Aircrew Performance and Safety While Using Protective Masks in Response to Coronavirus Disease

Kara M. Cave; Amanda M. Kelley; Kathryn A. Feltman; Jason A. Gerstner; Justin L. Stewart; John S. Crowley

INTRODUCTION: In response to the urgent need for safe aircrew respiratory protection due to the COVID-19 pandemic, three small descriptive evaluations were conducted with aircrew and air traffic controllers (ATC) that assessed the impact of mask use on safety and performance onboard rotary wing aircraft.

METHODS: A series of evaluations assessed aircrew performance using the 3M™ Model 1860 N95 respiratory protection mask, two aviation-specific cloth mask prototypes, and a commercial off-the-shelf aviation-specific cloth mask. The series of evaluations included different sets of subjects consisting of up to five Black Hawk helicopter aircrew members, air traffic control (ATC), and 12 CH-47 aircrew members. The Modified Rhyme Test was used to measure speech intelligibility and was administered in the UH-60 among crewmembers of the same aircraft, between pilots of different aircraft, and between the pilots and ATC. Measures of workload, usability, comfort, and pulse oximetry were also administered.

RESULTS: Results from the Modified Rhyme Test indicated that all subjects scored greater than 80% accuracy given the proper microphone positioning relative to the mask. With respect to workload, NASA-TLX total scores for the "perform radio communications" task was 50.83.

DISCUSSION: Despite an elevated perceived degree of workload on the communications flight task, results from the speech intelligibility test indicated that performance was maintained within the acceptable range as defined by MIL-STD-1474E, Design Criteria Standard Noise Limit. This abbreviated evaluation suggests that the face masks tested are safe for use by helicopter aircrew under the conditions tested.

KEYWORDS: COVID-19, safety, aircrew, speech intelligibility, workload, personal protective equipment, face mask.

Cave KM, Kelley AM, Feltman KA, Gerstner JA, Stewart JL, Crowley JS. *Aircrew performance and safety while using protective masks in response to coronavirus disease*. Aerosp Med Hum Perform. 2021; 92(4):274–280.

In response to the COVID-19 pandemic, a requirement for Army aircrew to use face-worn respiratory protection during infected patient transport as well as routine operations was introduced. However, the impact of facial coverings on aircrew performance, particularly the effects on communication, was unknown. As such, the U.S. Army Aeromedical Research Laboratory (USAARL) responded to multiple requests from different Army agencies to assess selected facial coverings on aircrew safety and performance. Results from a series of evaluations informed U.S. Army aviation leadership of performance changes associated with this personal protective equipment (PPE). Describing the performance effects enabled aviation leadership to render informed decisions regarding continuation of aviation operations amid infection control measures necessitated by the pandemic.

Communication within rotary wing aircraft presents considerable challenges to Army aircrew even without the possible challenges associated with face-worn respiratory protection. Ambient hazardous noise threatens auditory health and communication among the aircrew, with ambient noise levels requiring the use of double hearing protection.[2] The reduced signal-to-noise ratio inside the aircraft resulting from high ambient noise can lead to increased operational demands and listening effort for aircrew.[13] A review of accidents within the air Emergency Medical Services cited that degraded radio communication remained one of several contributing factors to

From the U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL, USA.

This manuscript was received for review in August 2020. It was accepted for publication in January 2021.

Address correspondence to: LTC Kara M. Cave, Ph.D., Warfighter Performance Group, U.S. Army Aeromedical Research Laboratory, 6901 Farrell Ave., Fort Rucker, AL 36362, USA; kara.m.cave.mil@mail.mil.

DOI: https://doi.org/10.3357/AMHP.5751.2021

accidents.[10] Given the already challenging communication environment, the effects of a physical barrier imposed by the required face mask on the speech transmission path warranted further evaluation.

Wearing respiratory protective masks presents other possible challenges to aircrew beyond those resulting from reduced auditory signal quality. Specifically, reduced auditory signal quality can yield secondary effects, such as increased workload and performance decrements on tasks requiring auditory input, or compound the effects of tasks already taxing performance. For example, increased workload, particularly in combination with reduced communication signal integrity (i.e., no electronic signal processing applied to the incoming communication), can degrade flight performance and speech intelligibility.[1] Similarly, when levels of signal-to-noise ratio are low, subjects' performance decreased on a recall task and visual reaction times.[11] Therefore, these results supported that increased listening effort and subsequent increased workload can deplete other cognitive resources that are unrelated to auditory detection. As such, use of a face mask that could potentially degrade performance among aircrew warranted further assessment of workload to accurately describe device-worn performance.

