



# Heart Rate, Heart Rate Variability and Subjective Workload of Commercial Pilots During Jet Airplane Handling Maneuvers

Samuel Clément-Coulson<sup>1</sup>(✉), Alaa Boutelaa<sup>1,2</sup>, Ramiya Veluppillai<sup>1,2</sup>, and Aaron P. Johnson<sup>1,2</sup>

<sup>1</sup> Concordia University, Montreal, Canada  
{sam.coulson, aaron.johnson}@concordia.ca

<sup>2</sup> Québec Vision Research Network, Montreal, Canada

**Abstract.** Cardiac function (measured by heart rate and heart rate variability) have been investigated as objective measures of pilot workload, by comparing cardiac function between different flight manoeuvres. However, no study has investigated whether cardiac function measures can detect workload changes between manoeuvres of the same type, but different levels of difficulty (e.g., normal turns and steep turns). Commercial pilots ( $n = 14$ ) flew a short scenario in a narrow-body jet simulator, which consisted of a normal turn, simple stall, steep turn, and complex stall. Heart rate, heart rate variability, NASA-TLX and flight path deviations were recorded. We found that heart rate and heart rate variability patterns were stable between participants and show clear differences between manoeuvre types. However, we did not observe any change due to manoeuvre difficulty. NASA-TLX ratings highlighted differences between manoeuvre difficulty, but not manoeuvre type. Our findings suggest that a combination of cardiac measures and subjective measures is best to understand workload. Additional research is required to establish guidelines for use of cardiac function as an indicator of workload.

**Keywords:** Workload · Aviation · Heart rate · Heart rate variability · NASA-TLX · Simulator

## 1 Introduction

Extensive research in human performance has investigated the association between operator states and performance. As pilots' primary role in flight operations is to perceive, process and act upon information from their environment, flight operations rely on the human operator's cognitive resources to accomplish this task (Endsley 1995; Rolfe and Lindsay 1973). A cognitive resource frequently discussed in aviation is workload. Workload arises when a human operator must dedicate physical and psychological resources towards performance on a task (Hart and Staveland 1988; Young et al. 2015). The association between operator workload and operator task performance is well established and can be described by an inverted U-shape: operator performance is poor when workload is either too low or too high (Young et al. 2015). As such, considerable efforts have

been dedicated by researchers to measuring workload. One of these approaches aims to infer operator workload from psychophysiological measures. Proponents of this approach highlight that, in addition to being objective measures, they are not intrusive to the primary task, allowing workload measurement in a naturalistic environment (Cain 2007; Lehrer et al. 2010). However, interpreting physiological measures in relation to flying tasks is difficult as these measures are a result of overall psychophysiological activity.

Of the various psychophysiological measures of workload, a large body of research has investigated cardiac function through electrocardiography (ECG; reviewed in Roscoe 1992). In aviation, this is commonly done by measuring heart rate and heart rate variability (HRV) across different phases of flight. Heart rate is higher in phases with increased workload such as take-off and landing, across different pilots and airplanes (Causse et al. 2012; Roscoe 1993). Even in trained fighter pilots, progressive increases in heart rate have been observed (Mansikka et al. 2015, 2016). Abnormal situations such as in-flight emergencies have also been shown to increase heart rate (Kinney and O'Hare 2020). Investigating HRV (i.e., the variation in the duration of the R-R interval), De Rivecourt et al. (2008) conducted a study of pilot candidates in an instrument flying simulator. Decreases in HRV from baseline, indicative of increased parasympathetic activation, were observed during manoeuvres considered to impose higher workload (Mansikka et al. 2015, 2019). Together, these experiments provide ample evidence that heart rate and HRV change across phases of flight.

While this research shows that cardiac function differs between separate phases of a flight, it remains difficult to establish that these changes are a direct function of differences in workload. It is conceivable that these changes are a result of other task-related factors (e.g., stress, attention). Roscoe (1978) reviewed early studies of cardiac response of two-crew operations (i.e.: one flying and one nonflying pilot). Because both pilots are exposed to the same psychological stressors, only workload is different between pilots. He suggested that cardiac responses changed consistently for the pilot flying, but nonflying pilot responses varied based on the degree of task involvement, supporting its use as an indicator of workload. Furthering the association between cardiac function and workload, cardiac function also correlates with subjective workload across different phases of flight (Lee and Liu 2003; Mansikka et al. 2019). Together, these studies provide further evidence that cardiac function changes in response to the workload imposed by flight. However, most of this research employs an experimental design where workload is compared between different types of flying tasks (e.g.: different phases of flight or different manoeuvres). Therefore, it is unknown whether heart rate and HRV are sensitive to workload changes between similar flight tasks, where workload is manipulated by increasing the task difficulty.

