International Journal of Industrial Ergonomics 54 (2016) 120–130

Contents lists available at ScienceDirect

International Journal of Industrial Ergonomics

journal homepage: www.elsevier.com/locate/ergon

Air traffic control: Ocular metrics reflect cognitive complexity

Mauro Marchitto [a, *], Simone Benedetto [b, c], Thierry Baccino [b, c], José Juan Cañas [a]

[a] *Cognitive Ergonomics Group, Mind, Brain and Behavior Research Centre (CIMCYC), University of Granada, Campus de Cartuja s/n, 18071, Granada, Spain*
[b] *Université Paris 8, Vincennes – St. Denis, 2 rue de la Liberté, 93200 Saint-Denis, France*
[c] *CHArt/LUTIN (EA 4004), Cité des Sciences et de l'Industrie de la Villette, 30 Avenue Corentin Cariou, 75930 Paris, France*

A R T I C L E I N F O

Article history:
Received 16 February 2015
Received in revised form
22 April 2016
Accepted 18 May 2016

Keywords:
Conflict detection
Traffic safety
Eye tracking
Mental workload
Cognitive strategy
Decision making

A B S T R A C T

The objective of the study was to evaluate effects of complexity on cognitive workload in a simulated air traffic control conflict detection task by means of eye movements recording. We manipulated two complexity factors, convergence angle and aircrafts minimum distance at closest approach, in a multi-dimensional workload assessment method based on psychophysiological, performance, and subjective measures. Conflict trials resulted more complex and time-consuming than no conflicts, requiring more frequent fixations and saccades. Moreover, large saccades showed reduced burst power with higher task complexity. A motion-based and a ratio-based strategy were suggested for conflicts and no conflicts on the basis of ocular metrics analysis: aircrafts differential speed and distance to convergence point at trial start were considered determinant for strategy adoption.

Relevance to industry: Eye metrics measurement for online workload assessment enhances better identification of workload-inducing scenarios and adopted strategy for traffic management. System design, as well as air traffic control operators training programs, might benefit from on line workload measurement.

1. Introduction

According to the International Civil Aviation Organization (ICAO) medium-term passenger traffic is expected to increase by 6.3% in 2015 (ICAO, 2012) and air traffic control (ATC) is going to be one of the most affected activities of the whole air transport system. ATC is a service offered by ground-based operators (ATCO), which aims at preventing conflicts between aircrafts, by managing the resulting complexity. It mainly consists of organizing the traffic flow, and providing information and support for pilots. In order to prevent collisions, ATCOs employ traffic separation rules, which constantly ensure the maintaining of a minimum amount of empty space around airplanes. The concept of complexity in the ATC domain has received considerable attention along the years, and a large amount of studies classify complexity at several levels of analysis: environmental, organizational, traffic, and display (e.g. Cummings and Tsonis, 2005; Mogford et al., 1995). These elements influence controllers' cognitive complexity, i.e. the perceived complexity at the individual level. Many studies have been carried out for identifying factors that affect cognitive complexity in air traffic scenarios (Hillburn, 2004; Histon et al., 2002), and to devise evaluation frameworks to be applied both in simulation and real work contexts (Majumdar and Ochieng, 2007; Pawlak et al., 1996). Other studies have focused on the effects of complexity on ATCO, showing that it affects mental workload, the allocation of mental resources for accomplishing task demand in a safe and efficient manner (Athènes et al., 2002; Li et al., 2010).

According to Leplat (1978) mental workload (also referred to as cognitive workload) is a multidimensional construct, rather than a unidimensional one. In this sense the multidimensional assessment of mental workload by triangulation from physiology, performance, and subjective assessments is a fruitful approach (Parasuraman et al., 2008; Wierwille and Eggemeier, 1993). In ATC, high performance standards are usually maintained, independently of task complexity. Therefore, the monitoring of controller's performance cannot always convey the real cognitive demand imposed by the task. For this reason the need of online measurement of mental workload is an opportunity to seek. When dealing with safety-critical work contexts, the best option for workload assessment relies on unobtrusive techniques (Langan-Fox et al., 2009) and psychophysiology seems to be one of the most promising fields for

* Corresponding author.
E-mail addresses: maurom@ugr.es (M. Marchitto), benedetto.simone@gmail.com (S. Benedetto), baccino@lutin-userlab.fr (T. Baccino), delagado@ugr.es (J.J. Cañas).

http://dx.doi.org/10.1016/j.ergon.2016.05.010

the online measurement of the operator's state (Wilson and Russell, 2003). The main advantage of psychophysiological responses is that they do not require an overt response by the operator, and can be collected continuously with relatively low obtrusiveness. In this framework, workload has been also investigated by means of real time research techniques such as electroencephalography (Weiland et al., 2013), and optical brain imaging (Ayaz et al., 2010), but the level of intrusiveness of such techniques still represents a limitation. In contrast, the increased sophistication and accessibility of eye tracking technologies have generated a great deal of interest around eye measures (Ahlstrom and Friedman-Berg, 2006). While in the past they required invasive equipment, unsuitable for most applied settings, nowadays large advances in technology have made the equipment much more portable and capable.

In the framework of conflict detection in ATC, we explored the effects of complexity on mental workload by means of a multidimensional assessment method based on psychophysiological (eye movements), performance (response time and accuracy), and subjective measures. With respect to previous investigations, the main contribution of this study lies on the employment of eye movements for the online assessment of workload during a simulated ATC task.

Conflict detection is one of the main activities performed by ATCOs (Kallus et al., 1999). It consists of comparing trajectories of converging aircrafts and estimating the probability of a future simultaneous violation of vertical and lateral separation standards, which are commonly set to 1000 ft (feet) and 5 nm (nautical miles, the unit of distance used in aviation, which correspond to 1852 m), respectively (Loft et al., 2009). Conflict detection is carried out on radar displays and involves several subtasks such as information location, change detection, and short- and long-term predictions in a complex and dynamic environment (Li et al., 2010). To the best of our knowledge, studies in ATC domain employing psychophysiological data for workload assessment are quite limited and those dealing with conflict detection by means of ocular behaviour recording are even sparser (for a review see Langan-Fox et al., 2009). Eye tracking has been used to explore attention allocation in control tasks (Martin et al., 2011; Lokhande and Reynolds, 2012), task discrimination (Imants and de Greef, 2011), system usability (Jacob and Karn, 2003), and it has been employed as head-free input device (Alonso et al., 2013).

According to the literature dealing with conflict detection, convergence angle (CA) and minimum distance at closest approach (MD) are key factors in determining cognitive complexity. CA is a geometry factor that influences visual information acquisition, the basic cognitive process that enables successive high-level cognitive elaboration and decision making. CA directly influences distance between converging aircrafts. While wider angles increase the separation between aircraft, smaller ones reduce it. For example, two aircrafts flying with the same speed that converge with CA $= 135^\circ$ will keep lateral separation for a longer period with respect to aircrafts converging with CA $= 60^\circ$. As a result, the eyes must perform wider saccadic movements to transit from one aircraft to the other with wider CA. In this respect, Marchitto et al. (2012) showed that an increase of CA affects both conflict detection times and ocular movements, reducing the peak velocity of large reaching saccades. According to Remington et al. (2000), trajectory comparison is faster and more accurate for smaller angles, which are usually associated to higher probability of intervention by ATCO (Loft et al., 2009).

MD is a traffic complexity factor that affects the cognitive simulation process, which involves the projection of future aircrafts positions, the estimation of the distance between them, and the comparison of such distance to the separation standards (i.e. vertical and lateral). Controllers can apply different perceptual and cognitive methods (i.e. time- or space-based strategies) for estimating future relative positions of aircrafts on the basis of relevant flight information such as speed, distance to convergence point, heading, and movement observation (Xu and Rantanen, 2003). In conflict detection, MD has been shown to predict response time, probability of intervention (Loft et al., 2009; Stankovic et al., 2008), and subjective ratings of difficulty and complexity (Boag et al., 2006).

This paper is organized as follows: information about participants, stimuli, apparatus, and a detailed description of the dependent variables with relative hypotheses is provided in Section 2. Results are presented and discussed in Section 3 with a correlational analysis followed by a factorial ANCOVA with CA and MD as predictors, and response time as covariate. Finally, conclusions are drawn in Section 4, together with practical applications and some perspectives for future work.

2. Materials and methods

2.1. Participants

Twenty-six students (22 women, mean age = 22 years, SD = 2 years) from the University of Granada—Faculty of Psychology, volunteered for course credits after signing an informed consent. They all had normal or corrected-to-normal vision (contact lenses were accepted, but not glasses). None of them had previous ATC experience. An internal committee board approved the study, which was performed in keeping with the Declaration of Helsinki.