To address rapid-response inquiries from the Army aviation community, USAARL conducted assessments of speech intelligibility and workload on aircrew while wearing respiratory PPE. Additionally, other potentially important factors were examined; specifically, measures of usability, user experience, pulse oximetry, and comfort. The series of evaluations described herein were conducted with U.S. Army aircrew wearing an N95 respirator, Aviation Center Logistics Command prototype cloth masks, and an aviation-specific commercial-off-the-shelf (COTS) flame-retardant cloth mask. The series of evaluations sought to describe performance-related issues with Army helicopter aircrew wearing respiratory protective masks.

METHODS

Subjects

All subjects were Army aircrew and air traffic controllers (ATC) assigned to Fort Rucker, AL, USA, and deemed medically fit for their respective duties. In the first of three evaluations, a male HH60M pilot served as the talker. Listeners consisted of a female rear-stationed crewmember, a male HH60M copilot, and a male UH60A/L pilot. In the second evaluation, two HH60M pilots (one woman), one male ATC, and one female rear-stationed crewmember served as talkers. Three HH60M pilots (one woman who also served as a talker) and an ATC (the same one who served as a talker) served as listeners. In the third evaluation, 12 CH-47 male aircrew [4 instructor pilots (IPs), 4 flight students, and 4 flight engineers] completed qualitative assessments. Medical Evacuation (MEDEVAC) aircrew and an ATC completed speech intelligibility testing. The MEDEVAC aircrew also completed the same qualitative assessments as the CH-47 aircrew. The MEDEVAC aircrew was comprised of two HH60M pilots (both men), two JUH-60A+ pilots (one woman), a female rear-stationed crewmember, and a male ATC. Two HH60M pilots, a male JUH-60A+ pilot, and the rear-stationed aircrew member (flight surgeon) served as talkers. Both HH60M pilots, the male JUH-60A+ pilot, and ATC served as listeners.

Materials

The three mask evaluations described herein incorporated commonly used, validated measures of speech intelligibility, workload, comfort, and user experience. Speech intelligibility measurements complied with those outlined in MIL-STD-1474E,[9] Design Criteria Standard Noise Limit, using the Modified Rhyme Test (MRT).[7] The MRT is a multiple-choice test that consists of 50 monosyllabic words in 6 lists (300 words total). Each list is presented as 50 ensembles of 6 related words. The words in each ensemble all share a core vowel and either start or end with the same consonantal phoneme (e.g., tent, bent, went, lent). The listener then identifies which of the six words in the ensemble was transmitted by circling the word on the response sheet. In these evaluations, one crewmember read the 50-word list with the carrier phrase, "Say the word [MRT stimulus]" while the listener circled their choice via hardcopy score sheet.

The NASA Task Load Index (NASA-TLX)[6] was used to measure subjective workload. Using a 100-point scale, respondents rated mental demand, physical demand, temporal demand, performance, effort, and frustration subscales. Typically, respondents then provide assessments of which factors contributed to workload through pairwise comparisons. The ratings and weights are used to calculate weighted total scores. Only ratings were used in this evaluation. In order to assess comfort associated with use of the mask, the Wong-Baker Visual Analog Scale[14] was used to assess pain level [no pain (0) to worst pain possible (10)] and administered pre- and postflight. Two additional questionnaires were administered to assess user experience and comfort postflight: Usability Metric for User Experience-Lite (UMUX-Lite)[8] and an adapted version of Schumacher et al.'s comfort and experience questionnaire.[12] The UMUX-Lite consists of a response scale of strongly disagree (1) to strongly agree (7) that is used to respond to two statements: "This face mask's capabilities met my requirements" and "This face mask is easy to use." Finally, pulse oximetry was measured to evaluate changes in blood oxygen saturation associated with mask use and was collected using a Nonin (Plymouth, MN, USA) 7600 oximeter.

Procedure

This project was reviewed and locally determined by the U.S. Army Aeromedical Research Laboratory's Research Compliance Office to be a nonresearch project, and did not require IRB approval. In all evaluations, the MRT was administered during grounded flight operations just prior to takeoff. The NASA-TLX was administered to aircrew after performing 12 flight tasks and subsequent shutdown. The rear-stationed crewmembers completed the NASA-TLX for a subset of these tasks. However, only the "perform radio communications procedures" task is presented here given that communication was the

focus of the evaluation. All aircrew wore the HGU-56/P helmet and communications earplugs (CEP). All flights were performed around the vicinity of Fort Rucker, AL, USA.

