Defence Research and Development Canada | Recherche et développement pour la défense Canada

DRDC | RDDC
technologyssciencetechnologie

Stereoacuity Distribution for Royal Canadian Air Force (RCAF) Aircrew

Mackenzie Glaholt
Paul Kim
DRDC – Toronto Research Centre

Defence Research and Development Canada
Scientific Report
DRDC-RDDC-2019-R117
August 2019

Canada

IMPORTANT INFORMATIVE STATEMENTS

Abstract

Clinical measurement of stereoacuity was added to the RCAF biannual medical for aircrew in 2012, but at this time there is no standard or cut-off for stereoacuity applied to the RCAF aircrew occupations. In contrast, several allied air forces, such as the United States Air Force (USAF) and the Royal Air Force (RAF), apply a stereoacuity standard. The primary goal of the present study was to characterize the stereoacuity distribution of RCAF aircrew, and to determine the potential impact of various possible minimum stereoacuity standards given the present RCAF population. Stereoacuity measurements were extracted from a subset of the RCAF aircrew medical records and distributional analyses were conducted. We estimated the effect of applying the USAF and RAF stereoacuity standards, as well as other medically-motivated cutoff standards. Our analysis indicates that the adoption of the USAF and USN stereoacuity standard would likely affect large proportions of the RCAF population, while the RAF standard would have a much more modest impact, as might be expected given that these standards are relatively conservative and liberal, respectively. Given these potentially large effects on the RCAF population, the motivation for a stereoacuity standard must be carefully considered. While there is converging evidence that stereopsis is relevant to certain aviation sub-tasks, there is little direct evidence to support a specific stereoacuity cut-off for aircrew. This situation tends to support a liberal cut-off standard that ensures functioning stereopsis in aircrew. However, further research must be conducted to establish a direct link between stereoacuity and operational performance and determine whether or not a more conservative stereoacuity requirement should be in place.

Significance to Defence and Security

This research provides information that supports aeromedical decision making for the RCAF. Stereopsis research is a CAF Surgeon General research priority, and this project was endorsed by the Surgeon General Health Research Program.

Résumé

En 2012, la mesure clinique de l’acuité visuelle stéréoscopique a été ajoutée à l’examen médical bisannuel du personnel navigant de l’ARC, mais à présent, aucune norme ou aucun seuil ne s’applique pour l’acuité visuelle stéréoscopique des groupes professionnels du personnel navigant de l’ARC. Cependant, plusieurs forces aériennes alliées, comme la United States Air Force (USAF) et la Royal Air Force (RAF), appliquent une norme d’acuité visuelle stéréoscopique. L’objectif principal de la présente étude visait à caractériser la répartition de l’acuité visuelle stéréoscopique parmi le personnel naviguant de l’ARC et de déterminer l’incidence potentielle de diverses normes minimales possibles de l’acuité visuelle stéréoscopique de la population actuelle de l’ARC. Les mesures d’acuité visuelle stéréoscopique ont été extraites d’un sous-ensemble de dossiers médicaux du personnel navigant de l’ARC et des analyses de répartition ont été menées. Nous avons estimé l’effet de l’application des normes d’acuité visuelle stéréoscopique de la USAF et de la RAF, ainsi que d’autres normes de seuil qui sont liées à une condition médicale. Notre analyse indique que l’adoption de la norme de l’acuité visuelle stéréoscopique de la USAF et de la USN aurait des répercussions sur un grand pourcentage de la population de l’ARC, tandis que la norme de la RAF aurait une incidence beaucoup plus modeste, comme nous pourrions nous y attendre puisque ces normes sont relativement prudentes et libérales, respectivement. Étant donné que ces répercussions pourraient être importantes sur la population de l’ARC, il faut examiner attentivement ce qui motive l’adoption d’une norme d’acuité visuelle stéréoscopique. Bien qu’il existe des données convergentes sur la pertinence de la stéréopsie pour certaines tâches secondaires de l’aviation, il existe peu d’éléments de preuves directs pour appuyer un seuil particulier d’acuité visuelle stéréoscopique pour le personnel navigant. Cette situation peut appuyer un seuil libéral qui assure une bonne stéréopsie du personnel navigant. Toutefois, d’autres recherches doivent être menées pour établir un lien direct entre l’acuité visuelle stéréoscopique et le rendement opérationnel et pour déterminer si une exigence plus prudente en matière d’acuité visuelle stéréoscopique devrait être en place ou non.

Importance pour la défense et la sécurité

Cette recherche permet de recueillir des renseignements qui appuient la prise de décisions dans le domaine aéromédical au sein de l’ARC. La recherche en stéréopsie est une priorité pour le médecin général des FAC. Le Programme de recherche en santé du médecin général approuve ce projet.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Major Erin Smith contributed to the study design and Captain Kathryn Brett helped coordinate the data collection.

1 Introduction

Stereopsis is the human perceptual ability to discern depth information based on binocular visual cues and it depends upon the perceptual fusion of retinal images from the two eyes. Accordingly, monocular individuals are not capable of stereopsis and it is also compromised in certain eye diseases such as strabismus and amblyopia (Ehrlich, Reinecke, & Simmons, 1983; Marsh, Rawlings, & Mumma, 1980), where poor oculomotor alignment interferes with perceptual fusion of the two retinal images. Stereopsis has long been suspected to be relevant to aviation (Wilmer, 1919), and though there is a modest body of empirical evidence to support this (Deas, Allison, Hartle, Irving, Glaholt, & Wilcox, 2017; Hartle, Sudhama, Deas, Allison, Irving, Glaholt, & Wilcox, 2019; Grosslight, Harold, Fletcher, Masterton, & Hagen, 1978; Jongbloed, 1935; Lewis & Krier, 1969; Pfaffman, 1948; Winterbottom, Gaska, Wright, Lloyd, Gao, Tey, & McIntire, 2014; Winterbottom, Lloyd, Gaska, Williams, Shoda, & Hadley, 2017; for a review see Wright, Gooch, & Hadley, 2013), it is still controversial and no study to date has presented a strong empirical case to link operational performance to a specific level of stereoacuity.