2.2. Stimuli and apparatus

The ATC-lab^{Advanced} software (Fothergill et al., 2009) was employed for building the air traffic scenarios, which consisted of a central point, i.e. convergence point, and two aircrafts with related flight information moving on predefined routes, as represented in Fig. 1. Four different types of routes were employed, i.e. vertical, horizontal, and two oblique. Aircrafts' position was updated every 5 s. Around the convergence point, a circle with a radius equal to the lateral separation standard for conflict definition (5 nm) was always presented. In this experiment the conflict detection task dealt with leveled aircrafts flying at the same altitude, thus only lateral problems were considered (Loft et al., 2009). Stimuli were

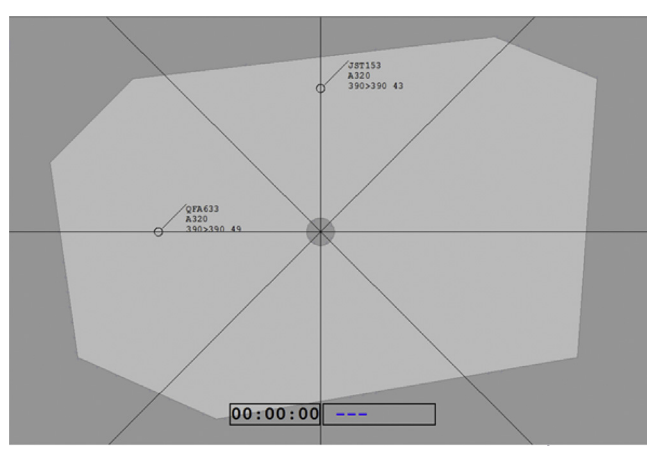

Fig. 1. Airspace structure representation. The circle around the central convergence point has a radius of 5 nm, and it subtends 1° of visual angle from a viewing distance of approximately 60 cm. Aircrafts are represented by a small circle, with relative tag containing flight data: call sign on the first line; type of aircraft on the second line; current altitude, assigned altitude, and speed on the third line.

presented on a CRT monitor (21 inches) with 160 Hz refresh capability, vertically.

Eye movements were recorded with a 500 Hz infrared video-based eye tracker (Eyelink II – SR Research, Ontario, Canada). A nine-point calibration process was carried out before each recording session and a drift correction was performed before the beginning of each trial. In order to limit head movements, we used a chin rest during data acquisition.

2.3. Experimental design and procedure

Two CA levels (90°; 135°) and five MD levels (0; 1; 6; 10; 12 nm) were manipulated in a 2 × 5 within subject design. On the basis of lateral separation standard adopted in the study (5 nm), two levels of MD indicated a conflict (0; 1 nm), while the remaining three a no conflict (6; 10; 12 nm). In total, 10 experimental conditions resulted from factors' combination. In order to cover both right and left halves of display with aircraft position at trial start, we built two scenarios for each condition. Therefore, a set of 20 trials was created and finally uploaded in Experiment Builder software (SR Research, Ontario, Canada). All the scenarios had the same time to minimum separation (i.e. 7 min, blind information). Screenshots from two examples are reported in Fig. 2.

The conflict detection task consisted of judging whether two converging aircrafts would have lost lateral separation through time, i.e. whether the distance between them would have been below (conflict) or remained above (no conflict) the critical separation standard. After task instruction explanations, four training trials were performed (without recording eye tracking data), and accuracy feedbacks were given to improve learning. Experiment began with the randomized presentation of experimental trials. Participants performed the conflict detection task by means of a PC mouse whose buttons were colored in red (conflict, left button) and green (no conflict, right button), respectively. After each trial, participants were asked to verbally express their perceived mental workload. Experimental session lasted for approximately 1 h. A schematic representation of the procedure is provided in Fig. 3.

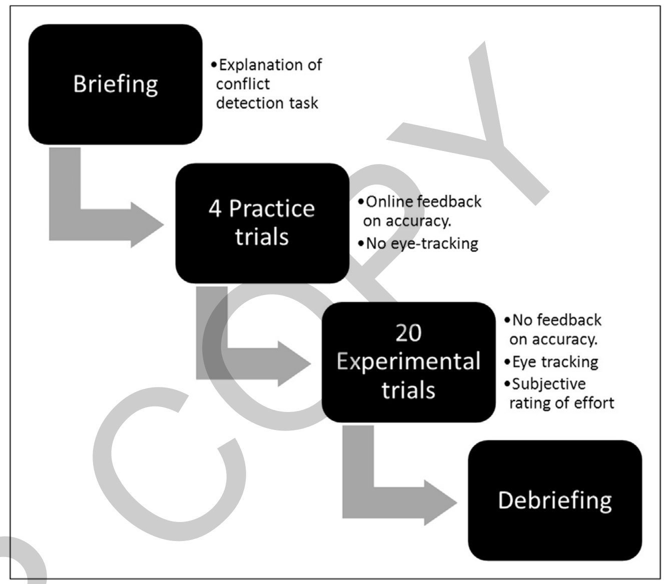

Fig. 3. Experimental procedure breakdown. Participants were briefed on conflict detection task and had the opportunity of practicing it in four practice trials. After the experimental session, a final debriefing on study's hypotheses was provided.

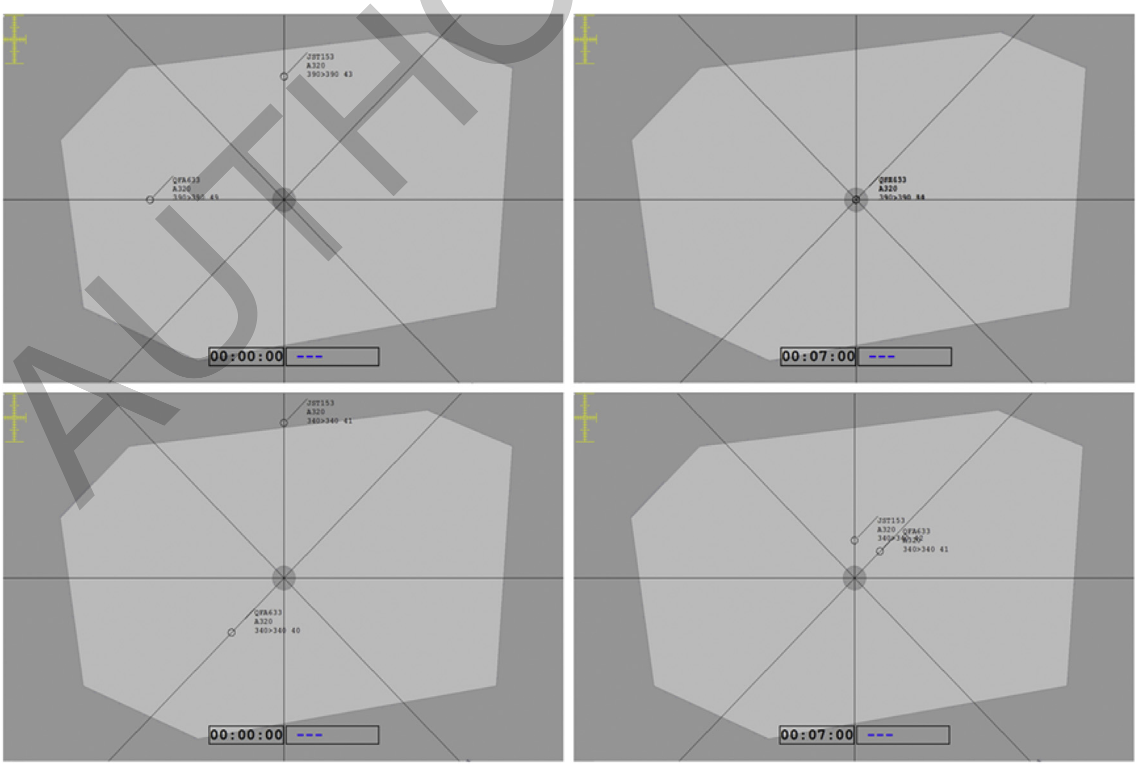

Fig. 2. Examples of traffic scenarios. Left side represents trial start; right side represents aircrafts at closest approach, after 7 min. Top panel: conflict (MD = 0 nm, CA = 90°). Bottom panel: no conflict (MD = 12 nm, CA = 135°).

2.4. Dependent variables

Dependent variables are briefly described and relative predictions formulated in the following sections. A schematic representation of hypotheses is provided in Fig. 4.

2.4.1. Ocular measures

On the basis of the so-called eye-mind relationship, eye tracking is a useful technique for monitoring operators' attention (Poole and Ball, 2005). Fixations and saccades constitute the basic core events of ocular behaviour. Fixations represent an overt visual attention process that directly inform about attentional resources allocation (Just and Carpenter, 1976; Henderson, 2003).

Concerning fixations, two metrics were selected: total number of fixations (*NrF*), and mean fixation duration (*FD*). Several studies have shown the positive correlation between *NrF* and workload (Goldberg and Kotval, 1998; Ha et al., 2006; Lin et al., 2003). According to Henderson (2007), the comprehension of the visual world is strongly dependent on the fixated points, and, *FD* is a direct expression of the time spent during information processing. Longer *FD* has been associated to observer's difficulty in extracting information from a display (Fitts et al., 1950; Goldberg and Kotval, 1998; Callan, 1998). In conflict detection task, MD has been showed to be strongly associated to mental workload (Loft et al., 2009; Xu and Rantanen, 2007; Xu, 2003). In this study we hypothesized that when MD value was close to the separation standard, conflict detection would have been more demanding if compared to MD values considerably above or below the separation criteria. Consequently, we expected that MD = 6 nm would generate higher *NrF* and longer *FD* with respect to remaining MD levels.