The first evaluation conducted assessed the effects of respiratory protective masks (N95) on performance and safety of MEDEVAC aircrew in the HH60M and UH60A/L aircraft. Speech intelligibility was assessed between pilots within the same aircraft (HH60M), between the pilot and rear-stationed crewmember (HH60M), and between aircraft (HH60M to UH60A/L). Due to severe shortages in PPE at the time of this evaluation, a single 3M™ Model 1860 N95 respiratory protection mask was worn by one Army aviator while conducting flight operations in an HH60M Black Hawk helicopter. Of note, respiratory fit-check testing did not occur prior to flight. The pilot wearing the N95 mask served as the talker whereas the three other subjects (copilot, crewmember, and pilot of a separate aircraft) did not wear a mask and served as listeners. Three mask conditions were tested: the pilot wearing an N95 mask and maxillofacial shield, the pilot wearing an N95 mask only (no face shield), and the pilot wearing neither the maxillofacial shield nor the mask (control). The flight lasted approximately 1.5 h and following aircraft shutdown, the N95 wearer (talker) and two pilot listeners completed the NASA-TLX. The two pilot listeners reported only the workload associated with the "perform radio communications procedures" flight task, whereas the talker completed a NASA-TLX associated with all 12 flight tasks. The talker also completed the modified Wong-Baker Visual Analog Scale pre- and postflight. Postflight, the talker completed the UMUX-Lite and adapted comfort and experience questionnaire as well.

In the second evaluation, onboard aircrew performance was assessed while using two types of cloth prototype protective masks produced by the U.S. Army Aviation Center Logistics Command. The evaluation was conducted using two UH60M aircraft. The smaller cloth mask prototypes were worn by the pilot talker in the primary aircraft (Pilot A), the crewmember in the primary aircraft, and ATC, whereas the larger one was worn by the pilot talker of the second aircraft (Pilot C). The copilots in both aircraft (Pilots B and D) served as listeners. The maxillofacial shield routinely mounted on the crew chief flight helmet as crash protection for Army aircrew was incorporated in this study as a mask condition. Finally, the effect of the boom microphone position relative to the mask (inside the mask, outside the mask) on speech intelligibility was evaluated. There were 15 configurations systematically varying the role of talker, listener, use of the maxillofacial shield, and microphone position relative to the mask which were tested. The pilots and the rear-stationed crewmember wearing the mask prototypes completed the pre- and postflight measures in the same manner as in the first evaluation.

The third evaluation assessed onboard performance of CH-47 and UH-60 aircrew using a COTS cloth fire-retardant facemask. Speech intelligibility measures were collected only on the UH-60 aircrew. The evaluation involved a two-part evaluation. In part one, 12 CH-47 aircrew evaluated the mask over 3 d, with the aircrew each completing three flights. The IPs and students completed a baseline administration of the modified Wong-Baker Visual Analog Scale before the first flight. The flight engineers were unable to do so at this time (due to schedules) and, however not ideal, completed the assessment before the second flight occurred. After the third day of flying, the same postflight measures used in the previous 2 evaluations were completed by 11 crewmembers. One flight engineer filled out his postflight measures after the second day due to a projected absence on the third day. The MRT was not administered in this evaluation. Prior to donning the masks, a demonstration was provided of the four configurations of the mask: 1) straps under the helmet, microphone inside the mask; 2) straps under the helmet, microphone outside the mask; 3) straps over the helmet, microphone inside the mask; and 4) straps over the helmet, microphone outside the mask. Aircrew were advised to wear the mask in the configuration they found most comfortable, but that previous evaluations demonstrated that the microphone inside the mask yielded the best speech intelligibility results.

Part two of the third evaluation tested the same configurations and followed the same procedures as the second evaluation, with the addition of pulse oximetry and without use of the maxillofacial shield, to evaluate the COTS masks. The four pilots (Pilots A, B, C, and D) and a rear-stationed aircrew member (flight surgeon) donned the COTS mask. One ATC participated as a listener and wore a mask of his choosing.