In sum, research in aviation and similar domains suggest that heart rate and HRV are reflecting the demands of flight on pilots. However, the majority of this research investigates workload by comparing measurements between manoeuvres, phases of flight, or pilots' task. To the authors knowledge, heart rate and HRV have seldom been used to measure workload across different variants of the same manoeuvre. Thus, the present study aims to investigate the use of ECG as a measure of workload through a different aviation paradigm where workload is induced by increasing the difficulty of a manoeuvre.

## 2 Method

The sample included commercial ( $n = 5$ ) and airline ( $n = 9$ ) pilots (males  $n = 12$ , Age<sub>Range</sub> 23–42, Median = 26.5). All pilots held a valid medical certificate, multi-engine qualification, and instrument flying rating. Participants flew an Ascent XJ flight training device (Mechtronix, Montréal, Canada); a fixed base simulator configured as a narrow body transport jet. Participants acted as pilot flying, while a qualified pilot researcher acted as pilot monitoring.

Electrocardiogram data were collected using a Polar H10 heart rate monitor (Polar Electro, Kempele, Finland) sampling at 1000 Hz. The Elite HRV app (Polar Electro, Kempele, Finland; Elite HRV LLC, Asheville, USA) exported the sequential R-R interval durations in milliseconds. Subjective workload was measured with the NASA-TLX (Hart and Staveland 1988) application on an Apple iPad Pro (Apple, Cupertino, USA).

### 2.1 Procedure

Study objectives, procedures, and simulator handling characteristics were briefed, along with informed consent. Then, pilots completed two flights. The first flight familiarized participants with the simulator's handling characteristics and ensure harmony between the pilot flying and pilot monitoring. The second flight was the experiment flight. Take-off and landing were conducted for ecological validity but were not evaluated. Table 1 describes the manoeuvre sequence for both flights. Pilots completed the NASA-TLX between manoeuvres.

**Table 1.** Manoeuvre sequence for practice and experimental flight.

Seq	Practice flight	Experiment flight	Details
0	Take off	Take off	
1	Normal Turn (NT)	Normal Turn (NT)	30° Bank Turn
2	Simple Approach to Stall (SAS)	Simple Approach to Stall (SAS)	Approach to Stall – Clean Configuration
3	Landing	Steep Turn (ST)	45° Bank Turn
4		Complex Approach to Stall (CAS)	Approach to Stall – Landing Configuration
5		Landing	

### 2.2 Data Handling

Due to the small participant sample size, missing data were excluded as interpolation or mean substitution would be unreliable. ECG data was missing for one participant. NASA-TLX subjective workload ratings were missing for two participants. Interbeat interval data were analyzed in Kubios HRV 3.3.1 (Kubios, Kuopio, Finland). Heart

rate and heart rate variability (SDNN) were calculated for time windows corresponding to each manoeuvre. The simulator data was extracted via video recording of a screen displaying simulator parameters via optical character recognition using a custom R program implementing Google's Tesseract OCR engine (Google, Mountain View, USA). Flight path deviations were calculated as an objective performance measure for each manoeuvre. For turns, this refers to the time, in seconds, spent outside acceptable airspeed (240 knots – 260 knots) and altitude (4900 feet – 5100 feet) (Transport Canada 2017). For stalls, this refers to the interval between the stall warning onset and the point when safe airspeed (200 knots) and altitude (4800 feet) were attained.