Saccades are rapid eye movements occurring between fixations, and are performed to bring task relevant information to the fovea, the retinal area with greatest visual acuity. By means of saccades, attentional requirements are accomplished. No encoding takes place during saccades (saccadic suppression), but saccade dynamics are indicative of oculomotor system's effort during task completion (Goldberg and Kotval, 1999). In this experiment, four metrics were considered: total number of saccades (*NrS*), average saccadic amplitude (*SA*), average saccadic peak velocity (*SPV*), and average saccadic duration (*SD*). The last two parameters are related in a nonlinear manner to *SA*: according to the main sequence relationship, *SD* and *SPV* will increase as long as *SA* also increases (Bahill and Stark, 1975). Main sequence indicates the relationships between the size of saccades and their duration, as well as between size and peak velocity. It shows a linear trend for saccades with *SA* up to 15° or 20°, while reaching a soft saturation for larger saccades. This saturation effect is mainly due to ocular motoneurons firing at their maximum rates when realizing large saccadic movements, up to human *SPV* thresholds. Many studies evidenced how contextual factors such as fatigue (Schmidt et al., 1979; Cazzoli et al., 2014), arousal (Di Stasi et al., 2013; Benedetto et al., 2014), or task difficulty (Di Stasi et al., 2010) can affect saccadic velocities, in particular those of large saccades. On the basis of these evidences, we expected significant main effects of MD and CA on the selected saccade metrics. As to MD, we predicted that correct conflict detections would have required more time, and consequently produced more saccades, with higher complexity trials (MD = 6 nm). Furthermore, on the basis of previous findings (Marchitto et al., 2012), we predicted that large saccades ($SA > 15°$) for high complexity trials would have triggered slower *SPV*, as to indicate the decreased average burst power of oculomotor system in such situations. As to CA, we expected that larger saccades would have been needed for higher CA, i.e. when aircrafts would have been farther each other in the display. According to this assumption, we predicted higher *SA* for wider convergence angles, together with higher *SD* and *SPV*.

2.4.2. Performance measures

As to performance measures, two metrics were selected: mean error rate (*ERR*) and response time (*RT*). The total number of errors was computed for each subject and transformed in error rate (out of a total of 20 experimental trials per subject). For instance, a participant that made 2 errors out of 20 trials, showed an error rate of 0.1. Average *ERR* across all participants was then calculated for each experimental condition. Errors were produced both when no conflicts were gauged as conflict trials (false alarm) and when conflicts were assessed as no conflicts (miss). Although in real ATC situations misses are more dangerous than false alarms, in this experiment both type of errors were considered together, and a unique accuracy metric was computed. In line with previous statements on ocular measures, we expected that more complex trials (MD = 6 nm), would have increased both *ERR* and *RT*. With respect to *RT*, we considered the time between traffic scenario onset (the beginning of each trial) until participant press the button ("yes" or "no") response: less values have been already found for MD = 0; 1; 10; 12 nm in comparison with MD = 6; 8 nm (Loft et al., 2009).

2.4.3. Subjective measures

We employed the Verbal Online Subjective Opinion (VOSO) for subjective workload estimation (*SUB*). The VOSO (Miller, 2001) is an eleven-point (from 0 to 10) unidimensional rating scale, which requires a verbal response of the perceived mental workload experienced during the task. It was administered at the end of each experimental trial. In line with previous predictions on ocular and performance measures, higher VOSO scores were expected for MD = 6 nm, lower for conflicts (MD = 0; 1 nm), and even lower for remaining no-conflict trials (MD = 10; 12 nm), on the basis of MD difference with the lateral separation standard.

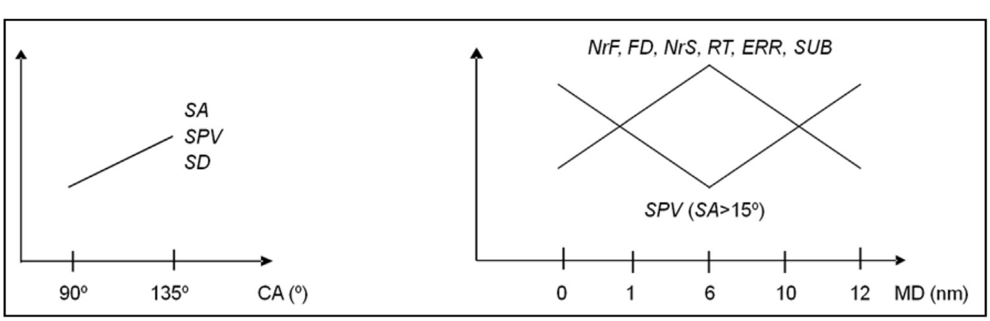

Fig. 4. Graphical representation of hypotheses.

3. Results and discussion

Concerning data analysis, participants with performance accuracy below 70% (i.e. more than 6 errors out of 20 decision trials) were excluded. Such a threshold was employed mainly because in a binary decision task in which conflicts were present in 40% of trials, accuracy levels around 50% could have been reached by guessing. As a result, 9 participants over 26 were excluded from the analysis. Furthermore, only correct trials were included in the analysis of ocular, performance, and subjective data.

In relation to ocular measures, some further thresholds were applied in order to exclude abnormal values. As to fixations, *FD* ranged from 100 m to 1000 m. Fixations around blinks were dropped from the sample data. As to saccades, detection required a deflection in eye position greater than 0.1°, with a minimum velocity of 30°/s and a minimum acceleration of 8000°/s, maintained for at least 4 m. Thresholds for duration (10–250 m) and peak velocity (35–900°/s) were also applied, and saccades around blink eliminated.

This present section is organized as follows. Results of the correlational and an inferential analysis are presented and discussed in sections 3.1 and 3.2 respectively. Finally, implications about adopted strategies are discussed in section 3.3, and a summary of results is provided in section 3.4.

3.1. Correlational study

The correlational analysis was carried out to investigate the occurrence of considerable associations between ocular, performance, and subjective measures. Due to the predictable high positive correlations between *RT* and ocular events frequencies (*NrF* and *NrS*) partial correlations were preferred: *RT* was set as continuous control variable and its effects controlled. Resultant correlations among these variables are reported in Table 1.

As to fixations, *NrF* showed a positive relationship with *NrS* (r = 0.95), while negative with *ERR* (r = −0.38). The alternation of fixations and saccades explains this strong positive association. Moreover, the significant negative correlation with *ERR* suggested that larger frequencies of fixations and saccades were associated to higher accuracy. As to *FD,* a significant negative correlation was found with *NrS* (r = −0.20). This association might be interpreted in terms of different visual behaviors that might be adopted in a visual task. Shorter *FD* and higher saccadic frequency might be typical of exploratory visual behaviour, while longer *FD* and lower saccadic frequency might be expression of focused ocular behaviour. In this sense *FD* confirms to be associated with attentional and cognitive effort in visual tasks.

In relation to saccades, *NrS* correlated negatively with *ERR* (r = −0.39). In this experiment, participants were asked to perform the conflict detection task as best they could. Therefore, the overall quantity of eye events (saccades and fixations) was associated with the accuracy of decision. In particular, it seemed that, independently of time needed to make a decision, conflict detection performance improved with higher frequency of fixations and saccades, i.e. more detailed exploration and frequent processing. In a dynamic decision task as conflict detection, the probability of making the correct decision increases with time, while available time for implementing a conflict solution decreases accordingly. Partial correlations controlled for this effect.

The remaining significant correlations between saccadic metrics were essentially due to main sequence relationship: *SA* highly correlated with *SPV* (r = 0.57), and *SD* (r = 0.40). Similarly, *SPV* and *SD* correlated positively (r = 0.41). Furthermore, *SD* was positively correlated with *SUB* (r = 0.24), as to indicate that less workload was estimated in presence of shorter saccades: short *SD* are associated to short *SA*, and thus to local exploration. In relation to *ERR*, there was a significant positive correlation with *SUB* (r = 0.35), meaning that lower accuracy was associated to higher subjective workload.

In summary, results of the correlation study confirmed that ocular metrics were associated to perceived complexity and resulting cognitive workload. In fact, subjective estimations correlated positively with saccadic parameters, suggesting that oculomotor system effort is taken into account when rating subjective workload. Performance accuracy (*ERR*) correlated negatively with the number of ocular events.

Table 1
Partial correlations matrix. *RT* was set as control variable. Marked correlations (*) significant at $p < 0.05$, and marked correlation (**) significant at $p < 0.01$; N = 17.

	NrF	FD	NrS	SA	SD	SPV	SUB
FD	−0.135						
NrS	0.954**	−0.204**					
SA	−0.013	0.023	−0.065				
SD	−0.065	0.046	−0.092	0.395**			
SPV	−0.094	−0.057	−0.120	0.573**	0.411**		
SUB	−0.091	0.044	−0.098	0.091	0.242**	0.159*	
ERR	−0.379**	0.032	−0.392**	−0.094	0.113	0.083	0.348**

3.2. Inferential study

After computing partial correlations for dependent variables, we wanted to test the effects of CA and MD complexity factors while controlling the for *RT.* A 2 × 5 repeated measures ANCOVA was performed for each of the dependent variables, with *RT* as covariate (Mean *RT* = 28 s; Standard Error, SE = 4.8 s). The significance level α was set at 0.05 for all statistical analyses. When the Mauchly sphericity test was significant, we applied the Greenhouse-Geisser correction. As to MD factor, we performed planned comparisons of MD = 6 nm level with each other one, according to hypotheses. In relation to multiple post-hoc comparisons, we used Bonferroni correction. Means and standard deviations for each dependent variable are reported in Table 2.

3.2.1. Ocular measures

Concerning *NrF* and *FD,* there were not significant main effects of CA and MD factors, as well as CAxMD interactions. Therefore, fixation metrics in the study could be predicted by *RT*: higher *RT* was assumed to indicate higher cognitive complexity of conflict detection and thus, increased frequency and duration of information collection episodes.