Statistical Analysis

The data from all three evaluations for each outcome measure was summarized and descriptive statistics are presented. No inferential statistics were conducted due to the small number of observations and the variability between the evaluations in this operational evaluation.

RESULTS

Speech intelligibility was measured with the MRT and results were scored according to MIL-STD-1474E by determining the percent correct and using a correction factor to control for guessing. The results from the three mask evaluations are furnished in **Table I**. Overall, use of the N95 respirator and maxillofacial shield yielded results that exceeded the MIL-STD's 1474E acceptable speech intelligibility criterion value of 80%. For the cloth prototype and COTS masks evaluations, MRT scores were within acceptable limits with the microphone configured inside the mask, with the exception of the pilot to ATC condition with the cloth prototype.

Perceived workload was measured with the NASA-TLX and responses specific to the "perform radio communications" flight task are summarized in **Fig. 1** for the N95 and cloth mask prototype evaluations. As no specific criteria exist for what defines "high" and "low" workload with NASA-TLX scores, these values were compared to mean values from past studies where subjects performed similar tasks:[12,13] as can be seen in Fig. 1, the N95 talker's NASA-TLX scores were near or exceeded

Table I. MRT Scores for the Three Mask Evaluations.

TALKER	LISTENER	TALKER MASK TYPE	MICROPHONE CONFIGURATION	COMMUNICATION PARTNER LOCATION(S)	MRT % CORRECT
N95 Respirator					
Pilot 1	Copilot	None	not applicable (n/a)	same aircraft	92.8
Pilot 1	Crew	N95 + Shield	Outside mask	same	88.0
Pilot 1	Pilot 2	None	n/a	different	92.8
Pilot 1	Pilot 2	N95 + Shield	Outside mask	different	100
Pilot 1	Pilot 2	N95	Outside mask	different	90.4
Pilot 1	Copilot	N95+ Shield	Outside mask	same	97.6
Cloth Mask					
Pilot 3	Pilot 4	Small prototype	Inside mask	same	97.6
Pilot 3	Pilot 4	Small prototype	Outside mask	same	90.4
Pilot 5	Pilot 3	Large prototype + Shield	Inside mask	different	83.2
Pilot 5	Pilot 3	Large prototype + Shield	Outside mask	different	**73.6**
Pilot 5	Pilot 3	Shield	n/a	different	80.8
Pilot 5	Pilot 6	Large prototype + Shield	Inside mask	same	100
Pilot 5	Pilot 6	Large prototype + Shield	Outside mask	same	80.8
Pilot 5	Pilot 6	Shield	(n/a)	same	97.6
Pilot 3	ATC	Small prototype + Shield	Inside mask	UH60M to ATC	**78.4**
Pilot 3	ATC	Small prototype + Shield	Outside mask	UH60M to ATC	83.2
Pilot 3	ATC	Shield	n/a	UH60M to ATC	95.2
ATC	Pilot 3	Small prototype	n/a	ATC to UH60M	90.4
ATC	Pilot 3	Small prototype	n/a	ATC to UH60M	97.6
ATC	Pilot 3	Small prototype	n/a	ATC to UH60M	92.8
Crew	Pilot 3	Small prototype	Inside mask	same	88.0
Crew	Pilot 3	Small prototype	Outside mask	same	**76.0**
Crew	Pilot 3	Small prototype + Shield	n/a	same	**76.0**
COTS Mask					
Pilot 7	Pilot 8	COTS	Inside mask	same	100
Pilot 7	Pilot 8	COTS	Outside mask	same	90.4
Pilot 7	Pilot 8	Shield	n/a	same	100
Pilot 9	Pilot 7	COTS	Inside mask	different	90.4
Pilot 9	Pilot 7	COTS	Outside mask	different	**56.8**
Pilot 9	Pilot 7	Shield	n/a	different	88.0
Pilot 10	Pilot 9	COTS	Inside mask	same	85.6
Pilot 10	Pilot 9	COTS	Outside mask	same	**71.2**
Pilot 10	Pilot 9	Shield	n/a	same	88.0
Pilot 9	ATC	COTS	Inside mask	different	80.8
Pilot 9	ATC	COTS	Outside mask	different	**59.2**
Pilot 9	ATC	Shield	n/a	different	85.6
Crew	Pilot 7	COTS	Inside mask	same	92.8
Crew	Pilot 7	COTS	Outside mask	same	**66.4**
Crew	Pilot 7	Shield	n/a	same	95.2