### 3 Results

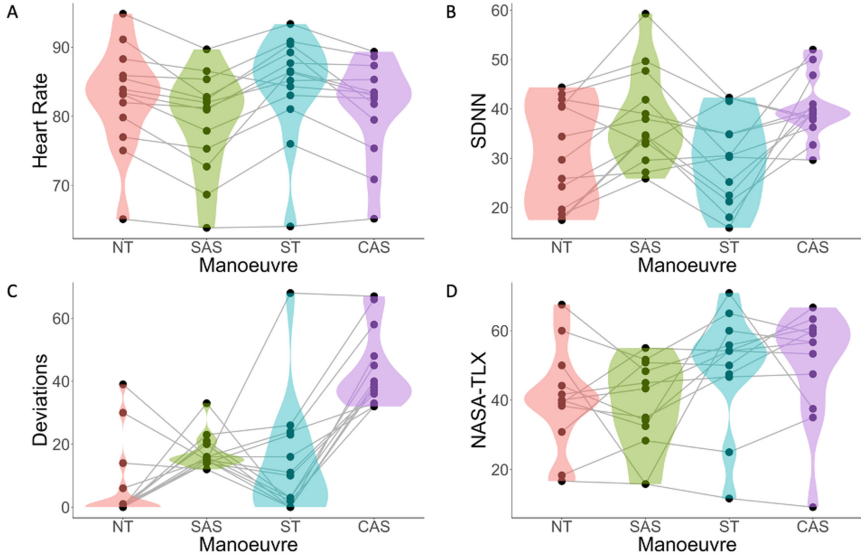
Descriptive statistics and distributions plots were calculated for flight path deviations to ensure the complex manoeuvres were more difficult, showed by increased deviations. Mean flight path deviations were higher during steep turns ( $M = 13.29$ ,  $SD = 18.47$ ) than during normal turns ( $M = 6.50$ ,  $SD = 12.60$ ,  $t = -1.57$ ). Mean recovery time for the complex approach to stall was higher ( $M = 43.50$ ,  $SD = 11.95$ ) than for the simple approach to stall ( $M = 17.50$ ,  $SD = 5.37$ ;  $t = -8.19$ ). The distribution plots suggest that greater range in performance between participants was greater for the complex variants. Altogether, these data suggest that the complex manoeuvres were more difficult than their simple counterparts.

**Table 2.** Descriptive Statistics for heart rate, heart rate variability (SDNN), flight path deviations and NASA-TLX

	Heart rate (BPM)				SDNN			
	CAS	NT	SAS	ST	CAS	NT	SAS	ST
<i>n</i>	13	13	13	13	13	13	13	13
<i>M</i>	81.21	82.70	79.20	84.47	40.15	30.62	37.96	29.98
<i>SD</i>	7.03	7.52	7.34	7.63	6.30	10.59	9.59	9.04
Min.	65.18	65.12	63.86	64.06	29.63	17.47	25.88	15.89
Max.	89.35	94.85	89.70	93.37	52.01	44.35	59.27	42.28
	Flight path deviations				NASA-TLX			
	CAS	NT	SAS	ST	CAS	NT	SAS	ST
<i>n</i>	14	14	14	14	12	12	12	12
<i>M</i>	43.50	6.50	17.50	13.29	50.49	40.56	37.99	49.58
<i>SD</i>	11.95	12.60	5.37	18.47	16.26	14.69	13.33	16.39
Min.	32.00	0.00	12.00	0.00	9.17	16.67	15.83	11.67
Max.	67.00	39.00	33.00	68.00	66.67	67.50	55.00	70.83

*Note.* CAS: complex approach to stall, NT: normal turn, SAS: simple approach to stall, ST: steep turn.

Descriptive statistics and distributions plots were calculated for heart rate, heart rate variability, and NASA-TLX ratings (see Table 2 & Fig. 1). Qualitative observation of the graphs suggests that heart rate and heart rate variability are different between manoeuvre types, but not between the levels of difficulty.



**Fig. 1.** Distribution plots for each manoeuvre. Each dot represents a single participant, and lines connect within-subject observations. (A) Heart rate in beats per minute (B) Heart rate variability (SDNN) in milliseconds (C) Time outside ideal flight path, in seconds. (D) Compound NASA-TLX subjective workload ratings, in percentage. *Note:* NT: normal turn, SAS: simple approach to stall (clean configuration), ST: steep turn, CAS: complex approach to stall (landing configuration).

Repeated-measures analyses of variance (ANOVA) were conducted to investigate between-manoevrue differences in cardiac function, and NASA-TLX ratings. Planned post-hoc, uncorrected, dependent samples t-tests are reported to compare manoeuvre pairs, contrasting the turns (i.e., normal vs. steep turn) and approaches to stall (i.e., simple vs. complex). Effect sizes, Bayes Factors and confidence intervals will be reported, but not p-values. Interpretations of Bayes Factors follow general guidelines (Wetzels et al. 2011). Due to the sample size, these statistical tests are considered exploratory and should be interpreted with caution.