Concerning saccades, a further factor was introduced in the analysis, i.e. interval of saccadic amplitude (Bin), and consequently *NrS, SD,* and *SPV* were analyzed as a function of Bin, CA, and MD. Five amplitude intervals were selected: 0.1°–1° (microsaccades), 1°–5° (short saccades), 5°–10° (medium-short saccades), 10°–15° (medium-large saccades), and 15°–53° (large saccades). Wider saccades were expected with higher CA. On the basis of main sequence relationship, we assumed that *SD* and *SPV* should have varied similarly. However, we also hypothesized lower values of *SPV* in most difficult trials (MD = 6 nm) for large saccades (Bin 5). Significant interactions for saccadic metrics are reported in Fig. 5.

As to *NrS*, results replicated those of fixation metrics. Once the effect of *RT* has been controlled by covariance analysis, there was not any remaining effect of complexity factors CA and MD on saccadic frequency. As to *SA*, wider saccades were found for CA = 135°, as hypothesized [$F(1, 15) = 19.73$; $p < 0.001$; $\eta^2_p = 0.59$]. Consequently, the geometrical cost for oculomotor displacement with higher CA was observed in *SA* data: with wider CA aircrafts at trial start are placed further each other, and saccades with bigger

Table 2
Means and standard deviations (italic) for each of the dependent variables and experimental conditions. Mean error rates (ERR) were calculated by dividing number of errors in each experimental condition by number of observations (2 trials in each condition × 17 participants = 34 observations). Subjective workload (SUB) was verbally rated on VOSO scale (from 0 to 10).

DV	CA = 90°					CA = 135°				
	MD (nm)									
	0	1	6	10	12	0	1	6	10	12
NrF	165.3	138.2	100.4	121.1	84.0	122.8	119.6	110.9	86.3	53.3
	(143.2)	*(109.6)*	*(79.1)*	*(114.9)*	*(60.6)*	*(91.6)*	*(93.0)*	*(93.1)*	*(79.6)*	*(31.0)*
FD (ms)	274.8	277.4	264.6	272.9	256.0	264.4	264.4	254.2	252.9	246.2
	(47.8)	*(45.7)*	*(28.9)*	*(35.9)*	*(45.0)*	*(38.9)*	*(35.1)*	*(38.9)*	*(43.8)*	*(35.3)*
NrS	185.1	154.4	114.9	136.0	93.7	146.0	137.3	123.6	96.2	56.6
	(154.0)	*(129.0)*	*(94.8)*	*(133.4)*	*(75.0)*	*(119.8)*	*(109.5)*	*(110.0)*	*(98.2)*	*(32.0)*
SA(°)	4.3	4.4	4.0	3.8	4.6	4.5	4.8	4.7	4.6	5.3
	(0.8)	*(1.1)*	*(0.6)*	*(0.8)*	*(0.9)*	*(0.7)*	*(1.1)*	*(1.0)*	*(1.0)*	*(1.1)*
SD (ms)	42.7	44.1	47.6	46.6	48.5	44.1	43.5	42.7	43.0	44.8
	(19.2)	*(23.8)*	*(28.2)*	*(27.5)*	*(33.1)*	*(25.1)*	*(23.9)*	*(21.0)*	*(22.3)*	*(25.8)*
SPV(°/s)	284.0	303.7	342.2	312.5	309.5	300.0	301.9	305.9	302.1	322.0
	(152.6)	*(181.7)*	*(212.1)*	*(188.5)*	*(196.0)*	*(181.6)*	*(175.5)*	*(183.2)*	*(179.9)*	*(202.6)*
ERR	0.2	0.2	0.6	0.2	0	0.2	0.3	0.1	0	0
	(0.3)	*(0.3)*	*(0.4)*	*(0.4)*	-	*(0.3)*	*(0.3)*	*(0.2)*	-	-
SUB	6.1	5.6	4.9	4.2	3.4	5.7	5.8	3.9	3.4	2.0
	(1.6)	*(1.5)*	*(1.4)*	*(1.8)*	*(1.2)*	*(1.7)*	*(1.4)*	*(2.2)*	*(1.6)*	*(0.9)*
RT(s)	43.1	33.7	26.8	33.2	18.7	34.1	35.4	24.8	18.5	9.9
	(36.4)	*(28.4)*	*(19.0)*	*(31.0)*	*(19.5)*	*(28.4)*	*(25.5)*	*(25.1)*	*(23.7)*	*(6.6)*

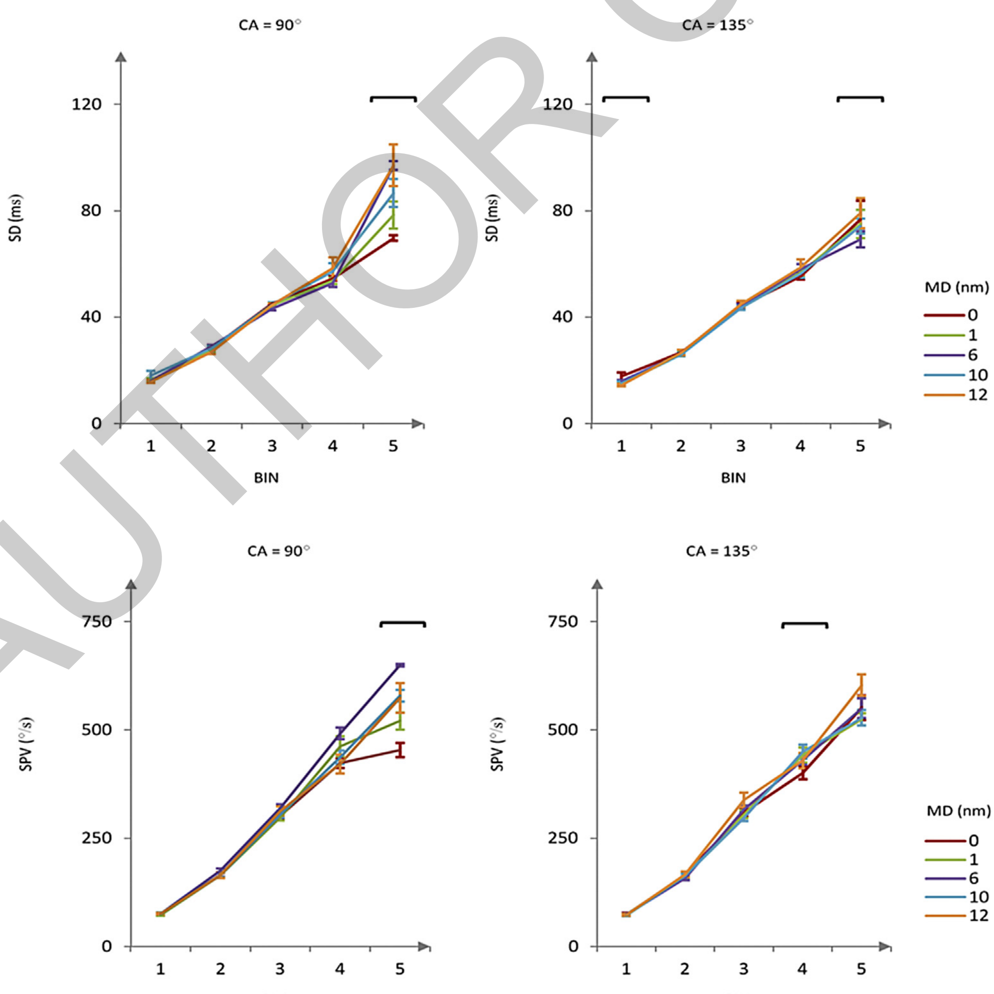

Fig. 5. Saccadic metrics significant interactions. Top panel shows simple effects analyses for CAxMDxBin interaction on *SD*, bottom panel shows the same analysis in relation to *SPV*: CA = 90° on the left and CA = 135° on the right. Vertical bars denote SE. Horizontal lines indicate Bins with significant MD effects.

amplitude need to be performed in order to locate task relevant information, like position and speed.

As to *SD*, bigger values were found for CA = 90° [$F(1, 15) = 16.6$; $p = 0.001$; $\eta^2_p = 0.52$]. Combined *SA* and *SD* results suggest that, on average, saccades were performed more slowly with right-angles. Main effect of MD was not significant, while there was a strong Bin effect [$F(4, 60) = 214.1$; $p < 0.001$; $\eta^2_p = 0.94$], due to main sequence. Several interactions were significant, namely CA × MD [$F = (4, 60) = 3.18$; $p = 0.02$; $\eta^2_p = 0.18$], CA × Bin [$F = (4, 60) = 15.5$; $p = 0.001$; $\eta^2_p = 0.51$], and CA × MD × Bin [$F(16, 240) = 3.76$; $p < 0.012$; $\eta^2_p = 0.20$]. We performed two separate 5 × 5 (MD × Bin) ANOVA in each level of CA as simple effects analysis for the second order interaction. As to CA = 90°, there were significant differences for large saccades (Bin 5), which showed lower *SD* in MD = 0 nm with respect to MD = 6; 12 nm (both $p < 0.001$). The other conflict trials (MD = 1 nm) presented the same trend (both $p = 0.06$). As to CA = 135°, significant differences were found for microsaccades (Bin 1), which showed higher *SD* in MD = 0 nm with respect to MD = 12 nm ($p = 0.01$) and, similarly to CA = 90°, for large saccades (Bin 5), although this effect was weaker ($p = 0.05$), with no significant post hoc comparisons. Therefore, duration of large saccades decreased and that of microsaccades increased with higher cognitive complexity, as in conflict trials. These results further support microsaccades as index of workload and experienced fatigue (Benedetto et al., 2011; Di Stasi et al., 2013a,b).