MRT: Modified Rhyme Test; ATC: air traffic controller; COTS: commercial off-the-shelf. Conditions incorporating the maxillofacial shield are referred to as "shield." Results that fell below acceptable standards are in bold.

the mean "high" workload scores used for comparison. **Fig. 2** presents the responses for the COTS cloth mask evaluation for each aircrew group and airframe. Workload ratings reported by nonstudent pilots and instructor pilots were "low," whereas ratings provided by flight engineers and students were primarily "high" in comparison. Overall, the N95 talker reported the greatest degree of workload compared to the other conditions. The cloth mask prototypes demonstrated the second highest workload rating compared to all other conditions, with the exception of the frustration subscale recorded during the COTS study.

Results from the Modified Wong-Baker Visual Analog scale varied across the evaluations. For the N95 evaluation, the pilot wearing the mask indicated no discomfort preflight and moderate discomfort postflight. In the cloth mask prototype evaluation, this same measure showed disparate results among the pilots and crewmembers. Specifically, the pilot wearing the small mask reported no discomfort and no change in rating from pre- to postflight, whereas the pilot wearing the large mask reported moderate discomfort pre- and postflight with an increase in discomfort by 1 point on the scale. The crewmember wearing the small mask reported moderate discomfort for preflight which decreased to a mild level postflight (difference in 3 points on the scale). For the COTS mask evaluation, the CH-47 aircrew preflight comfort ratings were on average considered mild, but increased slightly by 1 point postflight, remaining within the mild rating range. The UH-60 aircrew ratings of the COTS mask indicated no discomfort, with only one respondent providing a rating greater than zero, but postflight ratings decreased to zero.

The Comfort and Experience Questionnaire results are provided in **Table II**. Ratings suggest that the cloth COTS and

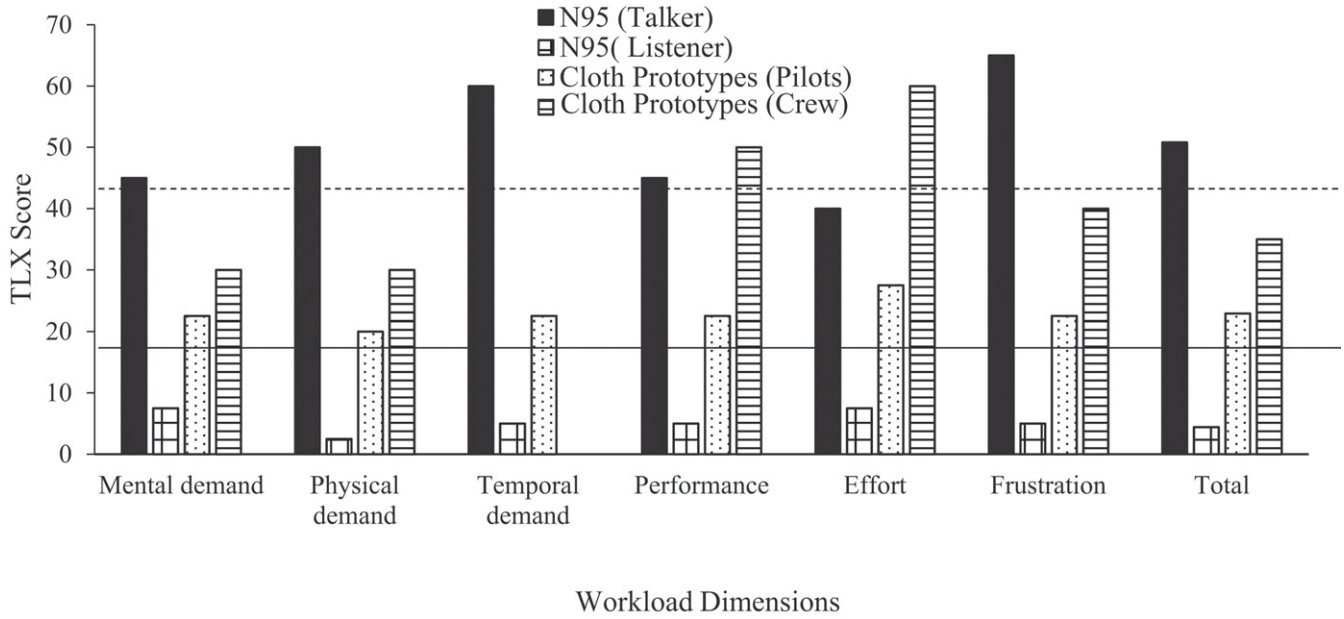

Fig. 1. NASA-TLX results for the N95 and cloth prototype mask conditions and aircraft. The dashed and solid lines indicate a comparison workload value reported in Feltman et al.[3] for high and low workload, respectively. The mean value for the two listeners who were unmasked in the N95 condition is reported.

prototype masks resulted in lower heat ratings than the N95 mask. However, the N95 mask yielded the least limitations to the field of view. Speech intelligibility ratings were best for the COTS mask compared to the other mask types evaluated.

