.2 Subjective Ratings

Participants provided ratings and rankings for perceived performance and preference for each of the four display conditions. Data were gathered using 7 point Likert scale in a survey administered immediately after the experiment. The use of ratings and rankings may be considered redundant but were included to insure consistency in responses.

A Wilcoxon test was used to analyze these non-parametric data separately for each of the four rating evaluations across participants. For the four treatments and 10 rank orders, the critical value is 14.8 for $p < = .05$, and 18 for $p < = .01$ [15]. Table 1 below indicates a significant difference ($p < = .01$) between the virtual switches without audio condition versus the physical switches with audio for perceived performance; Table 2 indicates similar results for preference. Other condition comparisons were only significant at $p < = .1$.

We also tested the association between our subjective and objective measures. A Spearman Rank Correlation coefficient was calculated by transforming the performance data into rank values (interval to ordinal conversion) and then comparing these ranks to the ranked opinion data.

Looking at each of the four conditions individually, the correlation between response time and preference rating was 0.72. The correlation between response time and perceived performance rating was 0.68. This indicates a moderately strong correlation between the quality metrics of perceived performance and preference, compared to the objective performance measure of response time (0.68 and 0.72 are near the value of 0.7 for the $p = .05$ significance level (using n-2 degrees of freedom). Similar values were found for response time to “priority” messages ($\rho = 0.62, 0.61$ for perceived preference and performance, respectively).

Additional analyses were conducted by averaging the response times for the two different audio conditions and then ranking according to interface condition, as well as by averaging the two different interface types and ranking according to audio condition. The justification is based on the significant effect found for the performance data for these main effects and the lack of interaction. The correlation between response time and both preference and perceived

Table 2. Wilcoxon pairwise comparison for preference ratings.

PREFERENCE RATING

	virtual-no audio	physical-no audio	virtual-w/audio	physical-w/audio	
	14.5	23	28	34.5	
virtual-no audio	14.5	8.5	13.5	**20**	**significant**
physical-no audio	23		5	11.5	
virtual-w/audio	28			6.5	
physical-audio	34.5				

PREFERENCE RANKING

	virtual-no audio	physical-no audio	virtual-w/audio	physical-w/audio	
	15	24	25	36	
virtual-no audio	15	9	10	**21**	**significant**
physical-no audio	24		1	12	
virtual-w/audio	25			11	
physical-audio	36				

performance ratings increases to 0.98 for normal messages when combining. For "priority" messages, the correlation between response time and both preference and perceived performance ratings increases to 0.96. The correlation between subjective and objective measures is therefore quite high when accounting for the main objective effect.

4 DISCUSSION

From the standpoint of interface design optimization, these results address NextGen prioritized concerns, as expressed in the National Research Council's "Decadal Survey of Civil Aeronautics" for "interfaces that ensure effective information sharing and coordination among ground-based and airborne human and machine agents," and "interfaces and procedures that support human operators in effective task and attention management." The results of the "trade study" aspect of this investigation indicated that physical switches were significantly superior to virtual switches in terms of response time to an *individual message* by about 0.6 – 1.5 s. Audio feedback provided an advantage of about 0.6 s. When dealing with *multiple messages* in a real world context under high workload, these individual time advantages may combine to ensure safer, more efficient operations that foster effective information sharing and information management.

The subjective results show a significant effect of preference and impression of superior performance for physical pressure sensitive switches having audio feedback, compared to touch panel virtual switches. The correlation between objective measures of performance and subjective ratings of preference and performance was shown to be high.

Overall, the results indicate that any replacement of physical controls by virtual touch screens in NextGen flight deck controls must be considered carefully, and should include audio feedback. Additionally, the use of a five-channel message storage system with an on-demand playback interface shows promise for enabling pilots to successfully manage a complex set of NextGen data link audio messages.

5. ACKNOWLEDGMENTS

This work was supported by NASA's System-wide Safety Assurance Technologies (SSAT) Project. We appreciate the assistance of our colleagues in the Advanced Controls and Displays Laboratory at NASA Ames' Human Systems Integration Division.

6 REFERENCES

[1] *National Research Council Decadel Survey of Civil Aeronautics: Foundation for the Future* (Washington, DC: National Academies Press). Retrieved from http://www.nap.edu/ (2006).

