Vision Research 193 (2022) 107977

Contents lists available at ScienceDirect

Vision Research

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

Influence of mental workload on motion perception: A direct comparison of luminance-based and contrast-based stimuli

Julien Audiffren *, Jean-Luc Bloechle, Jean-Pierre Bresciani

Control and Perception laboratory, University of Fribourg, Switzerland

ARTICLE INFO

Keywords:
Second order motion
Feature tracking
Motion sensors
Cognitive workload

ABSTRACT

In order to study the impact of increased mental workload on motion detection, twenty-four observers performed a motion discrimination task in which they had to detect odd moving patches. Two types of moving patches were used, namely luminance-based and contrast-based patches. For both types of patches, the motion discrimination task was performed with and without an additional N-Back task aimed at increasing the mental workload. The dual task decreased discrimination performance for both types of patches, but the difference was significantly larger for contrast-based patches, i.e., for second-order motion stimuli, both as an absolute and relative increment. This suggests that motion discrimination requires larger cognitive resources for contrast-based than for luminance-based stimuli, thereby hinting at the higher complexity of the cognitive mechanisms underlying second-order motion detection.

1. Introduction

Visual motion perception plays a crucial role in many human activities, such as navigating around an environment, driving, or playing sports. Visual motion perception mostly derives from the optic flow, which is the pattern of local motion fields that results from the relative displacement between surrounding objects and our eyes. The properties of local motion fields can vary significantly (translation, rotation, etc), and are generally regrouped in two classes depending on their nature: first-order signals, which derive from changes in luminance, and second-order signals, where the changes are related to another property of the field, such as contrast (Cavanagh and Mather, 1989). In line with this, the study of the mechanisms underlying motion perception in humans has revealed the existence of two different yet complementary systems to detect visual motion (Derrington and Badcock, 1985). On the one hand, motion sensors are a first-order motion filtering system (Braddick, 1974). It is commonly accepted that motion sensors are local, orientation selective, low-level sensors (Adelson and Bergen, 1985) that cover all areas of the visual field (Morrone et al., 1995). Motion sensors have notably been shown to efficiently detect first-order motion (i.e. luminance based), even with very short exposure time (lower than 100 ms) (Derrington et al., 1993), and to be susceptible to after-effect illusions (see e.g. (Wohlgemuth, 1911; Anstis et al., 1998)). On the other hand, the feature tracking system is a higher-level system that can track a significantly wider range of features over time (Ullman, 1979). Among other things, feature tracking has been shown to operate significantly slower than motion sensors (Derrington et al., 1992; Derrington et al., 1993), and to be able to track motion in the presence of perturbations such as blank intervals (Georgeson and Shackleton, 1989).

Most importantly for this work, feature tracking has been suggested to involve higher level processes, thus being more resource-demanding than motion-sensors-based motion detection (Cavanagh, 1992; Allen and Derrington, 2000; Ashida et al., 2001; Derrington et al., 2004). Specifically, the detection of second-order motion seems to require more cognitive resources, in particular attentional resources (Ashida et al., 2001). For instance, Allen and colleagues (Allen and Derrington, 2000) noted that: “Observers reported that they often felt that they were checking each patch in turn when presented with the contrast-defined patterns whereas the differences between the luminance-defined patterns were immediately obvious.” Ashida and colleagues (Ashida et al., 2001) also noted that slow and effort-full visual search was required to perceive local second-order motion. These authors concluded that this type of visual search may heavily tax attentional resources. In line with this, the current study was designed to test the impact of Mental Workload on motion perception. In particular, we directly compared the effect of Mental Workload on first-order and second-order motion perception.