Results from the UMUX-Lite showed that the pilot wearing the N95 mask in the first evaluation agreed with the statement that the mask met his requirements and was neutral on the statement that the mask was easy to use. In the cloth prototype mask evaluation, responses varied among the aircrew such that the pilots who served as talkers agreed that the mask met their requirements, whereas the crewmember reported disagreement with the statement (M = 5.50, SD = 0.71). Pilot A, wearing the small mask, responded that the mask was easy to use, whereas Pilot C, wearing the large mask, reported a neutral response, and the crewmember (small mask) disagreed with this statement (M = 4.50, SD = 0.71). In the third evaluation

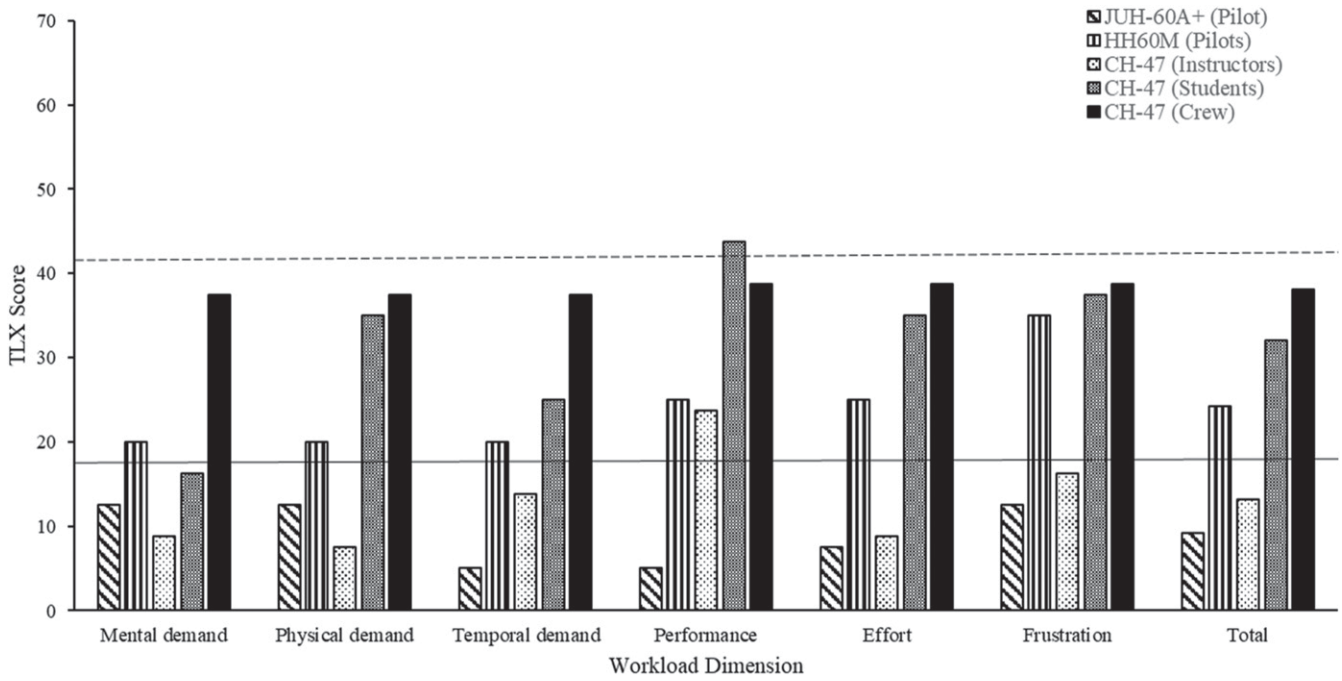

Fig. 2. NASA-TLX results for the COTS mask condition according to aircraft and crewmember type. "Crew" refers to the flight engineers. The dashed and solid lines indicate a comparison workload value reported in Feltman et al.[3] for high (dashed horizontal line) and low (solid horizontal line) workload, respectively.