First, an ANOVA was conducted to investigate overall heart rate differences between manoeuvres, which suggests an effect of manoeuvre on heart rate,  $F(3,36) = 9.97$ ,  $\eta^2 = .45$ , with decisive evidence for the research hypothesis ( $BF_M = 324.58$ ). Pairwise comparisons found anecdotal evidence suggesting a medium increase in heart rate between normal and steep turns ( $BF_{10} = 1.15$ ,  $t = 1.92$ ,  $d = .53$ , 95% CI  $[-.06, 1.11]$ ) and between the simple and complex approaches to stall ( $BF_{10} = 1.09$ ,  $t = 1.88$ ,  $d = .52$ , 95% CI  $[-.07., 1.09]$ ). However, the Bayes Factors and qualitative observation of the descriptives plot suggests these differences are not large enough to be meaningful.

Second, an ANOVA was conducted to investigate overall differences in SDNN between manoeuvres, which suggests a moderate effect of manoeuvre on SDNN ( $F(3,36) = 8.70, \eta^2 = .42$ ) with decisive evidence for the research hypothesis ( $BF_M = 164.82$ ). Two non-directional, dependent sample t-tests were conducted to compare simple and complex manoeuvre variants. Anecdotal evidence was found suggesting no difference in SDNN between normal and steep turns ( $BF_{10} = .30, t = .39, d = .11, 95\% \text{ CI} [-.44, .65]$ ) and between the complex and simple stalls ( $BF_{10} = .38, t = -.87, d = -.44, 95\% \text{ CI} [-.79, .32]$ ). However, qualitative observation of the descriptives plot suggested differences between manoeuvre types (Fig. 1). Dependent, non-directional t-tests were therefore conducted contrasting the normal turn with the simple approach to stall, and the steep turn with complex approach to stall. Substantial evidence was found suggesting that the SDNN was lower during the turns than during the stalls (for simple variants:  $BF_{10} = 9.22, t = -.36, d = -.93, 95\% \text{ CI} [-1.58, .26]$ ; for complex variants:  $BF_{10} = 9.91, t = -3.41, d = -.95, 95\% \text{ CI} [-1.59, -.27]$ ).

Last, an ANOVA was conducted to investigate overall differences in subjective workload, which found substantial evidence supporting differences in subjective workload between manoeuvres ( $F(3, 33) = 5.24, \eta^2 = .32, BF_M = 9.78$ ). Two non-directional, dependent sample t-tests were conducted to compare simple and complex manoeuvre variants. Substantial evidence was found suggesting a difference in subjective workload between normal and steep turns ( $BF_{10} = 4.22, t = 2.88, d = .83, 95\% \text{ CI} [.15, 1.48]$ ) and between the complex and simple stalls ( $BF_{10} = 9.24, t = 3.41, d = .99, 95\% \text{ CI} [.27, 1.67]$ ).

## 4 Discussion

Here we aimed to investigate the usability of heart rate and HRV to detect workload changes between maneuvers of varying complexity. We hypothesized that heart rate, HRV and NASA-TLX would be able to differentiate between high and low workload manoeuvres, and between manoeuvre types. Our results provide partial support for our hypothesis. Heart rate and HRV were different between turns and stalls, but did not vary meaningfully between the simple and complex variants. On the other hand, NASA-TLX scores were different between the simple and complex variants, but not between manoeuvre types.

Increases in heart rate and decrease in HRV (measured by SDNN) are associated with increased arousal and workload (Jorna 1993; Luque-Casado et al. 2016). Therefore, we expected to observe an increase in heart rate and a decrease in HRV in the complex exercises. However, we observed clear differences between manoeuvre types (i.e., turn vs stall), but little difference was found between the complex and simple variants of each manoeuvre. Pilots' heart rate was higher during the turns than during the approaches to stall, and HRV was lower during turns. When contrasting the complex manoeuvres to their simple variant, heart rate and HRV were similar. Given the limited sample size, we were unable to conduct traditional parametric statistical tests on these between-condition differences. Nevertheless, the general pattern of the results suggests cardiac function was different between manoeuvre types (i.e., turns vs. approaches to stall), but not between different difficulties of the same manoeuvre type.