Mental workload (MWL), also called cognitive workload, is a

* Corresponding author.
E-mail address: julien.audiffren@unifr.ch (J. Audiffren).

https://doi.org/10.1016/j.visres.2021.107977
Received 5 July 2021; Received in revised form 21 October 2021; Accepted 15 November 2021
Available online 13 December 2021

complex, multidimensional notion (Young et al., 2015) for which there is no unified theory to this day. In the following, we use MWL as defined by Oviatt and colleagues, i.e., as a subjective physiological experience that results from the interplay between the available cognitive resources and the cognitive demands of the task (we kindly refer the reader to Oviatt et al. (2018) for a more in-depth discussion about the definition of MWL). MWL is closely related to attention and perception (Kantowitz, 1987). In particular, an increased MWL has been associated to a decrease in overall performance, notably for tasks that require shared resources (see for instance (Young et al., 2015)).

The goal of this study was to determine if, as indirectly suggested by Ashida and colleagues (Ashida et al., 2001), second-order motion sensing indeed requires more cognitive resources, and if so, to which extent the presence of a dual task aimed at manipulating the MWL impacts motion perception. To test that, participants performed a motion discrimination task with two types of visual stimuli, namely luminance-based and contrast-based moving patches. The luminance-based stimuli were constructed using moving gratings designed to stimulate first-order motion sensors. The contrast-based stimuli were constructed using moving beats designed to stimulate second-order motion perception. Both types of visual stimuli were presented either alone (single task condition) or in a dual task setting in which a MWL-increasing task had to be performed in parallel with the motion discrimination task (dual task condition). To manipulate the MWL, we used the N-Back task (Obinata et al., 2008), which has been shown to induce an increased cognitive workload (Soveri et al., 2017) and is a staple of neuroimaging studies focusing on working memory (Braver et al., 1997). The N-Back task also has the advantage of relying on auditory stimuli, resulting in less interference between the main visual task (here motion perception) and the secondary task (Wickens, 2002; Wickens, 2008). Given the observations made in previous studies, our assumption was that an increased MWL should strongly impact second-order motion sensing but have little effect on first-order motion sensing. In particular, we reasoned that an impact of the dual task setting on the perception of second-order motion would hint at a heavier use of cognitive resources during feature tracking, notably as compared to first-order motion sensing.

2. Method

2.1. Participants

Twenty-four (24) participants were included in the study (average age: 27 years, standard deviation 4.5 y., 14 males). All participants had normal or corrected-to-normal vision, and they were naïve as to the purpose of the research. The study was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and approved by the Ethics Committee of the University of Fribourg. Participants had the option to withdraw from the study at any time without penalty and without having to give a reason.

2.2. Apparatus

The experiment took place in a dimly lit, sound-attenuated room. The visual stimuli were presented in a full screen display window on a ViewPixx display (Resolution 1920 × 1200, 55.88 cm diagonal, 100 Hz, no latency). The display was tuned to provide an average luminance of 50 $cd.m^{-2}$ during the experiment. The participant was positioned 70 cm in front of the screen. The visual stimuli were created using the Psychopy v2.3 (Python 3.7.6) library (Peirce et al., 2019), and the display was adapted using the provided calibration tools and a luxmeter. The audio stimuli were realised using two identical small loudspeaker rectangles (10 cm in diagonal) that were positioned 15 cm to either side of the screen, at the same distance from the participant. The audio stimuli consisted in a sequence of numbers ranging from 1 to 5. A number was presented every two seconds, and each number was presented for 0.5 s at a volume of 70 dB (as measured from the participant's head position).

2.3. Visual Stimuli

We used two different types of local-motion visual stimuli: a luminance-based optic flow, and a contrast-based optic flow, similarly to (Allen and Derrington, 2000). These stimuli were expansion patterns, aimed at stimulating first- and second-order motion perception, respectively (Derrington et al., 2004). Stimuli were restricted to the spatial and temporal frequency range that matched previously observed sensitivity to contrast-defined motion (Derrington and Badcock, 1985).