Table II. Comfort and Experience Questionnaire Responses on a Scale of 0 (Worst) to 5 (Best).

	MOBILITY	HEAT, OUTSIDE TEMPERATURE IN °C	VISION: FIELD OF VIEW	VISION: OBSCURATION	SPEECH INTELLIGIBILITY
N95	5.00	1.00, 34	5.00	3.00	2.00
Cloth prototype	3.00 (1.00)	1.67 (0.58), 30	4.00 (0.00)	3.67 (1.53)	2.67 (0.58)
COTS CH-47 Aircrew	4.00 (1.21)	2.42 (1.37), 27	3.33 (1.15)	2.50 (1.31)	3.00 (1.48)
COTS UH-60 Aircrew	4.60 (0.55)	3.40 (0.89), 24	3.80 (1.30)	3.80 (1.30)	4.00 (0.71)

COTS: commercial off-the-shelf. Standard deviations are in parentheses with the exception of the N95 evaluation due to only one evaluator.

(COTS cloth mask), the CH-47 aircrew tended to agree that the cloth mask met their requirements (M = 5.00, SD = 1.60) and that the mask was "easy to use" (M = 5.33, SD = 1.07). Similarly, the UH-60 aircrew agreed that the cloth mask met their requirements (M = 6.20, SD = 0.84) and was easy to use (M = 6.20, SD = 1.30).

Finally, in the cloth mask prototype evaluation, blood oxygen saturation levels were within normal limits (97% ± 2).[4] Specifically, the preflight to in-flight levels in Pilot A (98–97%), Pilot B (98–97%), and the rear-stationed crewmember (100–99%) decreased by 1%, but this degree of decrease is not considered clinically significant. Note that oxygen saturation is generally not affected by altitudes less than 4921 ft (1500 m), and the aircrew remained below this altitude.

DISCUSSION

The key finding across all three evaluations is that all mask types allowed for appropriate communications within aircraft, between aircrafts, and between the aircraft and ATC. However, to achieve communications that met current defined standards, as measured by the MRT and defined by MIL-STD-1474E, the microphone required placement adjustments dependent upon mask type. Specifically, the N95 3M™ 1860 mask allowed the clearest communications compared to the other two masks. Given this N95 model does not have a respirator valve device, the pilot was able to press and maintain the microphone boom closer to the mouth without apparent compromise to the mask seal. Similarly, pilots reported that the COTS cloth mask did not inhibit communications between crewmembers or ATCs when the microphone was inside the mask. The pilot was able to press and maintain the microphone boom close enough to the mouth without apparent compromise to communications.

Workload during the "perform radio communications" task was highest with the N95 mask compared to that reported for the other mask types. Furthermore, workload ratings reported by the talker were considered high in comparison to those provided in a study including a similar communication function,[3] and in a recent meta-analysis of NASA-TLX ratings which compiled data reported for 152 pilot/aircraft studies.[5] Alternatively, the listeners during radio communications procedures rated overall workload during this task as low. The disparity in workload between the talker and listeners in the first evaluation may be due to the pilot wearing the mask having to repeat himself multiple times. Of note, with the exception of the student pilots in the third evaluation, the aircrew performed routine tasks in a familiar airspace. Therefore, the effects of varied flight conditions on workload given the wear of the various masks remains unknown.

Regarding comfort and usability measures, wear of the N95 and prototype masks indicated some challenges with speech intelligibility and heat. However, the COTS mask ratings did not reveal adverse ratings on these attributes and wear of these masks was associated with the highest, or best, ratings for heat stress and speech intelligibility. Of note, the N95 mask showed the lowest (worst) rating for heat stress, but the ambient temperature was highest during that evaluation. The N95 was also associated with comfort issues and described as moderately uncomfortable postflight. The cloth prototypes received disparate comfort ratings among users, and the COTS mask was associated with mild discomfort on average.

This study was an operational test conducted under great time pressure, as guidance for aircrew was urgently needed by Army aviation leaders. In order to provide guidance, certain aspects of a typical well-controlled evaluation were sacrificed. Among these were sample size, selection, and range. Also, as the tests progressed, methods were evolved, making comparisons across mask types difficult. Finally, these tests did not attempt to evaluate the effectiveness of any of the protective mask variants in preventing the spread of coronavirus.