[2] S. A. Lozito, A. McGann, and K. Corker," Data Link Air Traffic Control and Flight Deck Environments: Experiments in Flight Crew PerformanceData Link Air Traffic Control and Flight Deck Environments: Experiments in Flight Crew Performance," *Proc. 7th International Symposium on Aviation Psychology*, pp. 1009–1015 (1993).

[3] D. R. Begault, "Guidelines for NextGen Auditory Displays," *J. Audio Eng. Soc.*, vol. 60, pp. 519–530 (2012 July/Aug.).

[4] A. J. Rehmann and R. H. Mogford, "Airborne Data Link Study Report," U.S. Federal Aviation Administration Technical Report DOT/FAA/CT-TN95/62 (1996).

[5] J. A. Lancaster and J. G. Casali, "Investigating Pilot Performance Using a Mixed-Modality Simualted Data Link," *Human Factors* vol. 50, pp. 183–193 (2008).

[6] R. W. Koteskey, S. Wu, V. Battiste, E. M. Wenzel, J. B. Lachter, D. R. Begault, J. H. Nguyen, M. S. Politowicz, and W. W. Johnson, "Enhanced Audio for NextGen Flight Deck," *Proc. 4th Int. Conference on Applied Human Factors and Ergonomics* (2012).

[7] A. R. Pritchett and R. Hansman, "Variations Among Pilots from Different Flight Operations in Party Line Information Requirements for Situational Awareness," *Air Traffic Control Quarterly* vol. 4, pp. 29–50 (1997).

[8] M. Godfroy, D. R. Begault, and E. M. Wenzel, "Speech Synthesis for Datalink: A Study of Overall Quality and Comprehension Effort," *Proc. 15th International Symposium on Aviation Psychology*, Dayton, OH, April 27–30 (2009).

[9] K. A. Latorella, "Effects of Modality on Interrupted Flight Deck Performance: Implications for Data Link," *Proc. Human Factors and Ergonomics Society Annual Meeting* vol. 42, pp. 87–91 (1998).

[10] N. Bilton, "F.A.A. Approves iPads in Cockpits, but Not for Passengers," New York Times. Retrieved from http://bits.blogs.nytimes.com (December 14, 2011).

[11] www.boeing.com (n.d.).

[12] S. Merchel, E. Altinsoy, and M. Stamm, "Touch the Sound: Audio-Driven Tactile Feedback for Audio Mixing Applications," *J. Audio Eng. Soc.*, vol. 60, pp. 47–53 (2012 Jan./Feb.).

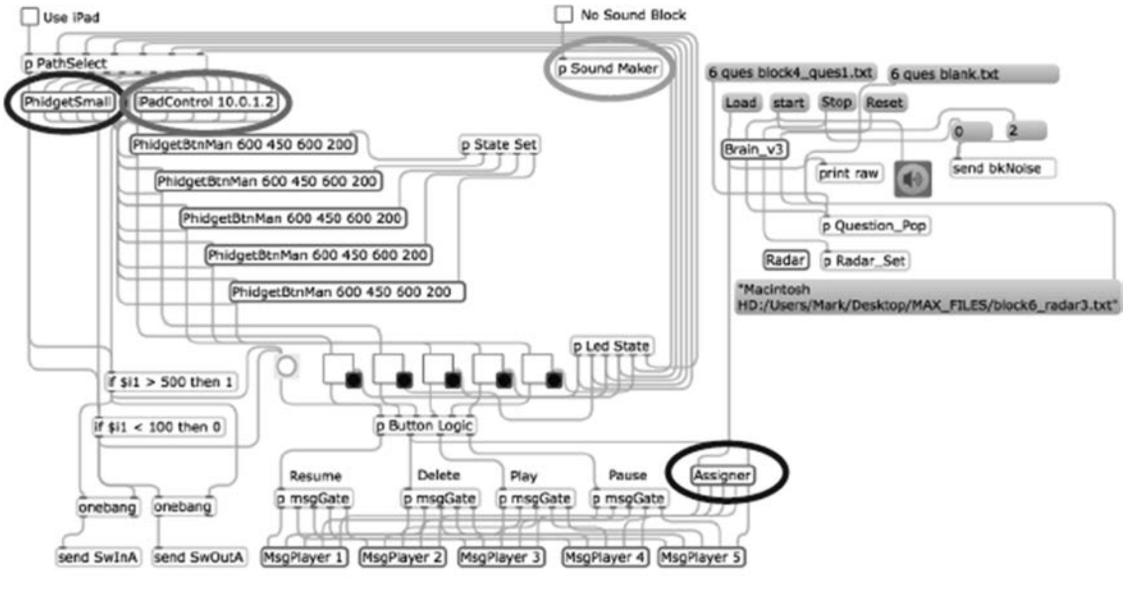
[13] E. Altinsoy, "The Quality of Auditory-Tactile Environments," *J. Audio Eng. Soc.*, vol. 60, pp. 38–46 (2012 Jan./Feb.).