As to *SPV*, CA main effect was not significant, while MD main effect it was [$F(4, 60) = 6.82$; $p < 0.001$; $\eta^2_p = 0.31$]: planned comparisons in relation to MD = 6 nm showed higher values with respect to MD = 0; 1; and 10 nm (all $p < 0.001$). No other differences were found after post hoc comparisons. Therefore, there was a considerable difference in *SPV* data of conflicts with respect to no conflicts, the latter presenting higher values. Planned comparisons related to MD = 6 nm in Bin 5 showed the same pattern as in MD main effect (all $p < 0.001$): this result was opposed to predictions, although it is a convergent evidence of the higher complexity and cognitive resources consumption of conflicts with respect to no conflicts. Bin effect also was significant [$F(4, 60) = 270.9$; $p < 0.001$; $\eta^2_p = 0.95$], as expected. Significant interactions were MD × Bin [$F(16, 240) = 6.35$; $p < 0.001$; $\eta^2_p = 0.30$] and CA × MD × Bin [$F(16, 240) = 2.67$; $p < 0.04$; $\eta^2_p = 0.15$]. As to the second order interaction, simple effect analysis revealed in CA = 90° trials significant lower *SPV* values of large saccades (Bin 5) for MD = 0 nm with respect to no conflicts (MD = 6; 10; 12 nm), as well as between MD = 1 nm and MD = 6 nm (all $p < 0.001$). As to CA = 135° trials, effects on *SPV* were weaker, and related to Bin 4: MD = 0 nm presented lower values with respect to MD = 10 nm ($p = 0.03$).

Therefore, *SPV* results resembled those of *SD*: large saccades in conflicts had lower values with respect to no conflicts, especially with vertical and horizontal routes. Hypotheses predicted CA effects on saccadic parameters, as geometrical complexity factor that, when increasing, requires oculomotor system to trigger movements of wider amplitude with congruent velocity and duration. This result was substantially confirmed for *SA*, but not for *SD* (which showed higher values with CA = 90°), neither *SPV* (CA main effect not significant). Effects on *SD* and *SPV* involved mainly large saccades with CA = 90° trials. We made a precise prediction on large saccades' *SPV* about a possible slowing down effect for increased complexity, since large saccades already proved to be sensitive to task difficulty (Marchitto et al., 2012). Results about large saccades showed a substantial difference in perceived complexity between conflicts and no conflicts independently of *RT*, whose effects were controlled as covariate, while effects of complexity factors on remaining ocular metrics were strictly dependent on conflict detection times. Large saccades in conflicts showed less burst power with respect to no conflicts, especially when CA = 90°.

These results constitute empirical evidence that *SPV* of large saccades decreased when cognitive complexity increased, as for example when dealing with horizontal and vertical routes. Since both large saccades *SD* and *SPV* decreased with higher complexity, they probably became more precise, with less burst power and reduced positional error, as a task adaptation process would suggest (Gancarz and Grossberg, 1999).

In summary, the geometry complexity factor CA affected as expected *SA*, while effects on *SD* presented a trend opposite to predictions: there was an intrinsic cognitive difficulty with CA = 90°, most probably due to horizontal and vertical flying routes. Perpendicular routes have been reported to be particularly difficult (Boag et al., 2006). The traffic complexity factor MD did not affected fixation parameters, once the effects of *RT* were removed. The expected effects of MD = 6 nm were not encountered, but a clear difference between conflicts and no conflicts emerged after ocular metrics analysis. Trials with the biggest MD difference compared to lateral separation standard of 5 nm (i.e. MD = 12 nm) resulted to be the least complex for conflict detection task. Fixation parameters are further confirmed as reliable metrics for on line workload assessment, especially in more complex traffic scenarios (e.g. with higher traffic loads), in which conflict detection *RTs* for single pairs of aircraft would be hardly partitioned. It is reasonable to conclude that in conflict detection task the most fixated elements are those related to highest cognitive complexity. More importantly, manipulation of MD affected saccadic parameters: large saccades showed shorter *SD* and lower *SPV* in conflicts with respect to no conflicts (especially with MD = 12 nm). Probably, as cognitive complexity augmented, large saccades became more precise, i.e. with shorter duration and weaker burst signal. Therefore, large saccades were affected by traffic and geometry complexity factors, increasing perceived cognitive complexity of the conflict detection task. Increased cognitive complexity augmented local exploration (more saccades with lower amplitude), information processing and thus, consequent workload. All these differences were highly confirmed by subjective and performance data. Next section reports on these findings.

3.2.2. Performance measures

We measured error rates (*ERR*) and response time (*RT*) as performance indices, and analyzed them separately in two repeated measures ANOVAs. As to *ERR*, no errors were made in three conditions, as reported in Table 2, namely when MD = 12 nm in both CA levels, and when MD = 10 nm with CA = 135° level. These conditions had not variance and were not further considered in the analysis. As predicted, they were the easiest conditions for correct conflict detection. Consequently, a 2 × 3 ANOVA was performed.

There were higher *ERR* with CA = 90° [$F(1, 16) = 5.68$; $p = 0.03$; $\eta^2_p = 0.26$], and a significant CAxMD interaction [$F(2, 32) = 10.69$; $p < 0.001$; $\eta^2_p = 0.40$], due to higher *ERR* when MD = 6 nm for CA = 90° with respect to CA = 135°. Therefore, MD = 6 nm was effectively the most error-prone condition, with an intrinsic difficulty deriving from the little difference with lateral separation standard, as predicted. This effect was sensibly amplified in presence of vertical and horizontal routes. Results about *ERR* confirmed predictions.

As to *RT*, there were later responses for CA = 90° [$F(1, 16) = 18.05$; $p = 0.001$; $\eta^2_p = 0.53$]. Main effect of CA on *RT* was not explicitly predicted, but it demonstrated the inherent difficulty of CA = 90° situations, as ocular data already suggested. Moreover, MD main effect was significant [$F(4, 64) = 10.49$; $p < 0.001$; $\eta^2_p = 0.40$]. Planned comparisons revealed that *RT* in MD = 6 nm was higher than in MD = 12 nm ($p = 0.003$), but lower than in MD = 0 nm ($p = 0.049$). Post hoc comparisons revealed higher *RT* values for MD = 0; 1 nm vs. MD = 12 nm ($p < 0.001$ and $p = 0.01$, respectively). There was not significant interaction. Main effects of MD on *RT* are reported in Fig. 6. These results showed that there was a clear effect of MD as cognitive complexity factor that increased *RT* in conflict trials, requiring more cognitive effort and producing later responses, as opposed to no-conflict trials (especially MD = 12 nm), which were easier conditions that required little effort and allowed a very quick response. Furthermore, the particular geometry configuration with CA = 90°, that presented aircraft on vertical and horizontal routes determined a further cognitive cost, observable in performance data, ocular metrics, and subjective ratings, as reported in next section.

3.2.3. Subjective measures

Subjective ratings of workload (*SUB*) were collected after each trial using VOSO scale (Miller, 2001), and were analyzed introducing *RT* as covariate (as for ocular metrics), in order to remove effects on workload ratings due to longer times for correct conflict detection. Trials with CA = 90° generated significant higher workload ratings [$F(1,15) = 5.94$; $p = 0.028$; $\eta^2_p = 0.28$]. Main effect of MD was also significant [$F(4,60) = 6.38$; $p < 0.001$; $\eta^2_p = 0.30$], as reported in Fig. 6. Planned comparisons revealed that *SUB* in MD = 6 nm was higher than in MD = 12 nm ($p < 0.001$), but lower than in MD = 0; 1 nm ($p < 0.008$ and $p = 0.024$, respectively). Post hoc comparisons further revealed significant differences between conflicts and no conflicts: MD = 0 nm vs. MD = 10; 12 nm ($p = 0.008$ and $p < 0.001$, respectively); MD = 1 nm vs. MD = 10; 12 nm ($p = 0.046$ and $p < 0.001$, respectively). Interaction was not significant. In accord with ocular and performance variables, *SUB* results showed a substantial difference between conflicts and no conflicts and, among no conflicts, between MD = 12 nm and remaining levels.

These results further support the interpretation of ocular metrics results in terms of cognitive workload generated by the conflict detection task. Fixation and saccadic parameters captured the higher cognitive complexity of conflicts with respect to no conflicts (MD = 10; 12 nm), and of vertical and horizontal routes with respect to obtuse angle of convergence (CA = 135°). The most error-prone experimental condition was MD = 6 nm, as predicted. Nevertheless, conflicts trials required longer visual exploration and information collection, and generated higher subjective workload. Ocular metrics confirmed these trends, especially peak velocity of large saccades.

3.3. Adopted strategies in conflict detection task

These considerable quantitative differences might underlie qualitatively different cognitive processes for trajectory prediction and lateral separation estimation. As to conflicts, trajectory prediction was probably based on movement observation and aircraft position recurrent processing. Conversely, no conflicts presented fewer fixations (due to little *RT*), so that a strategy that allowed quick prediction after rapid computation of aircraft alphanumerical data and distance to convergence point was most probably adopted. Literature reports on several strategies for conflict detection task (Nunes and Mogford, 2003). It is normally assumed that the appropriate implementation of a strategy to a particular situation assures higher levels of performance.