One of the sources of uncertainty about the importance of stereopsis to aviation is that humans are known to be capable of perceiving depth based on monocular cues alone, suggesting that binocular depth perception might be redundant. In particular, depth can be inferred based on monocular visual cues such as size, parallax, texture gradients, and shading (for a review see Howard, 2012). However, stereopsis might be immune to certain illusions of monocular cues (e.g., shading can be an ambiguous monocular cue, apparent size of objects can be misleading), and can still function even when certain monocular cues are absent (e.g., no familiar objects in scene to infer relative size and depth) and therefore stereopsis might provide more reliable depth information under certain operational conditions. Even when monocular cues are available and reliable, the processing of stereoscopic depth cues might serve to enhance depth perception performance. Stereopsis is based on the disparity in the retinal images of objects across the two eyes. Objects in the near field (relative to the focal point) show greater disparity than those in the far field, and at long ranges objects' disparity diminishes towards zero. For this reason, stereospsis is particularly useful for perceiving the depth of objects that are relatively close to the observer. Nevertheless, stereopsis has been shown to contribute to depth judgments for objects over 100m from the observer (see Palmisano, Gillam, Govan, Allison, & Harris, 2010). Accordingly, the application stereoscopic depth perception in aviation is likely to be limited to certain subtasks that involve nearby objects or terrain, such as taxiing, take-off and landing, formation flying, aerial refueling, rotary wing low hover, etc. Indeed, recent work has shown that stereopsis might be particularly useful for depth estimation during rotary wing low hover (Deas et al., 2017; Hartle, et al., 2019) and under some circumstances might allow for a degree of performance that cannot be achieved with monocular cues alone.

Stereoacuity measures also have the potential to serve as an indicator of whether or not a person can use certain display technologies. For example, on certain aerial-refueling tankers the refueling boom is operated by way of a stereoscopic display that depends upon the user's ability to perceive depth using stereoscopic cues (Lloyd, 2012; Lloyd & Nigus, 2012). Measures of stereoacuity might also be useful for determining whether individuals will be able to use certain head-mounted display technologies. For example, binocular displays that require binocular image fusion impose a degree of image alignment error, and persons with weak binocular fusion might have difficulty using these devices. Therefore, stereoacuity screening might be useful for detecting binocular diseases that might preclude the use of advanced displays that depend upon stereoscopic fusion of displayed imagery.

Militaries around the world differ as to whether or not they impose a stereopsis standard on aircrew. For example, the Royal Canadian Air Force (RCAF) does not have a stereoacuity standard, while the United States Air Force (USAF), United States Navy (USN), and Royal Air Force (RAF) do. Moreover, the stereoacuity cut-off for the USAF, USN, and RAF are quite different (25 arcsecs, 40 arscecs, and 120 arcsecs respectively). Note that stereoacuity values describe the smallest disparity that is perceivable and accordingly smaller values are "better." There are only a few studies that have examined the general population distribution of stereopsis (Bosten, Goodbourn, Lawrance-Owen, Bargary, Hogg, & Mollon, 2015; Coutant & Westheimer, 1993; Piano, Tidbury, O'Connor, 2016). Across these studies, the median population stereoacuity is between 20 and 90 arcseconds, with the variation depending on the population and stereoacuity tests used. Aircrew are generally selected along a variety of individual differences variables, and vision variables in particular (e.g., near perfect acuity, lack of eye diseases) and accordingly the distribution of aircrew stereoacuity is expected to be different from the general population. The differences in the USAF and RAF standards might be interpreted to reflect two objectives in screening for stereoacuity in aircrew. In particular, the USAF standard might be set to select for individual with near-perfect stereoacuity, consistent with the visual acuity standard requiring perfect or near-perfect acuity in aviators. The RAF standard, on the other hand, might be set to select for individuals who possess functioning stereoacuity, and would exclude individuals with defective stereoacuity due to eye disease.

The goal of the present study was to measure the RCAF aircrew stereopsis distribution. This distribution is presently unknown, and its characteristics would be particularly relevant if a stereopsis standard is considered for RCAF aircrew. Accordingly, we sampled a random subset of aircrew medical records and extracted their stereoacuity data. The data were analyzed as a population distribution and were therefore anonymized. We estimated the distributional parameters for all aircrew, as well as for subsets of flying aircrew (e.g., pilots, flight engineers). We then computed the number of aircrew that would be excluded from each category based on five different hypothetical stereoacuity standards.

2 Method

2.1 Procedure

The ethics protocol for the study (2016-009 Amendment 1) was reviewed and approved both by the DRDC Human Research Ethics Committee (HREC) and the Canadian Forces Health Services (CFHS) Chief Privacy Officer. In particular, the ethics protocol describes the waiver of informed consent for individuals to participate in the study: due to the large number of subjects whose data will be considered in this study it was impractical to obtain informed consent from each member of the subject population. Furthermore, because the goal of the study is to obtain the distribution of stereopsis values for the RCAF aircrew population, it is desirable to access all records and thereby avoid any biases that might occur if participation was obtained through solicitation. For this reason we obtained permission from the HREC and the CFHS Chief Privacy Officer to waive the collection of informed consent. We obtained a list of RCAF members' service numbers with the following aircrew Military Occupation Structure Identification Description (MOSID): 00183 (Pilot), 00182 (Air Combat Systems Officer; ACSO), 00019 (Airborne Electronic Sensor Operator; AESOp), 00021 (Flight Engineer; FE), 00170 (Loadmaster; LM), 00101 (Search and Rescue Technician; SARTech).