[14] R. L. Deininger, "Human Factors Engineering Studies of the Design and Use of Pushbutton Telephone Sets," *Bell System Technical Journal* vol. 39, pp. 995–1012 (1960).

[15] L. Sachs, *Applied Statistics. A Handbook of Techniques* (Springer-Verlag, Berlin, 1984).

7. APPENDIX A: EXPERIMENT SOFTWARE OVERVIEW

The multi-line telephone used as a model for the study was simulated within the custom experimental control software BAM (Button Audio Manager). This appendix describes how the interface logic, visual indications, and auditory cues of the original telephone were synthesized using Cycling74's MAX/MSP version 5 software, including the logic for line selection, line hold, and "priority."



Primary Control Patch

Fig. A1. High-level MAX patch for BAM software.

The logic patches of the software include means for assigning and buffering incoming messages, depending on the current state of the device. Patches specific to the use of the touchscreen or button interface were selected depending on the experimental condition.

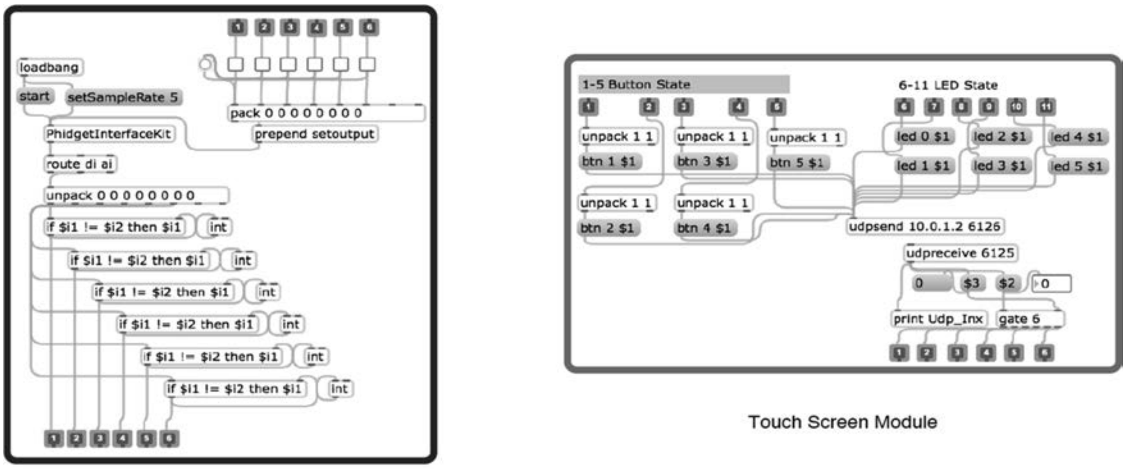
The BAM software is composed of a hierarchically arranged series of MAX patches that model the telephone's visual and auditory cues, run the experimental blocks, and collect the data for subsequent analyses. Fig. A.1 shows the highest level of the experimental control software. For each experimental block, the encapsulated object "Brain_v3" reads a script of time-ordered events that includes lists of audio messages, audio alerts, visual display changes, and user response questions that are displayed on a second iPad. All user actions with external interface hardware are recorded for subsequent analysis. Lower levels of the experimental control software are devoted to simulating the telephone: producing button sounds in reaction to touch, providing visual feedback in the form of flash rate and illumination, and indicating the state logic of the buttons (line engaged, disengaged, held, or pending).