According to Xu and Rantanen (2003), conflict detection task consists of two phases, performed not necessarily in a fixed order: a relative judgment task (RJ) to estimate which aircraft will reach the convergence point first, and a prediction-motion (PM) task to calculate when the aircraft will be at convergence point and to decide about their possible conflict situation after lateral separation estimation. In some special cases RJ is particularly easy, as when aircraft with the same distance from convergence point and same speed are rapidly estimated to collide at intersection point, i.e. to be in conflict. As to PM, some available strategies that simplify conflict detection are reported. For example, times to convergence point can be calculated for each aircraft in any instant by dividing distance to convergence point and speed, i.e. applying distance-to-speed ratio strategy. As a consequence, it is possible to make a prediction on lateral separation by projecting aircraft position once its motion has been included in a cognitive projection model, i.e. the cognitive simulation of flights future progress. This strategy is considerably helpful in conflict decision task when estimated times are very similar or considerably different: in fact, it would be relatively easy to estimate a conflict and a no conflict, respectively. Conversely, a medium difference in times to intersection point could make the estimation of lateral separation more difficult: in this case, distance-to-speed ratio strategy might be insufficient for correct conflict detection. Another strategy for PM is cognitive motion extrapolation (CME). This strategy does not imply time calculations, and it is based on motion representation in spatial dimension, leaving time dimension implicit. It assumes that the cognitive model of object's motion used for projection is explicitly spatial (DeLucia and Liddell, 1998). A recurrent processing of aircraft updated position is necessary for a correct spatial representation of motion.

It is normally assumed that the strategies that guarantee optimal result with minimal cognitive effort are applied whenever possible and that shifts towards more demanding strategies occur whether the former option is unfeasible. The difference between

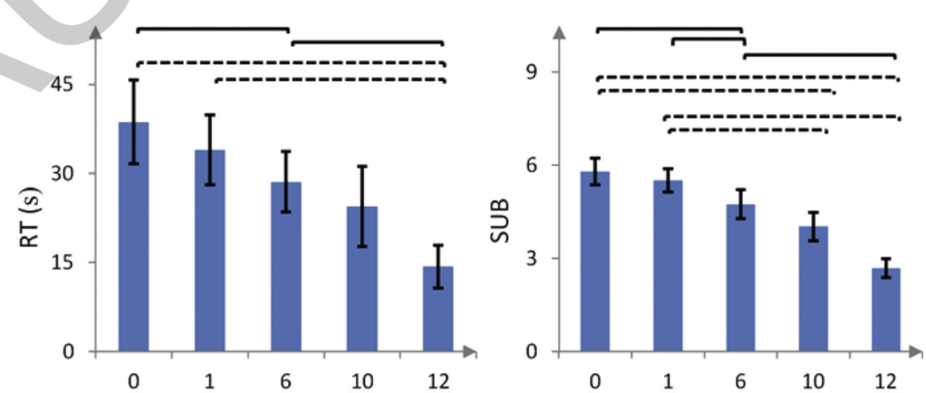

Fig. 6. Main effects of MD on performance (*RT*) and subjective ratings of workload (*SUB*). Vertical bars denote SE. Horizontal lines indicate significant planned comparisons; dotted lines refer to post hoc comparisons.

conflicts and no conflicts evidenced in data analysis could be a consequence of two different strategies that might have been adopted in no conflicts and conflicts. As to no conflicts, quick responses (after few fixations and saccades) might have been the result of adopting a ratio-based strategy after processing aircraft start position and flight data. As to conflict trials, a prediction-motion strategy based on spatial representation of motion was most probably adopted (e.g. CME): since flight data were kept constant through time, the higher number of fixations and saccades (predicted by late response time) were related to continuous aircraft position updates.

We suggested that there might have been a strategy shift from a ratio-based to a prediction-motion strategy in conflicts trials, due to the ratio-based strategy being unfeasible in such more uncertain situations. A quick decision on conflict presence might have been impeded by aircraft differences in speed and distance to convergence point, due to the difficulty of transforming the estimated difference in time of arrival into a reliable estimation of lateral separation. Conflicts were more difficult, demanding, and effortful, and this strategy shift might have contributed to final workload assessment. To support this conclusion, we computed speed and distance differences for each experimental condition at trial start, as reported in Table 3. These differential values were the consequence of traffic dynamic calculations for the experimental scenarios, after setting an MD level, fixing a time-to-MD value (7 min in this study), and choosing a speed value for one of the two aircrafts.

These differential values might have determined a strategy shift in conflict detection task. No conflicts (MD = 6; 12 nm) with CA = 135° were easily answered probably because distance differences were considerable and speed differences little: consequently, RJ and PM were relatively easy. Similarly, when MD = 10 nm with CA = 135°, a considerable speed difference simplified the task, also because the aircraft closer to intersection was faster: the distance difference with convergence point would have increased through time, further separating aircrafts in the lateral dimension. No conflicts with MD = 6 nm and CA = 90° presented a moderate distance difference and a little speed difference. Participants tended to erroneously interpret this situation as a conflict, since difference in times of arrival was probably estimated as close to zero. Conflicts also presented little distance differences and might have made particularly difficult the application of a ratio-based strategy. In fact, in almost all conflicts there was a considerable speed difference at trial start, and the aircraft closer to intersection was slower. Consequently, the distance difference with convergence point would have decreased as effect of speed: calculations of time to convergence point should have been considerably precise. In this case, RJ resulted quite difficult, since it should have been decided whether the quicker aircraft would have had time to arrive before the closer one. When facing this situation, participants might have switched to a strategy based on position recurrent computation and projection, instead of relying on difficult and error-prone calculations.

Table 3
Differential speeds in knots (kn) and differential distances to convergence point in nautical miles (nm) at trial start for each of the experimental conditions. Marked speed differences (*) mean that the aircraft closer to convergence point was slower than the other aircraft.

CA (°)	MD (nm)	Speed difference (kn)	Distance difference (nm)
90	0	60*	10
	1	20	0
	6	10	10
	10	20	15
	12	20	20
135	0	60*	10
	1	50*	12
	6	10*	23
	10	50	28
	12	10*	43

It is acknowledged in ATC literature that differential (or relative) values of speed, distance and flight level are central information for correct conflict detection (Leplat and Bisseret, 1966; Ahlstrom, 2005). Task performance is better when differential values are considerable high or close to zero (Kimball et al., 1973). Next section presents the conclusions drawn after quantitative analysis of ocular metrics discussed in this section.

3.4. Summary of important results

This section resumes the most important outcomes of the study. The correlational study showed that ocular metrics were associated to experienced cognitive workload: first, performance accuracy correlated positively with the number of fixations and saccades; second, subjective ratings correlated positively with saccadic parameters (duration and peak velocity), suggesting that oculomotor system effort is taken into account when rating subjective workload.

When analyzing data as a function of the two complexity factors (convergence angle and minimum distance at closest approach), a great difference emerged between conflict and no conflict trials. The former were more time-consuming and determined higher workload, while the latter required less effort for a correct decision. The presence of perpendicular routes (vertical and horizontal) raised complexity and further increased workload in conflict trials. In particular, both large saccades and microsaccades presented significant differences in conflict and in no conflict situations. As to large saccades, they showed decreased duration and peak velocity (burst power), and thus higher precision. Therefore, a form of cognitive control might have been employed on the realization of such neuromotor movements when higher complexity was perceived. Large saccades are the most expensive in terms of neural resources and they are performed by specific motor neurons firing at their maximum rate. Conversely, microsaccades presented increased durations. Fixation parameters (number and duration) increased accordingly when conflict detection task was more time-consuming (i.e. in conflicts), demonstrating longer processing times and bigger investment of cognitive resources in such situations.

The strong quantitative differences in ocular metrics between conflicts and no conflicts suggested that a qualitative difference might exist. The specific strategy adopted for detection task could have been different. We proposed an explanation of such differences. Participants first computed aircraft flight metrics (speed and position, i.e. distance to convergence center), with the aim of calculating the time to convergence point for the two aircraft, i.e. by employing a distance-to-speed ratio strategy. The latter could have been revealed as unfeasible, due to a difference in time of arrival considered insufficient to clearly exclude conflict presence, or even due to calculations gauged as difficult and error-prone. In that case the situation was perceived as more complex, and a shift towards a more reliable (and more demanding) strategy was needed. Consequently, more fixations and saccades were performed, and response time increased. Moreover, duration of microsaccades increased, and peak velocity and duration of large saccades decreased, meaning that they became more precise and that a sort of cognitive control might have been applied over these movements. All these variations in ocular metrics are in accord with the employment of a motion-based strategy for conflict detection task. In order to predict motion and project aircraft position into the future, aircraft moving forward must be used as basis for a cognitive projection model: this implied that longer and more frequent fixations were devoted to processing of aircraft position updates and, in parallel, that large transitions saccades performed between the two aircraft became as precise as possible. Although good performance level was maintained in conflict trials with respect to no conflict trials, greater workload was experienced and measured.

We considered that flight metric differential values at trial start determined whether a ratio-based strategy was retained reliable or whether a safer (motion-based) strategy might have been adopted. Future research should focus on the effects of flight differential values on conflict detection strategy and related experienced workload.