The health records of selected service members were accessed and collected by Canadian Armed Forces (CAF) medical personnel in accordance with CFHS Group Instruction 5020-30 Personal Health Information: Access, Use, and Disclosure. The principle risk to study subjects from this study would be in the event of a privacy breach. To mitigate this risk, the data were collected from Canadian Forces Health Information System (CFHIS) on a workstation located within the clinical consult services, and placed on a USB stored at the end of each day in accordance with Protected B guidelines. In total, three records (Records 1, 2, 3) were created for the data collection. The Record 1 consisted of the name, service numbers (SN), and MOSID of the Regular Force CAF aircrews whose health records were accessed. Record 2 is a copy of Record 1, but with a random subject code added to each entry to anonymize the individuals. Finally, for each entry in Record 2 that was sampled, the subject code was added to a third record (Record 3) along with the stereoacuity datum from the individual's medical record. Note that the sequence of sampling of Record 2 was randomized such that the aircrew population was randomly sampled. Records 1 and 2 were stored according to the guidelines for storage of Protected B information. Record 3 contained the anonymized data to be analyzed, and was stored according to the guidelines for storage of Protected A information.

Due to time constraints on data collection, we were only able to access a subset of 3103 subjects from the total population of 4466 RCAF aircrew members with these MOSIDs. The order in which the records in Record 1 were accessed was randomized to ensure random sampling of individuals across the aircrew population. Of the 3103 subjects whose records were accessed, 1466 (47%) were found to have stereoacuity data in their medical records. Because the selection of subjects with respect to MOSID was random, and consequently for some of the MOSIDS with smaller total numbers, additional data collection was required in order to estimate population parameters with sufficient power (see Table 1). For this additional sampling within MOSID the order of records accessed was also randomized.

Table 1: *Aircrew MOSIDs for which stereoacuity data were extracted.*

MOSID	Description	Total Population Number	Number Sampled	Number sampled that contained Stereoacuity data (% of total)
00183	Pilot	2271	1249	719 (32%)
00021	FE	364	352	206 (57%)
00170	LM	710	557	174 (25%)
00101	SARTech	191	179	67 (35%)
00182	ACSO	680	532	212 (31%)
00019	AESOp	250	234	88 (35%)

3 Results

In addition to analyzing the aircrew stereoacuity data as a whole, we also considered three sub-groups (see Table 2). Group 1 contained only the pilots (00183). Pilots are the subset of aircrew that have the most stringent vision standards for visual acuity and oculomotor health, and accordingly they might be expected to have a different distribution of stereoacuity from the other MOSIDs. Furthermore, it is possible that a stereoacuity standard might be applied only to that MOSID. Group 2 contained the FE, SARTech, and LM MOSIDs (00021, 00170, 00101), which are flying aircrew that are required to perform visual tasks both within and outside the aircraft, and accordingly stereopsis might also be relevant to those occupations. Group 3 contained the ACSO and AESOp MOSIDS (00182, 00019). These two occupational categories are primarily required to operate computer terminals on flying aircraft, and accordingly are least likely to have an occupational requirement for stereoacuity. However, if future aircraft incorporate stereoscopic displays, an operational requirement for stereopsis in Group 3 aircrew might emerge.

Our analysis was divided into two sections. The first section describes the distribution of stereoacuity values for each aircrew group. The second section investigates the application of various stereoacuity cut-offs to the three aircrew population groups. Data were manipulated and plotted in Matlab and statistical analyses were conducted in SPSS 23. For the purpose of our analysis we considered records with no stereoacuity information to be cases in which stereoacuity was not recorded. The reason for the substantial number of missing values is unknown, but is possibly related to a variety of factors such as the relative novelty of the stereoacuity measurement component to the aircrew medical procedure, the fact that there is no stereoacuity standard, the bi-annual scheduling of aircrew medical evaluations, and the availability of stereopsis testing equipment during aircrew medicals. Without information to the contrary, we assume that these missing values are randomly distributed, with individuals of different levels of stereoacuity being equally likely to have missing stereoacuity values in their aircrew medical record.

3.1 Stereoacuity Distribution

To characterize the stereoacuity distribution for the three groups, we computed the minimum and maximum values, the mean, median, and mode, the standard deviation, inter-quartile range, and estimates of skewness and kurtosis. These values are displayed in Table 2. We also produced histograms with 10 arcsec bins in order to show the distributions visually (see Figure 1). As can be seen in the figure, the distributions appears to have multiple modes, which likely reflect the discrete stereoacuity settings on booklet tests. All three distributions presented substantial skewness and kurtosis that is evident in the histograms: they are right skewed with a thin tail composed a small proportion of individuals with poor stereoacuity.