The local motion pattern were constructed as follows: four patches (discs, 5° of diameter) were positioned symmetrically around the center of the screen, at a distance of 3.7°. Each patch contained a pattern moving toward the center of the display. At each iteration of adaptive procedure (see the Procedure subsection below), one patch, randomly selected, had its motion inverted, i.e., the pattern was moving away from the center. Examples of the visual stimuli are shown in Fig. 1. The motion patterns were either luminance-defined or contrast-defined, depending on the condition being tested. Aside from these differences, the parameters of these patterns were chosen to reproduce the stimuli used in (Allen and Derrington, 2000). Briefly, luminance-defined patches were built using a luminance grating, described by the following equation:

$$L = C\sin(2\pi fx + 2\pi gt + \theta)$$

with $C = 0.1$ contrast factor, $f = 1$ c/deg spatial frequency, $g = 0.5$ Hz temporal frequency, and θ is a randomised phase. The contrast-defined patches were made by adding two luminance gratings moving in opposite directions, described by the following equation,

$$L = 2C\cos(2\pi f_e x + 2\pi g_e t + \theta_e)\sin(2\pi f_c x + \theta_c)$$

with $C = 0.1$ contrast factor, $f_1 = 1.7$ c/deg, $f_2 = 2.7$ c/deg, $f_e = 0.5(f_1 - f_2), f_c = 0.5(f_1 + f_2)$, and spatial frequencies of 1.7 and 2.7 c/deg.

2.4. Single vs dual task

In the single task condition, participants only performed the motion discrimination task. In the dual task condition, participants had to simultaneously perform the motion discrimination task and an additional task designed to increase the cognitive workload, namely a variant of the N-Back task (Kono et al., 2019; Obinata et al., 2008). For the N-Back task, integers ranging from 1 to 5 were called randomly every 2 s using two loudspeakers. Each call lasted 500 ms. Observers were asked to listen to these numbers, and to press a dedicated button if they thought that the latest number was also called among the N prior numbers. In our experiments, we used the 2-Back task, which is slightly harder than the task used in (Kono et al., 2019). Note that even the easier task used in (Kono et al., 2019) has been shown to successfully increase the MWL (Braver et al., 1997). Fig. 2 illustrates the workings of the N-Back task.

2.5. Procedure

The type of local motion stimulus (luminance-based vs contrast-based) and the type of task (single task vs dual task) constituted the two independent variables. Both were repeated measures variables. In total, each participant performed four blocks, corresponding to the four possible combinations between the levels of the two independent variables. The order of presentation of the four blocks was fully randomized between participants. For each condition, the dependent variable was the minimum exposition time needed for the participant to identify the odd patch, i.e., the patch moving outwards. This time was estimated

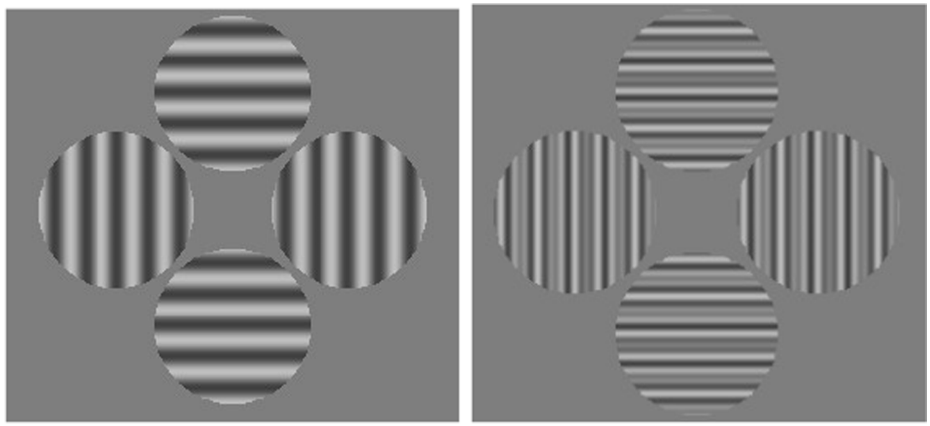

Fig. 1. Static, high contrast representations of local motion stimuli. (Left) luminance-defined stimulus, (Right) contrast-defined stimulus. Each patch was 5° in diameter, and was positioned symmetrically around the center.