These results do replicate the findings reported in similar studies (Causse et al. 2012; Hankins and Wilson 1998; Kinney and O'Hare 2020). Taken together with the results of the present experiment, heart rate appears to change predictably due to workload induced by different manoeuvre types, but not vary as a function of workload changes due to manoeuvre difficulty. It is probable that the manipulation of difficulty may not have been sufficient to create a significant change in heart rate (see workload and performance limitations later). Additional research is required to define contexts where psychophysiological measures are most appropriate measurements of workload in human factors research. For example, the changes in cardiac function related to workload imposed by basic processes required to complete a task, regardless of its difficulty level. One could hypothesize that the patterns of between-manoevrue changes in heart rate and HRV are caused by the inherent differences in pilot resources required for completing each manoeuvre. Turns require sustained attention, and consistently require updating a model in working memory with multiple parameters (i.e., bank angle, altitude, airspeed, engine thrust) and precision. Conversely, stalls require the prompt application of a trained rehearsed recovery procedure, with an emphasis placed on quickly and successfully exiting the situation, and not on the precision of flying inputs. Changes in psychophysiological measures may occur due to these different types of task demands.

Conversely, subjective workload evaluations are biased by previous exposure to similar manoeuvre and performance (Moore and Picou 2018). In the present experiment, pilots reported differences in workload between the simple and complex manoeuvres, but not between manoeuvre types. It is possible that pilots' workload estimates were derived from previous exposure to similar situations, not as a function of the resources required. This may explain subjective workload scores' ability to differentiate between manoeuvre difficulty levels, as flight path deviations were generally higher during the difficult manoeuvre, but not manoeuvre type.

This study does have limitations. We were unable to attain the planned sample size due to the COVID-19 pandemic. As such, the statistical tests included are considered exploratory. However, our sample size remains similar to that of aviation psychophysiological research (e.g.: Hankins and Wilson 1998; Hidalgo-Muñoz et al. 2018; Lee and Liu 2003; Mansikka et al. 2019). This study retains the limitations associated with the use of heart rate and HRV as objective workload indicators. Cardiac function is the result of multiple physiological processes occurring simultaneously, so its use as a workload metric remains confounded by external physiological factors. This renders interpretation of heart rate and HRV difficult in the aviation environment. Further research is required to identify guidelines for use of cardiac function as a workload indicator.

The present results add to the literature supporting the strategic use of cardiac function as an objective workload measure to compare aviation situations likely to require different types of cognitive and physiological resources on behalf of the pilot. The present study suggests that heart rate and HRV are limited indicators of workload in the cockpit, especially when different variants of the same manoeuvre are being studied. Nonetheless, cardiac function can complement subjective measures to better understand workload in the cockpit.

**Acknowledgements.** The authors would like to thank the Centre Québécois de Formation Aéronautique for their expertise and simulator access, without which this experiment would not have been possible.

## References

- Cain, B.: A review of the mental workload literature (2007). <https://apps.dtic.mil/dtic/tr/fulltext/u2/a474193.pdf>
- Causse, M., Dehais, F., Faaland, P.-O., Cauchard, F.: An analysis of mental workload and psychological stress in pilots during actual flight using heart rate and subjective measurements. In: Proceedings of the 5th International Conference on Research in Air Transportation (ICRAT 2012) (2012)
- De Rivecourt, M., Kuperus, M.N., Post, W.J., Mulder, L.J.M.: Cardiovascular and eye activity measures as indices for momentary changes in mental effort during simulated flight. *Ergonomics* **51**(9), 1295–1319 (2008). <https://doi.org/10.1080/00140130802120267>
- Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. *Hum. Factors: J. Hum. Factors Ergon. Soc.* **37**(1), 32–64 (1995). <https://doi.org/10.1518/001872095779049543>
- Hankins, T.C., Wilson, G.F.: A comparison of heart rate, eye activity, EEC and subjective measures of pilot mental workload during flight. *Aviat. Space Environ. Med.* **69**(4), 360–367 (1998)
- Hart, S.G., Staveland, L.E.: Development of NASA-TLX (task load index): results of empirical and theoretical research. In Hancock, P.A., Meshkati, N. (eds.) *Advances in Psychology*, no. 52, pp. 139–183. Elsevier Science Publishers (1988)
- Hidalgo-Muñoz, A.R., Mouratille, D., Matton, N., Causse, M., Rouillard, Y., El-Yagoubi, R.: Cardiovascular correlates of emotional state, cognitive workload and time-on-task effect during a realistic flight simulation. *Int. J. Psychophysiol.* **128**, 62–69 (2018). <https://doi.org/10.1016/J.IJPSYCHO.2018.04.002>
- Jorna, P.G.A.M.: Heart rate and workload variations in actual and simulated flight. *Ergonomics* **36**(9), 1043–1054 (1993). <https://doi.org/10.1080/00140139308967976>
- Kinney, L., O’Hare, D.: Responding to an unexpected in-flight event: physiological arousal, information processing, and performance. *Hum. Factors* **62**(5), 737–750 (2020). <https://doi.org/10.1177/0018720819854830>
- Lee, Y.H., Liu, B.S.: Inflight workload assessment: comparison of subjective and physiological measurements. *Aviat. Space Environ. Med.* **74**(10), 1078–1084 (2003)
- Lehrer, P., Karavidas, M., Lu, S., Vaschillo, E., Vaschillo, B., Cheng, A.: Cardiac data increase association between self-report and both expert ratings of task load and task performance in flight simulator tasks: an exploratory study. *Int. J. Psychophysiol.* **76**(2), 80–87 (2010). <https://doi.org/10.1016/j.ijpsycho.2010.02.006>
- Luque-Casado, A., Perales, J.C., Cárdenas, D., Sanabria, D.: Heart rate variability and cognitive processing: the autonomic response to task demands. *Biol. Psychol.* (2016). <https://doi.org/10.1016/j.biopsycho.2015.11.013>
- Mansikka, H., Simola, P., Virtanen, K., Harris, D., Oksama, L.: Fighter pilots’ heart rate, heart rate variation and performance during instrument approaches. *Ergonomics* **59**(10), 1344–1352 (2016). <https://doi.org/10.1080/00140139.2015.1136699>
- Mansikka, H., Virtanen, K., Harris, D.: Comparison of NASA-TLX scale, modified Cooper-Harper scale and mean inter-beat interval as measures of pilot mental workload during simulated flight tasks. *Ergonomics* **62**(2), 246–254 (2019). <https://doi.org/10.1080/00140139.2018.1471159>
- Mansikka, H., Virtanen, K., Harris, D., Simola, P.: Fighter pilots’ heart rate, heart rate variation and performance during an instrument flight rules proficiency test. *Appl. Ergon.* **56**, 213–219 (2015). <https://doi.org/10.1016/j.apergo.2016.04.006>