As can be seen in Table 2, Group 1 (pilots) had the lowest min, max, and mean stereoacuity, which is consistent with that group of individuals being subject to stringent selection based on visual acuity and oculomotor variables. Group 3 was very similar to Group 1 but with slightly higher min, max, and mean values. Group 2 had the largest mean stereoacuity value of the three groups. We conducted a series of pairwise Kolmogorov-Smirnov tests in order to determine whether the distributions differed. This analysis revealed that Group 1 had a significantly different distribution from Group 2 (K-S $Z = 2.16$, $p < 0.001$), and Group 2 differed significantly from Group 3 (K-S $Z = 3.08$, $p < 0.01$), but Group 1 and 3 did not differ significantly from one another (K-S $Z < 1$). As is apparent in the figure, Group 2 had a smaller proportion of individuals in the 20–29 arcsecond bin and a higher proportion of individuals in the 40–49 arcsecond bin compared to Groups 1 and 3. However all three groups appeared to have similar weight in the right tail of the distribution. Therefore, Group 2 would appear to have a smaller proportion of individuals with near-maximum stereoacuity and a greater proportion of individuals with moderately good stereoacuity than Groups 1 and 3.

Table 2: *Descriptive statistics for stereoacuity distribution for all subjects and for each group.*

Measure	All Subjects	Group 1	Group 2	Group 3
n	1466	719	447	300
Min (arcsec)	2.5	2.5	12	10
Max (arcsec)	800	550	800	800
Mean (arcsec)	41.18	38.2	46.2	40.84
Median (arcsec)	40	40	40	40
Mode (arcsec)	40	40	40	40
Standard Deviation (arcsec)	42.40	31.05	53.12	47.18
Inter-quartile Range (arcsec)	30–40	25–40	40–40	30–40
Skewness	11.77	11.19	9.32	14.13
Kurtosis	173.86	155.15	107.61	226.02

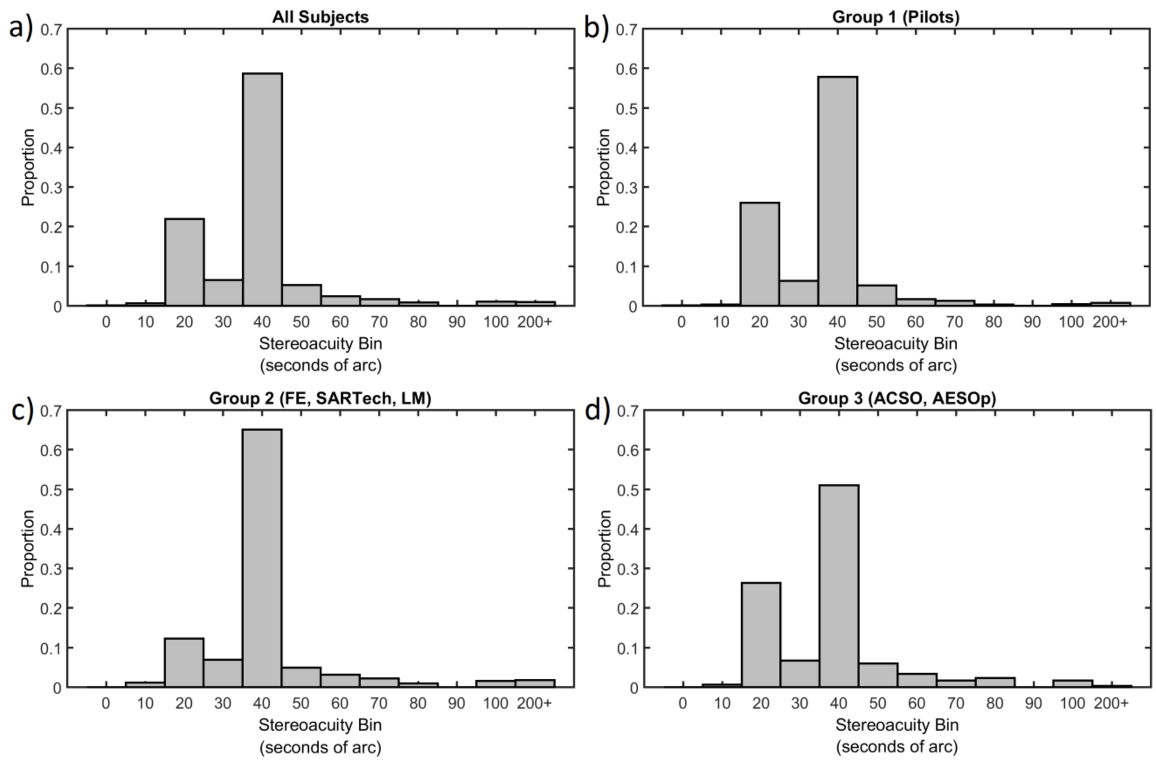

Figure 1: *Histograms of stereoacuity for all subjects (panel a), Group 1 (panel b), Group 2 (panel c), and Group 3 (panel d). The bins include values greater than or equal to the label value and less than the label value of the next bin to the right.*

3.2 Stereoacuity Cut-Off Criteria

We considered the effect of applying various stereoacuity cut-off criteria to the RCAF aircrew population. More specifically, we computed the proportion of RCAF aircrew in each group that would be excluded by each criterion. Because we did not sample the entire RCAF aircrew distribution (owing to time constraints and missing data), we had to estimate the number of aircrew that would have a stereoacuity greater than the cut-off value and thus be excluded. Accordingly, we applied the following computation of confidence intervals when estimating binomial probability for a finite population:

$$p \pm z_{\alpha/2} * \sqrt{\frac{p(1-p)}{n} * \frac{(N-n)}{(N-1)}}$$

where *p* is the sample probability, $z_{\alpha/2}$ is the *t*-distribution value for half the desired alpha (in this case alpha = 0.05, and therefore z = 1.96), *n* is the sample size, and *N* is the population size.