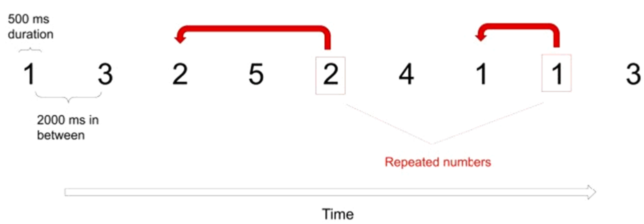

Fig. 2. Illustration of the *N*-back task. The black arrow indicates the progression of time. All numbers were presented sequentially to the observer as audio stimuli, with a 2 s pause between numbers. The red arrows indicate to the presence of repeated numbers, at which point the observer should press the dedicated button. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

using an adaptive procedure for psychometric function estimation, based on ideal temporal windows. Specifically, visual stimuli were presented to the participant sequentially, and after each presentation (i.e., after each trial), the participant was asked to indicate which patch was moving outwards using the arrow keys. For each trial, and based on the previous responses of the participant, the duration of presentation of the stimulus was chosen using the Questplus algorithm (Watson, 2017), which is a generalization of the Quest method (Watson and Pelli, 1983).

For the two blocks performed in dual task conditions, the participants were simultaneously asked to press a dedicated button if they thought that the latest number was also called among the N prior numbers. For all trials, the participants were asked to keep one hand above the response button for the N-Back task, and one hand above the arrow keys, which were used as response buttons for the visual stimuli. One experimental block consisted of 70 successive stimuli presentation. Participants rested for at least two minutes between each experimental block. Before each test block, the participants performed 30 training trials to familiarize themselves with the forthcoming task.

2.6. Statistical Analysis

As mentioned above, we used the QuestPlus procedure to choose the sequence of visual stimuli duration. QuestPlus is a Bayesian algorithm that uses the minimum entropy principle to choose the adaptive sequence of stimuli to present to the observer. We set the psychometric function to be a Weibull cumulative distribution function (1),

$$W(t|\gamma,\lambda,\alpha,\beta) = \gamma + (1-\lambda-\gamma)(1-\exp(-(t/\alpha)^{\beta})) \tag{1}$$

where γ denotes the guess rate, λ the lapse rate (see e.g. Wichmann and Hill (2001)), and α, β are the free parameters of the distribution. We used minimum and maximum possible durations of respectively 0.05 and 5 s, a $[0.05, 4.95]$ range of α values, and a $[0.05, 5]$ range of β values. Note that in this experiment, the guess rate γ was 25% (as there were four patches). For the other parameters, we used the values recommended by Watson (2017), and a modified version of the implementation provided by Peirce et al. (2019).

At the end of the experiment, we used the α and β produced by the Bayesian model to estimate the 50% accuracy threshold. When relevant, we also computed the success rate of the dual task in order to control the performance of the participants with respect to their compliance to the *N*-back task. All participants had a success rate above 90%, which was significantly higher than the 64% maximum performance of an oblivious observer.

Bayesian Models. To analyse our results, we used the following Bayesian model. We assumed that the thresholds τ measured during the psychometric experiments with stimulus s (luminance- or contrast-based) and task t (single task or dual task) were following the Normal distribution $\mathcal{N}$:

$$\tau(s,t) \sim \mathcal{N}\left(\mu_{s,t}, \sigma^2_{s,t}\right)$$

where $\mu_{s,t}$ (resp. $\sigma_{s,t}$) denotes the mean value (resp. the standard deviation) of the normal distribution. Both $\mu_{s,t}$ and $\sigma_{s,t}$ can take four different values, as summarized in Table 1. We used non informative priors (uniform distribution over the interval $[0, 5]$ for each $\mu_{s,t}$ and uniform distribution over the interval $[0, 10]$ for $\sigma_{s,t}$).