- Moore, T.M., Picou, E.M.: A potential bias in subjective ratings of mental effort. *J. Speech Lang. Hear. Res.* **61**(9), 2405–2421 (2018). [https://doi.org/10.1044/2018\\_JSLHR-H-17-0451](https://doi.org/10.1044/2018_JSLHR-H-17-0451)
- Rolfe, J.M., Lindsay, S.J.E.: Flight deck environment and pilot workload: biological measures of workload. In: *Applied Ergonomics* (1973). <https://pdf.sciencedirectassets.com/271441/1-s2.0-S0003687000X01391/1-s2.0-0003687073902159/main.pdf?x-amz-security-token=AgoJb3JpZ2luX2VjECUaCXVzLWVhc3QtMSJHMEUCIE8HuUIP70id%2BofcCfxdRRcptOGEFsZxbs042ATeX70zAiEA1OOTw1EEExc%2B%2BpX53gibtSX3qgMHmKLor6gFrHdQ>
- Roscoe, A.H.: Stress and workload in pilots. *Aviat. Space Environ. Med.* **49**(4), 630–636 (1978)
- Roscoe, A.H.: Assessing pilot workload. Why measure heart rate, HRV and respiration? *Biol. Psychol.* **34**(2–3), 259–287 (1992). [https://doi.org/10.1016/0301-0511\(92\)90018-P](https://doi.org/10.1016/0301-0511(92)90018-P)
- Roscoe, A.H.: Heart rate as a psychophysiological measure for in-flight workload assessment. *Ergonomics* **36**(9), 1055–1062 (1993). <https://doi.org/10.1080/00140139308967977>
- Transport Canada: Pilot proficiency check and aircraft type rating flight test guide (Aeroplanes) (Issue First Edition (Revision 1)). Government of Canada (2017). <https://www.tc.gc.ca/eng/civilaviation/publications/tp14727-menu-2709.htm>
- Wetzels, R., Matzke, D., Lee, M.D., Rouder, J.N., Iverson, G.J., Wagenmakers, E.: Statistical evidence in experimental psychology: an empirical comparison using 855 t tests. *Perspect. Psychol. Sci.* **6**(3), 291–298 (2011). <https://doi.org/10.1177/1745691611406923>
- Young, M.S., Brookhuis, K.A., Wickens, C.D., Hancock, P.A.: State of science: mental workload in ergonomics. *Ergonomics* **58**(1), 1–17 (2015). <https://doi.org/10.1080/00140139.2014.956151>