We considered the three aforementioned stereoacuity standards that are used by the USAF (25 arcsecs), the USN (40 arcsecs) and the RAF (120 arcsecs). In addition, we considered two medically-motivated criteria (100 and 200 arcsecs). The 100 arcsec criterion is based on the clinical intuition that individuals with stereacuity of 100 arcseconds or greater are suspected of having an oculomotor disease or disorder.

The 200 arcsecond criterion is considered here as a more liberal medical cut-off that is designed to detect the presence of stereopsis and eliminate individuals that are stereo-blind. In principle, one could apply a larger value to achieve this goal (e.g., 400 arcsecs), but given that subjects might guess the correct answer on the booklet tests that are used to measure stereoacuity, it is advisable to use lower stereoacuity value and thereby require them to answer more than one item correctly to pass (200 arcsecs is often the second test item following 400 arcsecs).

Table 3: *Analysis of proportions of each subject group that would be eliminated by each stereoacuity cut-off criterion. The confidence interval on the number of individuals excluded was rounded to the nearest whole number.*

Group	Criterion (arcsecs)	Proportion Excluded	Proportion Excluded 95% CI	Number Excluded	Number Excluded 95% CI
All Subjects	USAF (25)	0.774	0.757–0.792	3458	3379–3536
	USN (40)	0.123	0.109–0.137	548	487610
	Conservative Medical (100)	0.014	0.009–0.019	64	42–86
	RAF (120)	0.014	0.009–0.019	64	42–86
	Liberal Medical (200)	0.006	0.003–0.009	27	13–42
Group 1 (Pilots)	USAF (25)	0.736	0.709–0.762	1671	1610–1731
	USN (40)	0.095	0.077–0.112	215	175–255
	Conservative Medical (100)	0.010	0.004–0.016	22	9–36
	RAF (120)	0.010	0.004–0.016	22	9–36
	Liberal Medical (200)	0.004	0–0.008	9	1–18
Group 2 (FE, SARTech, LM)	USAF (25)	0.867	0.840–0.891	1095	1063–1127
	USN (40)	0.148	0.121–0.174	187	153–220
	Conservative Medical (100)	0.025	0.013–0.036	31	17–46

Group	Criterion (arcsecs)	Proportion Excluded	Proportion Excluded 95% CI	Number Excluded	Number Excluded 95% CI
	RAF (120)	0.025	0.013–0.036	31	17–46
	Liberal Medical (200)	0.011	0.003–0.019	14	4–24
Group 3 (ACSO, AESOp)	USAF (25)	0.730	0.689–0.771	679	640–717
	USN (40)	0.153	0.120–0.187	143	111–174
	Conservative Medical (100)	0.010	0.001–0.019	9	1–18
	RAF (120)	0.010	0.001–0.019	9	1–18
	Liberal Medical (200)	0.003	0–0.009	3	0–8

4 Discussion

In the present study we investigated the distribution of stereoacuity values for RCAF aircrew, and we also considered the effect of applying various stereoacuity standards for this population. Our analysis of the RCAF stereoacuity distribution produced parameter estimates that were consistent with prior studies. For example, our observed median of 40 arcseconds fell in the middle of the range of median values reported in prior work. Our population also tended to exhibit less variability than the data of Bosten et al. (2015). This might be due to the selection of aircrew according to vision health, though the mean, standard deviation, and inter-quartile range have been shown to vary substantially between tests (Piano et al., 2016). The value of detailed comparison with prior population distributions is limited by the fact that we do not know the test or tests used in RCAF aircrew medical evaluation. The test used for stereoacuity measurement was very rarely recorded in the aircrew medical records as there was no explicit field for this information on the examination chart. We assume that a booklet test was most often used (e.g., the Randot or the Butterfly test) given the prevalence of whole-number stereoacuity values that are found in the records. However, across all records the minimum stereoacuity recorded was 2.5 arcseconds, indicating that in some cases the medical assessment used tests other than the aforementioned booklets, which typically have a minimum stereoacuity setting of 20 (Randot) or 40 (Butterfly). We recommend that aircrew medical testing procedures be updated such that the medical chart contains the identity of the stereoacuity test used in addition to the stereoacuity score.

We were able to document some differences between the aircrew sub-groups. In particular, Group 1 (Pilots) and Group 3 (ACSO, Aesop) appeared to have the same distribution of stereoacuity values, while Group 2 (FE, SARTech, LM) differed. In particular, Group 2 tended to have a greater number of individuals with slightly poorer stereoacuity than Groups 1 and 3. This difference in the distributions was most notable in the proportion of individuals falling in the 20 and 40 arcsecond bins, with Group 2 exhibiting a shift in weight towards the latter bin when compared to Groups 1 and 3. Because the identity of the stereoacuity tests used for stereoacuity measurement is not known, it remains possible that this is the result of Group 2 individuals being tested more often on a booklet test with a minimum of 40 arcseconds. However, given that aircrew from each of these groups are co-located, it seems likely that they would be tested on the same aeromedical test apparatus. Another reason for the group differences might be because aircrew selection for Group 2 trades is less stringent with regards to vision variables than Groups 1 and 3, or that Group 2 trades tend to attract a slightly different population of individuals. If the observed difference or Group 2 is real, it is interesting to consider given that stereopsis is likely to be a relevant faculty for visual estimation of altitude in rotary wing operations (Deas et al., 2017; Hartle et al., 2019), which is a task that FEs and SARTechs engage in. However, it also is important to note that a stereoacuity of 40 arcseconds is not considered clinically abnormal, and furthermore there is no research to show that aircrew in Group 2 trades (or even other aircrew) would require a stereoacuity better than 40 arcseconds. However, if an operational requirement for a specific and stringent level of stereoacuity is shown, the distribution of Group 2 stereoacuity values should be reconsidered in this context.