The objective was twofold. First, for each combination of stimulus type and task, we aimed at estimating the posterior distributions $\mu_{s,t}$ and $\sigma_{s,t}$. To this end, we reported the mean value of the posterior distribution of each parameter, together with the 95 % High Density Interval (HDI). Second, to compare the posterior distribution of any two settings, we considered the distribution of threshold differences. For instance, to

Table 1
Parameters of the different distributions, as a function of both the nature of the visual stimuli and of the task. All likelihoods are assumed to be Gaussian distributions $\mathcal{N}(\mu, \sigma^2)$, with varying means and standard deviations. Similarly, all priors are uninformative uniform distributions $\mathcal{U}$ on the interval $[0, 5]$ for means and $[0, 10]$ for standard deviations.

Stimulus	Task	Likelihood	Prior Mean	Prior Std
Luminance	Single	$\mathcal{N}(\mu_{LS}, \sigma^2_{LS})$	$\mathcal{U}([0,5])$	$\mathcal{U}([0,10])$
	Dual	$\mathcal{N}(\mu_{LD}, \sigma^2_{LD})$	$\mathcal{U}([0,5])$	$\mathcal{U}([0,10])$
Contrast	Single	$\mathcal{N}(\mu_{CS}, \sigma^2_{CS})$	$\mathcal{U}([0,5])$	$\mathcal{U}([0,10])$
	Dual	$\mathcal{N}(\mu_{CD}, \sigma^2_{CD})$	$\mathcal{U}([0,5])$	$\mathcal{U}([0,10])$

study the impact of the dual task on the detection of the luminance-based stimuli, we studied the distribution of

$$\tau(s = L, t = D) - \tau(s = L, t = S) \sim \mathcal{N}\left(\mu_{LD} - \mu_{LS}, \sigma_{LS}^2 + \sigma_{LD}^2\right).$$

To study the hypothesis that $\mu_{LD} - \mu_{LS} > 0$ (i.e., the assumption that the dual task negatively impacted the detection of luminance-based stimuli), we reported the Bayes factor of the two hypotheses, which was estimated using the annealing sequential Monte Carlo sampling approach. Similarly, we assessed $\mu_{CD} - \mu_{CS} > 0$ (i.e., impact of the dual task on the detection of contrast-based stimuli), $\mu_{CS} - \mu_{LS} > 0$ (i.e., impact of the nature of the stimuli on the discrimination performance in the single task setting), $\mu_{CD} - \mu_{LD} > 0$ (i.e., impact of the nature of the stimuli on the discrimination performance in the dual task setting), and $(\mu_{CD} - \mu_{CS}) - (\mu_{LD} - \mu_{LS}) > 0$ (i.e., impact of the dual task more pronounced on contrast-based than on luminance-based stimuli). In addition to the absolute difference, we studied the relative increment of the threshold by performing a Bayesian analysis of the distributions of $\mu_{LD}/\mu_{LS}, \mu_{CD}/\mu_{CS}$ and their difference $\mu_{LD}/\mu_{LS} > 0$, $\mu_{CD}/\mu_{CS} > 0$. Note that contrarily to the absolute differences, the distributions related to relative increments are no longer Gaussian.

For the sake of convenience, when comparing populations, we also reported *p*-values that were obtained using the non parametric Mann–Whitney U-test with Bonferroni correction for multiple comparisons. We also studied the Pearson correlation coefficient of the different thresholds $\mu_{LD}, \mu_{LS}, \mu_{CD}$, and μ_{CS}. Note that the choice of the Pearson coefficient is motivated by the distribution of the threshold, which are assumed to be Gaussian. All statistical analyses were performed using python 3.8, and the scipy and pymc3 libraries (Salvatier et al., 2016).