Analysis of the original telephone indicates a complex set of "cause and effect" interactions that the software is required to simulate using the Phidget or iPad interface. For example, to "hold" line 1 and "answer" a new incoming call requires the software to accomplish the following tasks:

- The object "Assigner" (circled in the lower right of Fig. A.1) finds the next available line to assign to the incoming call (for example, if line 2 was already on hold, go to line 3). "Assigner" then sends a message to the object "Sound Maker" to activate ring tone (upper right circle, Fig. A.1) and to the object "Led State" to flash the light associated with the incoming line button.
- When the user pushes the hold button, the object "Button Logic":
  1) Activates the hold button audio and visual cue (red light),
  2) Pauses the message associated with the current activated line (via a command routed to the appropriate object "MsgPlayer"),
  3) Makes the line 1 light go to a rapid flash,
  4) Activates the audio cue corresponding to the release of the line button (activating the "onebang" command to the object "send SwInA" when exceeding a threshold of "500" from the hold button on the interface).
- When the user selects the incoming line button, the object "button logic" sends messages to the appropriate objects to activate the audio and visual cues associated with the button push (change the visual cue from a flashing light to a constant light; play back the new message).

The buttons of the Western Electric telephone were sampled for playback of five unique sounds: a relatively quiet "plastic rattle" corresponding to lightly touching the button; two sounds for button engage (corresponding to the application and release of finger force); and two sounds for button release. A constant stream of integers from 0–1000 is sent from each strain gauge button of the Phidget, depending on finger pressure. The iPad sends data based on duration of touch within the area of the button, scaled to 0–1000 via a custom application resident on the iPad. The integer values are used by the software to determine, for each button, when they are being touched, pushed to engage, or pushed to release.

The object "PhidgetBtnMan" in Fig. A.1 examines the state of each button. The first two arguments "600 450"



Phidget Patch

Touch Screen Module

Fig. A2. Details of the objects "PhidgetSmall" and "IpadControl."

refers to: when the finger pressure (or timing for the iPad) button exceeds 600, activate the first auditory cue (audio sample of application of finger force for button engagement), and activate the second auditory cue when the finger force is released from the button, corresponding to the threshold dropping from above 600 to below 450. The third and fourth arguments "600 200" activate the third and fourth auditory cues that correspond to the button release sounds. The equivalent data from the iPad corresponded to finger timing within a specific area. Fig. A.2 left and right detail the objects "PhidgetSmall" and "iPadControl" in the upper left of Fig. A.1. These objects control how messages are received and routed from the strain gauge device and the iPad, respectively. The outputs 1–6 are sent to the five objects shown in Fig. A.1 as "PhidgetBtnMan."

The final module in the MAX/MSP pipeline models the current state of the array of switches using multiple logical operators. Multi-line telephones allow only one line at a time to be audited, and engaging one line would "hang up" other lines unless they were first put on "hold" using a hold button. Because the participant may elect to handle the messages in any order, this module assigns the message to the lowest open line, or buffers the message if no line is available.

THE AUTHORS

Durand R. Begault

Rachel M. Bittner

Mark R. Anderson

Durand R. Begault Ph.D. (AES Fellow, 2002) works in the area of research and development of new audio and multi-modal technologies for aeronautic and space applications, including psychoacoustic research, human factors evaluation, sound quality, acoustical modeling, and communications engineering. He has been associated with the Human Systems Integration Division of NASA Ames Research Center, located at Moffett Field in the Silicon Valley, since 1988. He is a member of the Acoustical Society of America, the Audio Engineering Society, and the Institute of Noise Control Engineering. His activities at AES have included publications, organization of workshops and paper sessions, chairing of the Technical Committee on Perception and Subjective Evaluation of Audio Signals, membership on the AES *Journal's* review board since 1993, work as Associate Technical Editor of the *Journal*, and membership on the Board of Governors.

•

Rachel M. Bittner is currently enrolled in New York University's doctoral program in music technology with an emphasis in digital signal processing and music information retrieval. She received her M.S. in mathematics from New York University's Courant Institute in 2013 and has a B.S. in mathematics and a B.M. in music performance from U.C. Irvine. She has worked as a research associate in the Advanced Controls and Displays Laboratory at NASA Ames Research Center since 2011. She is a member of the American Mathematical Society, and has been a member of the AES since 2012.

•

Mark R. Anderson has been a software engineer in the Human Systems Integration Division at NASA Ames Research Center for the last 16 years. Prior to arriving at NASA, Mark applied his background in aeronautical engineering as both a maintenance base manager for American Eagle Airlines and as a technical support representative for Fairchild Aircraft. The skill sets practiced in private industry coupled with his B.S. in computer science from San Jose State University allow Mark to work across disciplines, both building physical hardware prototypes and providing the software that brings them to life.