Our analysis of potential cut-off criteria indicated that the USAF and USN standards would eliminate a large proportion of RCAF aircrew. The USAF criterion in particular was estimated to eliminate 77% of RCAF aircrew. One caveat to consider here is that the estimate of the number excluded based on a 25 arcsecond criterion could be artificially high if the RCAF aircrew are most often tested with a test that has a minimum greater than 25 arcseconds (e.g., Butterfly test with 40 arcseconds as a minimum score).

However we did find that 23% of RCAF aircrew met the 25 arcsecond criterion and thus it would appear that a test with a possible outcome of 25 arcseconds or lower is used at least some testing locations. Therefore, bearing this caveat in mind and based on the data that are available, it appears that a 25 arcsecond criterion would eliminate a substantial proportion of the RCAF population. It should also be noted, however, that the USAF will waive the 25 arcsecond stereoacuity standard and accept candidates with up to 120 arcseconds if there is no evidence of excessive phoria or microtropia. The 120 arcsecond criterion, by comparison, would eliminate a relatively small 1–2% of RCAF aircrew. The fact that the USAF allows waivers to the 25 arcsecond stereoacuity criterion raises additional uncertainty as to the exact operational requirement for stereoacuity in aircrew. Furthermore, at the time of this writing the USAF is apparently considering a change to a 40 arcsecond criterion (M. Winterbottom, Air Force Research Laboratory, personal communication). In contrast to the USAF's waiver system, the USN stereoacuity standard is applied firmly within that service, and if applied within the RCAF this criterion would eliminate 12% of aircrew. Accordingly, it would be important to derive clear evidence for the operational requirement for 40 arcseconds of stereoacuity if one were to exclude such a large proportion of candidates during selection. In our data the conservative medical cutoff (100 arcseconds) eliminated an identical proportion of aircrew to the RAF criterion (120 arcseconds). This is because in our data set there were no 120 arcsecond values. The 120 arcsecond stereoacuity value is present on some tests (e.g., TNO test, AO Vectorgraph test used by USAF) but not others (e.g., the Randot and Butterfly tests), and it is possible that none of the tests used during RCAF aircrew stereoacuity measurement had this setting. In general these two criteria are expected to eliminate a very similar proportion of any population. We also considered the effect of a 200 arcsecond criterion which was conceived as a liberal medical criterion (based on the typical booklet test settings) that would eliminate individuals without functioning stereopsis. This criterion was found to eliminate a very small proportion of RCAF aircrew (<1%).

At the present time it would appear that stereopsis is important for at least certain aspects of aviation, but the exact level of stereoacuity that is required is unknown. Accordingly, the adoption of a medically-motivated stereoacuity cut-off (e.g., 100–200 arcseconds) would tend to eliminate individuals that are stereoblind and who therefore would likely be at a disadvantage for certain aircrew tasks. More generally, the adoption of such a criterion would provide a screening measure during the medical component of aircrew selection where individuals with abnormal stereopsis as a result of oculomotor disorders could be excluded without having to undergo a complete evaluation, thereby improving the efficiency of aircrew selection. Prior to the adoption of a more stringent and specific stereoacuity standard, the operational requirements for higher stereoacuity should be demonstrated.

One of the challenges to implementing a stereoacuity standard is that there is no agreed upon gold standard for measurement of stereoacuity. Firstly, it is known that different tests have different stereoacuity distributions (see Piano et al., 2016), and therefore a specific stereoacuity cut-off must be tied to a specific test (i.e., 25 arcseconds on one test is not necessarily the same requirement as 25 arcseconds on another). Secondly, the booklet tests that are typically used might not be suitable for robust measurement of stereoacuity thresholds. In a prior study (Glaholt, Spivak, & Sacripanti, 2017) we found that popular booklet tests used to measure stereoacuity clinically correlate poorly with stereoacuity thresholds obtained from computer-based stereoacuity thresholding procedures. This suggests that these booklets might not be suitable for fine discriminations between individuals (e.g., deciding whether a person has 25 or 40 arcseconds of stereoacuity). Further research is required in this field to develop and validate a robust measure of threshold stereoacuity that can be used to make fine discriminations between individuals.

Rather than obtaining precise thresholds, the existing clinical tests of stereoacuity might be best used to detect the presence of stereopsis (i.e., as opposed to stereoblindness) and as such might be suited to implement a liberal stereoacuity criterion (e.g., 100, 200 arcseconds). However, it must be noted that for typical clinical stereo tests (e.g., Randot, Butterfly), the maximum threshold tested is often 400 arcseconds. Because there is a certain probability of guessing the correct answer on these tests (e.g., 25% on the Titmus, 33% on the Randot), there must be a number of stereoacuity settings greater than the desired threshold in order to ensure that the candidate does not pass the criterion by guessing. For example, for a test that is four-alternative and has a 100, 200 and 400 arcsecond setting, the chance of guessing all three answers correctly and thereby passing a 100 arcsecond criterion is 0.25^3, or 0.015. Another concern with the booklet-based stereo tests is that some of the high-disparity settings (e.g., 200, 400 arcseconds) have salient monocular cues that allow the correct answer to be inferred without binocular perception or stereopsis. For example, the Titmus stimulus (in the Butterfly test) consists of four circular stimuli arranged in a diamond shape. Each of the circles has an inner circle, and one of them has binocular disparity such that it appears, when viewed through the polarized lenses that accompany the test, to sink into the booklet while the other three have no disparity and appear flat on the page. At the highest disparity setting, the circle with disparity appears blurry when viewed binocularly without the polarized lenses, and it appears as off-centre when viewed through the polarized lenses with one eye closed. These monocular cues are less of a problem for tests that use stimuli composed of random dots. Therefore, while certain booklet tests might be adapted for use in implementing a liberal medical cutoff criterion in aircrew selection, we recommend that ultimately a robust stereoacuity thresholding test for aircrew should be developed.