4. Conclusion

This study explored ocular behaviour during simulated ATC conflict detection task. Ocular metrics were registered together with performance and subjective ratings of workload, in order to capture workload variations as a function of conflict detection task complexity. We manipulated two complexity factors in order to create different levels of task difficulty, i.e. convergence angle (CA), and minimum distance at closest approach (MD). Triangulation method proved to be very fruitful for multimodal workload measurement, supporting the validity of ocular metrics as sensitive online workload indices.

We hypothesized that the most complex situations were those with MD close to separation standard (i.e. MD = 6 nm). Hypotheses were confirmed only partially: MD = 6 nm was the most error-prone condition, but conflict trials were more time-consuming and rated as the most demanding scenarios. On the contrary, the remaining no conflict trials (MD = 10; 12 nm) resulted in lower subjective workload levels, more rapid responses, and little visual processing. Moreover, vertical and horizontal routes acted unexpectedly as further complexity factor. Ocular metrics were able to capture higher cognitive complexity. Effects on fixation metrics were absent once removed contemporary influence of response time. This substantially confirms the usefulness of fixations metrics recording for cognitive workload when it is difficult to isolate response time to a single aircraft pair, e.g. with higher traffic loads. As to saccadic metrics, we observed that large saccades ($SA > 15°$) were affected by MD factor, showing shorter duration and lower peak velocity in conflicts. It seemed that large saccades were more precise, i.e. more rapid and with weaker burst power for conflict trials, which were the most difficult situations. Velocity parameters of large saccades were sensitive to task difficulty, in accord with several recent findings (Di Stasi et al., 2010; 2011; 2013a,b; Marchitto et al., 2012).

The substantial difference in cognitive complexity between conflicts and no-conflicts was interpreted as the probable adoption of different strategies in the conditions. In no conflicts (especially with MD = 12 nm) it was suggested that a ratio-based strategy was applied, at least at a rough level, that allowed quick estimations of differential times of arrival to convergence point. In conflict trials, a strategy shift might have occurred, by adopting a motion-prediction strategy. Fixation metrics and saccadic parameters of large saccades captured this cognitive cost. We concluded that participants first checked for differential speeds and distances to convergence point, and then made a quick decision if a considerable difference in times to convergence point was estimated, as in no conflicts. Conversely, the particular combination of differential speeds and distances to intersection in conflicts probably determined the adoption of a motion-based strategy, which required more cognitive effort with respect to a ratio-based strategy and increased decision times. As a consequence, conflicts were perceived as more complex and demanding.

Ocular metrics allowed for online workload monitoring and are a promising method for strategy characterization (Hunter and Parush, 2009). Workload modeling as well as its management would benefit from the refinement of temporal and spatial analysis of ocular indices, for example in ATCO training (Kang and Landry, 2014). The incorporation of quantitative and qualitative assessment of visual behaviour in novice ATCOs learning process might enhance the identification of high workload scenarios, task relevant information and adopted strategy for traffic management. Ocular metrics measurement for online workload assessment could also be successfully applied in the field of adaptive technology, to identify overload episodes and to support controller by automating ATC subtasks. The manipulation of differential (or even relative) speed, distance to convergence point, and also flight level might help further in identifying the conditions for strategies adoption, by means of scanpath visualization. For example, the use of specific Areas of Interests (AOI) could help us to better address cognitive complexity in conflict detection task (Kang and Landry, 2010).

Acknowledgements

This work was supported by Spanish Ministry of Economy and Competitiveness, PSI2012-39246. We are very grateful to all the researchers and colleagues who have reviewed and commented on this manuscript at various stages. The authors have no conflicts of interests to declare.

References

Ahlstrom, U., 2005. Work domain analysis for air traffic controller weather displays. J. Saf. Res. 36 (2), 159–169. http://dx.doi.org/10.1016/j.jsr.2005.03.00.

Ahlstrom, U., Friedman-Berg, F.J., 2006. Using eye movement activity as a correlate of cognitive workload. Int. J. Ind. Ergon. 36 (7), 623–636. http://dx.doi.org/10.1016/j.ergon.2006.04.002.

Alonso, R., Causse, M., Vachon, F., Parise, R., Dehais, F., Terrier, P., 2013. Evaluation of head-free eye tracking as an input device for air traffic control. Ergonomics 56 (2), 246–255. http://dx.doi.org/10.1080/00140139.2012.744473.

Athènes, S., Averty, P., Puechmorel, S., Delahaye, D., Collet, C., 2002. ATC complexity and controller workload: trying to bridge the gap. In: Proceedings of the International Conference on HCI in Aeronautics, pp. 56–60.

Ayaz, H., Willems, B., Bunce, B., Shewokis, P.A., Izzetoglu, K., Hah, S., Deshmukh, A., Onaral, B., 2010. Cognitive Workload Assessment of Air Traffic Controllers Using Optical Brain Imaging Sensors. In: Marek, T., Karwowski, W., Rice, V. (Eds.), Advances in Understanding Human Performance: Neuroergonomics, Human Factors Design, and Special Populations. CRC Press, Taylor & Francis Group, pp. 21–32. http://dx.doi.org/10.1201/EBK1439835012-c3.

Bahill, A.T., Stark, L., 1975. Overlapping saccades and glissades are produced by fatigue in the saccadic eye movement system. Exp. Neurol. 48 (1), 95–106.

Benedetto, S., Pedrotti, M., Bridgeman, B., 2011. Microsaccades and exploratory saccades in a naturalistic environment. J. Eye Mov. Res. 4 (2), 1–10.

Benedetto, S., Carbone, A., Drai-Zerbib, V., Pedrotti, M., Baccino, T., 2014. Effects of luminance and illuminance on visual fatigue and arousal during digital reading. Comput. Hum. Behav. 41, 112–119. http://dx.doi.org/10.1016/j.chb.2014.09.023.

Boag, C., Neal, A., Loft, S., Halford, G., 2006. An analysis of relational complexity in an air traffic control conflict detection task. Ergonomics 14, 1508–1526. http://dx.doi.org/10.1080/00140130600779744.

Callan, D., 1998. Eye movement relationships to excessive performance error in aviation. In: Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting. Human Factors and Ergonomics Society, Santa Monica, CA, pp. 1132–1136. http://dx.doi.org/10.1177/154193129804201516.

Cazzoli, D., Antoniades, C.A., Kennard, C., Nyffeler, T., Bassetti, C.L., Müri, R.M., 2014. Eye movements discriminate fatigue due to chronotypical factors and time spent on task–a double dissociation. PLoS One 9 (1), e87146. http://dx.doi.org/10.1371/journal.pone.0087146.

Cummings, M.L., Tsonis, C., 2005, September. Deconstructing complexity in air traffic control. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 49, No. 1. Sage Publications, pp. 25–29. http://dx.doi.org/10.1177/154193120504900107.

DeLucia, P.R., Liddell, G.W., 1998. Cognitive motion extrapolation and cognitive clocking in prediction motion tasks. J. Exp. Psychol. Hum. Percept. Perform. 24 (3), 901. http://dx.doi.org/10.1037/0096-1523.24.3.901.

Di Stasi, L.L., Antolí, A., Cañas, J.J., 2011. Main sequence: an index for detecting mental workload variation in complex tasks. Appl. Ergon. 42 (6), 807–813. http://dx.doi.org/10.1016/j.apergo.2011.01.003.