Although some discomfort and degradation to speech intelligibility were noted, overall, all three mask types resulted in speech intelligibility well within current defined standards given proper microphone configuration relative to the mask. Additionally, while workload increased during the radio communications tasks, the TLX scores reported were considered within acceptable ranges. This abbreviated evaluation suggests that the face masks tested are safe for use by helicopter aircrew under the conditions tested.

ACKNOWLEDGMENTS

The authors would like to acknowledge the aircrew who assisted the research team in the writing of this document: MAJ Audrey Boenker, COL Susan Fondy, CPT Lance Randles, CW4 Christopher O'Brien, Mr. Russell Schmidt, and CW2 Rosa Trujillo.

The opinions, interpretations, conclusions, and recommendations contained in this report are those of the authors and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citation of trade names in this report does not constitute an official Department of the Army endorsement or approval of the use of such commercial items.

Financial Disclosure Statement: The authors have no competing interests to declare.

Authors and affiliation: Kara M. Cave, M.S., Ph.D., Amanda M. Kelley, M.S., Ph.D., Kathryn A. Feltman, M.A., Ph.D., Jason A. Gerstner, B.S., Justin L. Stewart, B.S., and John S. Crowley, M.P.H., M.D., U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL, USA.

REFERENCES

1. Casto KL, Casali JG. Effects of headset, flight workload, hearing ability, and communications message quality on pilot performance. Hum Factors. 2013; 55(3):486–498.
2. Department of the Army. Department of the Army Pamphlet 40-501: Army Hearing Program. Washington (DC): Headquaters Department of the Army; 2015.
3. Feltman KA, Kelley AM, Bernhardt K, Britt T, Matthews C. Real-time operator state monitoring during periods of high and low workload across four workload domains. USAARL Report No. 2019-12. Fort Rucker (AL): U.S. Army Aeromedical Research Laboratory; 2019.
4. Goldberg S, Buhbut E, Mimouni FB, Joseph L, Picard E. Effect of moderate elevation above sea level on blood oxygen saturation in healthy young adults. Respiration. 2012; 84(3):207–211.
5. Grier RA. How high is high? A meta-analysis of NASA-TLX global workload scores. Proc Hum Factors Ergon Soc Annu Meet. 2015; 59(1):1727–1731.
6. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Advances in Psychology. 1988; 52:139–183.
7. House AS, Williams C, Hecker MH, Kryter KD. Psychoacoustic speech tests: a modified rhyme test. J Acoust Soc Am. 1963; 35(11):1899.
8. Lewis JR, Utesch BS, Maher DE. Investigating the correspondence between UMUX-LITE and SUS scores. In: Marcus A, editor. Design, User Experience, and Usability: Design Discourse. Cham (Switzerland): Springer International Publishing; 2015:204–211.
9. MIL-STD-1427E. Design criteria standard noise limit. Washington (DC): U.S. Department of Defense; 2015. [Accessed 2020 July 23]. Available from https://www.arl.army.mil/wp-content/uploads/2019/12/ahaah-MIL-STD-1474E-Final-15Apr2015.pdf.
10. National Transportation Safety Board. Event summary: aviation special investigation report—Emergency Medical Services. 2006 Jan. 25. [Accessed 2019 Sept. 2]. Available from https://www.ntsb.gov/news/events/Pages/Aviation_Special_Investigation_Report_Emergency_Medical_Services.aspx.
11. Sarampalis A, Kalluri S, Edwards B, Hafter E. Objective measures of listening effort: effects of background noise and noise reduction. J Speech Lang Hear Res. 2009; 52(5):1230–1240.
12. Schumacher J, Arlidge J, Garnham F, Ahmad I. A randomised cross-over simulation study comparing the impact of chemical, biological, radiological or nuclear substance personal protection equipment on the performance of advanced life support interventions. Anaesthesia. 2017; 72(5):592–597.
13. Webb CM, Gaydos SJ, Estrada A, Milam LS. Toward an operational definition of workload: a workload assessment of aviation maneuvers. USAARL Report No. 2010-15. Fort Rucker (AL): U.S. Army Aeromedical Research Lab; 2010. [Accessed January 2021]. Available from https://apps.dtic.mil/sti/pdfs/ADA528341.pdf.
14. Wong DL, Baker CM. Pain in children: comparison of assessment scales. Pediatr Nurs. 1988; 14(1):9–17.