5 Conclusion

The present study investigated the stereoacuity distribution of the RCAF aircrew population and estimated the impact of various potential stereoacuity cut-off criteria. The RCAF aircrew population was found to have a stereoacuity distribution that is consistent with prior civilian population studies. In addition, we found that Pilots, ACSOs, and AESOps tended to have slightly better stereoacuity than the FEs, SARTechs, and LMs. When considering the effect of various potential stereoacuity cut-off criteria, we found that the USAF and USN stereoacuity standards would eliminate a substantial proportion of the RCAF aircrew population. Accordingly, prior to adopting such criteria it would be prudent to establish, through additional research, the operational requirement for such stringent standards for aircrew. In contrast, we found that both the conservative medical criterion (and RAF criterion) as well as the liberal medical criterion eliminated a very small proportion of RCAF aircrew. Given that the available scientific evidence supports the relevance of stereopsis to certain aircrew tasks, the adoption of one of these relatively liberal criteria could be justified at present. This would also increase the likelihood that all aircrew would be able to use future display technology that has a stereopsis requirement. Moreover, a stereoacuity standard could help to expedite aircrew medical selection because it could eliminate certain cases of oculomotor disorder without the need for in-depth examination. Finally, we recommend that during RCAF aircrew vision evaluations, the identity of the test used to assess stereopsis should be recorded on the medical chart.

References

Ehrlich, M. I., Reinecke, R. D., & Simons, K. (1983). Preschool vision screening for amblyopia and strabismus. Programs, methods, guidelines. *Survey of Ophthalmology*, 28, 145–163.

Hartle, B., Sudhama, A., Deas, L. M., Allison, R. S., Irving, E. L., Glaholt, M. G., & Wilcox, L. M. (2019). Contributions of stereopsis and aviation experience to rotary wing altitude estimation. *Human Factors*. DOI: 10.1177/0018720819853479.

Howard, I. P. (2012). Perceiving in Depth, Volume 1. New York: Oxford University Press.

Glaholt, M. G., Spivak, T., & Sacripanti, B. (2017). Evaluation of clinical stereopsis tests for use in aircrew vision assessment. Defence Research and Development Canada, Scientific Report, DRDC-RDDC-2017-R032.

Grosslight, J. H., Fletcher, H. J., Masterton, B., & Hagen, R. (1978). Monocular vision and landing performance in general aviation pilots: Cyclops revisited. *Human Factors*, 20(1), 27–33.

Jongbloed, J. (1935). Landing carried out by experienced aviators with the use of one eye only. Acta Brevia Neerland, 5, 123–125.Karlsberg, R. C., Karlsberg, F. S., & Rubin, M. (1971). Aerospace Physiological Optics: 1. Depth Perception. Aerospace Medicine, 1080–1085.

Lewis Jr., C. E., & Krier, G. E. (1969). Flight research program. XIV. Landing performance in jet aircraft after the loss of binocular vision. *Aerospace Medicine*, 40(9), 957.

Lloyd, C. J. (2012). On the utility of stereoscopic displays for simulation training. Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2012, 1–9.

Lloyd, C. J., & Nigus, S. G. (2012). Effects of stereopsis, collimation, and head tracking on air refueling boom operator performance. Proceedings of IMAGE 2012 Conference, 1–11.

Marsh, W. R., Rawlings, S. C., & Mumma, J. V. (1980). Evaluation of clinical stereoacuity tests. *Ophthalmology*, 87, 1265–1272.

Pfaffman, C. (1948). Aircraft landings without binocular cues: A study based upon observations made in flight. *The American Journal of Psychology*, 61(3), 323–334.

Wilmer, W. H. (1919). The eye and aviation. *Arch. Neurol. Psychiatry,* 1(2), 162–166.

Winterbottom, M., Gaska, J., Wright, S., Lloyd, C., Gao, H., Tey, F., & Mcintire, J. (2014). Operational Based Vision Assessment Research: Depth Perception. *Journal of the Australasian Society of Aerospace Medicine*, 9, 33–41.

Winterbottom, M., Lloyd, C., Gaska, J., Williams, L., Shoda, E., & Hadley, S. (2017). Investigating aircrew depth perception standards using a stereoscopic simulation environment. *Stereoscopic Displays and Applications XXVIII*. Proceedings of the IS&T International Symposium on Electronic Imaging, 29–40.

Wright, S., Gooch, J. M., & Hadley, S. (2013). The role of stereopsis in aviation: Literature review. Technical Report, Air Force Research Laboratory: Wright Patterson AFB, OH, USA.

List of Symbols/Abbreviations/Acronyms/Initialisms

AESOp	Airborne Electronic Sensor Operator
ASCO	Air Combat Systems Officer
CAF	Canadian Armed Forces
CFHIS	Canadian Forces Health Information System
CFHS	Canadian Forces Health Services
FE	Flight Engineer
HREC	Human Research Ethics Committee
LM	Load Master
MOSID	Military Occupational Structure Identification Description
RAF	Royal Air Force
RCAF	Royal Canadian Air Force
SARTech	Search and Rescue Technician
USAF	United States Air Force
USN	United States Navy

DOCUMENT CONTROL DATA

*Security markings for the title, authors, abstract and keywords must be entered when the document is sensitive

1. ORIGINATOR (Name and address of the organization preparing the document. A DRDC Centre sponsoring a contractor's report, or tasking agency, is entered in Section 8.) DRDC – Toronto Research Centre Defence Research and Development Canada 1133 Sheppard Avenue West Toronto, Ontario M3K 2C9 Canada	2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.) CAN UNCLASSIFIED 2b. CONTROLLED GOODS NON-CONTROLLED GOODS DMC A

3. TITLE (The document title and sub-title as indicated on the title page.)

Stereoacuity Distribution for Royal Canadian Air Force (RCAF) Aircrew

4. AUTHORS (Last name, followed by initials – ranks, titles, etc., not to be used)

Glaholt, M.; Kim, P.