Di Stasi, L.L., Catena, A., Cañas, J.J., Macknik, S.L., Martinez-Conde, S., 2013a. Saccadic velocity as an arousal index in naturalistic tasks. Neurosci. Biobehav. Rev. 37 (5), 968–975. http://dx.doi.org/10.1016/j.neubiorev.2013.03.011.
Di Stasi, L.L., McCamy, M.B., Catena, A., Macknik, S.L., Cañas, J.J., Martinez-Conde, S., 2013b. Microsaccade and drift dynamics reflect mental fatigue. Eur. J. Neurosci. 38, 2389–2398. http://dx.doi.org/10.1111/ejn.12248.
Fitts, P.M., Jones, R.E., Milton, J.L., 1950. Eye movements of aircraft pilots during instrument-landing approaches. Aeronaut. Eng. Rev. 9 (2), 24–29.
Fothergill, S., Loft, S., Neal, A., 2009. ATC-labAdvanced: an air traffic control simulator with realism and control. Behav. Res. Methods 41 (1), 118–127. http://dx.doi.org/10.3758/BRM.41.1.118.
Gancarz, G., Grossberg, S., 1999. A neural model of saccadic eye movement control explains task-specific adaptation. Vis. Res. 39 (18), 3123–3143. http://dx.doi.org/10.1016/S0042-6989(99)00049-8.
Goldberg, J.H., Kotval, X.P., 1998. Eye movement-based evaluation of the computer interface. Adv. Occup. Ergon. Saf. 529–532.
Goldberg, J.H., Kotval, X.P., 1999. Computer interface evaluation using eye movements: methods and constructs. Int. J. Ind. Ergon. 24, 631–645. http://dx.doi.org/10.1016/S0169-8141(98)00068-7.
Ha, C.H., Kim, J.H., Lee, S.J., Seong, P.H., 2006. Investigation on relationship between information flow rate and mental workload of accident diagnosis tasks in NPPs. Nucl. Sci. IEEE Trans. 53 (3), 1450–1459. http://dx.doi.org/10.1109/TNS.2006.874189.
Henderson, J.M., 2003. Human gaze control during real-world scene perception. Trends Cogn. Sci. 7 (11), 498–504. http://dx.doi.org/10.1016/j.tics.2003.09.006.
Henderson, J.M., 2007. Regarding scenes. Curr. Dir. Psychol. Sci. 16 (4), 219–222. http://dx.doi.org/10.1111/j.1467-8721.2007.00507.x.
Hillburn, B., 2004. Cognitive complexity in air traffic control: a literature review. EEC Note 4 (04).
Histon, J.M., Hansman, R.J., Gottlieb, B., Kleinwaks, H., Yenson, S., Delahaye, D., Puechmorel, S., 2002. Structural considerations and cognitive complexity in air traffic control. In: Digital Avionics Systems Conference, 2002. Proceedings. The 21st, vol. 1. IEEE, p. 1C2-1. http://dx.doi.org/10.1109/DASC.2002.1067894.
Hunter, A.C., Parush, A., 2009. Using eye movements to uncover conflict detection strategies. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 53, No. 22. Sage Publications, pp. 1729–1733. http://dx.doi.org/10.1177/154193120905302215.
ICAO, International, 2012. Civil aviation organization. Forecasts of scheduled passenger traffic. Retrieved on November, 6, 2014. http://www.icao.int/sustainability/pages/eap_fp_forecastmed.aspx.
Imants, P., de Greef, T., 2011, August. In: Using Eye Tracker Data in Air Traffic Control. In Proceedings of the 29th Annual European Conference on Cognitive Ergonomics. ACM, pp. 259–260. http://dx.doi.org/10.1145/2074712.2074769.
Jacob, R.J., Karn, K.S., 2003. Eye tracking in human-computer interaction and usability research: ready to deliver the promises. In: The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research, pp. 573–603.
Just, M.A., Carpenter, P.A., 1976. Eye fixations and cognitive processes. Cogn. Psychol. 8 (4), 441–480. http://dx.doi.org/10.1016/0010-0285(76)90015-3.
Kallus, K.W., Van Damme, D., Dittman, A., 1999. Integrated Job and Task Analysis of Air Traffic Controllers: phase 2. Task Analysis of En-route Controllers (European Air Traffic Management Programme Rep. No. HUM. ET1. ST01.1000-REP-04). Belgium, Brussels (Eurocontrol).
Kang, Z., Landry, S.J., 2010, September. Capturing and analyzing visual groupings of multiple moving targets in an aircraft conflict detection task using eye movements. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 54, No. 23. SAGE Publications, pp. 1906–1910. http://dx.doi.org/10.1177/154193121005402301.
Kang, Z., Landry, S.J., 2014. Using scanpaths as a learning method for a conflict detection task of multiple target tracking. Hum. Factors J. Hum. Factors Ergon. Soc. 56 (6), 1150–1162. http://dx.doi.org/10.1177/0018720814523066.
Kimball, K.A., Hofmann, M.A., Nossaman, R.O., 1973. Differential velocity and time prediction of motion. Percept. Mot. Ski. 36 (3), 935–945.
Langan-Fox, J., Sankey, M.J., Canty, J.M., 2009. Human factors measurement for future air traffic control systems. Hum. Factors J. Hum. Factors Ergon. Soc. 51 (5), 595–637. http://dx.doi.org/10.1177/0018720809355278.
Leplat, J., 1978. The factors affecting workload. Ergonomics 21, 143–149.
Leplat, J., Bisseret, A., 1966. Analysis of the processes involved in the treatment of information by the air traffic controller. Control 5 (1), 13–22.
Li, L., Cho, H., Hansman, R.J., Palacios, R., 2010. Aircraft-based complexity assessment for radar controllers in the multi-sector planner experiment. In: 10th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, pp. 1–9. http://dx.doi.org/10.2514/6.2010-9005.
Lin, Y., Zhang, W.J., Watson, L.G., 2003. Using eye movement parameters for evaluating human–machine interface frameworks under normal control operation and fault detection situations. Int. J. Hum.-Comput. Stud. 59 (6), 837–873. http://dx.doi.org/10.1016/S1071-5819(03)00122-8.
Loft, S., Bolland, S., Humphreys, M.S., Neal, A., 2009. A theory and model of conflict detection in air traffic control: incorporating environmental constraints. J. Exp. Psychol. Appl. 15 (2), 106. http://dx.doi.org/10.1037/a0016118.
Lokhande, K., Reynolds, H.J.D., 2012, September. Cognitive workload and visual attention analyses of the air traffic control tower flight data manager (TFDM) prototype demonstration. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 56, No. 1. Sage Publications, pp. 105–109. http://dx.doi.org/10.1177/1071181312561042.
Majumdar, A., Ochieng, W.Y., 2007. Air traffic control complexity and safety: framework for sector design based on controller interviews of complexity factors. Transp. Res. Rec. J. Transp. Res. Board 2007 (1), 70–80. http://dx.doi.org/10.3141/2007-09.
Marchitto, M., Di Stasi, L.L., Cañas, J.J., 2012. Ocular movements under taskload manipulations: influence of geometry on saccades in air traffic control simulated tasks. Hum. Factors Man. 22, 407–419. http://dx.doi.org/10.1002/hfm.20320.
Martin, C., Cegarra, J., Averty, P., 2011. Analysis of mental workload during en-route air traffic control task execution based on eye-tracking technique. In: Engineering Psychology and Cognitive Ergonomics. Springer Berlin Heidelberg, pp. 592–597.
Miller, S., 2001. Workload Measures. National Advanced Driving Simulator. The University of Iowa, Oakland, IA.
Mogford, R.H., Guttman, J.A., Morrow, S.L., Kopardekar, P., 1995. The Complexity Construct in Air Traffic Control: a Review and Synthesis of the Literature. CTA Inc, McKee city NJ.
Nunes, A., Mogford, R.H., 2003, October. Identifying controller strategies that support the 'picture'. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 47, No. 1. SAGE Publications, pp. 71–75. http://dx.doi.org/10.1177/154193120304700115.
Parasuraman, R., Sheridan, T.B., Wickens, C.D., 2008. Situation awareness, mental workload, and trust in automation: viable, empirically supported cognitive engineering constructs. J. Cogn. Eng. Decis. Mak. 2 (2), 140–160. http://dx.doi.org/10.1518/155534308X284417.
Pawlak, W.S., Brinton, C.R., Crouch, K., Lancaster, K.M., 1996, July. A framework for the evaluation of air traffic control complexity. San Diego, CA. In: Proceedings of the AIAA Guidance Navigation and Control Conference. http://dx.doi.org/10.2514/6.1996-3856.
Poole, A., Ball, L.J., 2005. Eye tracking in human-computer interaction and usability research: current status and future. In: Ghaoui, C. (Ed.), Encyclopedia of Human-computer Interaction. Pennsylvania: Idea Group, Inc, pp. 211–219.
Remington, R.W., Johnston, J.C., Ruthruff, E., Gold, M., Romera, M., 2000. Visual search in complex displays: factors affecting conflict detection by air traffic controllers. Hum. Factors 42, 349–366. http://dx.doi.org/10.1518/001872000779698105.
Schmidt, D., Abel, L.A., Dell'Osso, L.F., Daroff, R.B., 1979. Saccadic velocity characteristics: intrinsic variability and fatigue. Aviat. Space Environ. Med. 50 (4), 393–395.
Stankovic, S., Raufaste, E., Averty, P., 2008. Determinants of conflict detection: a model of risk judgments in air traffic control. Hum. Factors 50, 121–134. http://dx.doi.org/10.1518/001872008X250584.
Di Stasi, L.L., Marchitto, M., Antolí, A., Baccino, T., Cañas, J.J., 2010. Approximation of on-line mental workload index in ATC simulated multitasks. J. Air Transp. Manag. 16 (6), 330–333. http://dx.doi.org/10.1016/j.jairtraman.2010.02.004.
Weiland, M.Z., Roberts, D.M., Fine, M.S., Caywood, M.S., 2013, September. Real time research methods monitoring air traffic controller workload during simulation studies using electroencephalography (EEG). In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 57, No. 1. SAGE Publications, pp. 1615–1619. http://dx.doi.org/10.1177/1541931213571359.
Wierwille, W.W., Eggemeier, F.T., 1993. Recommendations for mental workload measurement in a test and evaluation environment. Hum. Factors 35 (2), 263–281. http://dx.doi.org/10.1177/001872089303500205.
Wilson, G.F., Russell, C.A., 2003. Operator functional state classification using multiple psychophysiological features in an air traffic control task. Hum. Factors J. Hum. Factors Ergon. Soc. 45 (3), 381–389. http://dx.doi.org/10.1518/hfes.45.3.381.27252.
Xu, X., 2003, October. Conflict detection with cockpit display of traffic information: what is it, what has been found, and what needs to be done?. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 47, No. 1. SAGE Publications, pp. 194–198. http://dx.doi.org/10.1177/154193120304700141.
Xu, X., Rantanen, E.M., 2003. Conflict detection in air traffic control: a task analysis, a literature review, and a need for further research. In: Jensen, R.S. (Ed.), Proceedings of the 12th International Symposium on Aviation Psychology. Wright State University Press, Dayton, OH, pp. 1289–1295.
Xu, X., Rantanen, E.M., 2007. Effects of air traffic geometry on pilots' conflict detection with cockpit display of traffic information. Hum. Factors J. Hum. Factors Ergon. Soc. 49 (3), 358–375. http://dx.doi.org/10.1518/001872007X197008.