5. DATE OF PUBLICATION (Month and year of publication of document.)	6a. NO. OF PAGES (Total pages, including Annexes, excluding DCD, covering and verso pages.)	6b. NO. OF REFS (Total references cited.)
August 2019	22	15

7. DOCUMENT CATEGORY (e.g., Scientific Report, Contract Report, Scientific Letter.)

Scientific Report

8. SPONSORING CENTRE (The name and address of the department project office or laboratory sponsoring the research and development.)

DRDC – Toronto Research Centre
Defence Research and Development Canada
1133 Sheppard Avenue West
Toronto, Ontario M3K 2C9
Canada

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
04k	

10a. DRDC PUBLICATION NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)
DRDC-RDDC-2019-R117	

11a. FUTURE DISTRIBUTION WITHIN CANADA (Approval for further dissemination of the document. Security classification must also be considered.)

Defence departments and their contractors

11b. FUTURE DISTRIBUTION OUTSIDE CANADA (Approval for further dissemination of the document. Security classification must also be considered.)

12. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Use semi-colon as a delimiter.)

Human Vision; Stereopsis; Aviation Medicine; Depth Perception

13. ABSTRACT (When available in the document, the French version of the abstract must be included here.)

Clinical measurement of stereoacuity was added to the RCAF biannual medical for aircrew in 2012, but at this time there is no standard or cut-off for stereoacuity applied to the RCAF aircrew occupations. In contrast, several allied air forces, such as the United States Air Force (USAF) and the Royal Air Force (RAF), apply a stereoacuity standard. The primary goal of the present study was to characterize the stereoacuity distribution of RCAF aircrew, and to determine the potential impact of various possible minimum stereoacuity standards given the present RCAF population. Stereoacuity measurements were extracted from a subset of the RCAF aircrew medical records and distributional analyses were conducted. We estimated the effect of applying the USAF and RAF stereoacuity standards, as well as other medically-motivated cutoff standards. Our analysis indicates that the adoption of the USAF and USN stereoacuity standard would likely affect large proportions of the RCAF population, while the RAF standard would have a much more modest impact, as might be expected given that these standards are relatively conservative and liberal, respectively. Given these potentially large effects on the RCAF population, the motivation for a stereoacuity standard must be carefully considered. While there is converging evidence that stereopsis is relevant to certain aviation sub-tasks, there is little direct evidence to support a specific stereoacuity cut-off for aircrew. This situation tends to support a liberal cut-off standard that ensures functioning stereopsis in aircrew. However, further research must be conducted to establish a direct link between stereoacuity and operational performance and determine whether or not a more conservative stereoacuity requirement should be in place.

En 2012, la mesure clinique de l’acuité visuelle stéréoscopique a été ajoutée à l’examen médical bisannuel du personnel navigant de l’ARC, mais à présent, aucune norme ou aucun seuil ne s’applique pour l’acuité visuelle stéréoscopique des groupes professionnels du personnel navigant de l’ARC. Cependant, plusieurs forces aériennes alliées, comme la United States Air Force (USAF) et la Royal Air Force (RAF), appliquent une norme d’acuité visuelle stéréoscopique. L’objectif principal de la présente étude visait à caractériser la répartition de l’acuité visuelle stéréoscopique parmi le personnel naviguant de l’ARC et de déterminer l’incidence potentielle de diverses normes minimales possibles de l’acuité visuelle stéréoscopique de la population actuelle de l’ARC. Les mesures d’acuité visuelle stéréoscopique ont été extraites d’un sous-ensemble de dossiers médicaux du personnel navigant de l’ARC et des analyses de répartition ont été menées. Nous avons estimé l’effet de l’application des normes d’acuité visuelle stéréoscopique de la USAF et de la RAF, ainsi que d’autres normes de seuil qui sont liées à une condition médicale. Notre analyse indique que l’adoption de la norme de l’acuité visuelle stéréoscopique de la USAF et de la USN aurait des répercussions sur un grand pourcentage de la population de l’ARC, tandis que la norme de la RAF aurait une incidence beaucoup plus modeste, comme nous pourrions nous y attendre puisque ces normes sont relativement prudentes et libérales, respectivement. Étant donné que ces répercussions pourraient être importantes sur la population de l’ARC, il faut examiner attentivement ce qui motive l’adoption d’une norme d’acuité visuelle stéréoscopique. Bien qu’il existe des données convergentes sur la pertinence de la stéréopsie pour certaines tâches secondaires de l’aviation, il existe peu d’éléments de preuves directs pour appuyer un seuil particulier d’acuité visuelle stéréoscopique pour le personnel navigant. Cette situation peut appuyer un seuil libéral qui assure une bonne stéréopsie du personnel navigant. Toutefois, d’autres recherches doivent être menées pour établir un lien direct entre l’acuité visuelle stéréoscopique et le rendement opérationnel et pour déterminer si une exigence plus prudente en matière d’acuité visuelle stéréoscopique devrait être en place ou non.