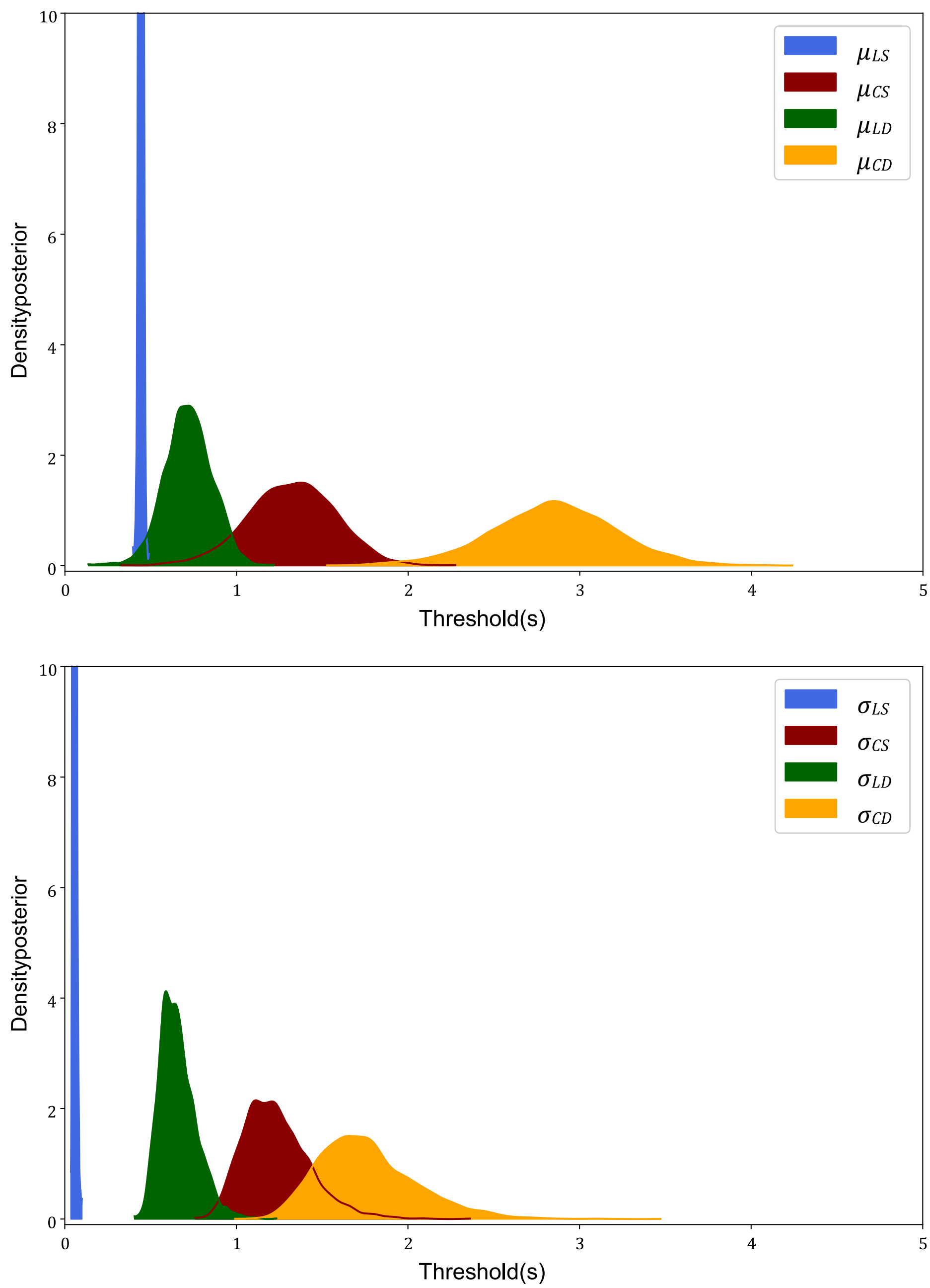

Fig. 3. Posterior distribution of the mean parameters (top) and the standard deviation parameters (bottom) of the Bayesian model, for each possible combination of experimental conditions.

3. Result

As aforementioned, and in line with previous studies on the mechanisms underlying motion perception (Allen and Derrington, 2000; Derrington et al., 2004), we expected the detection threshold to be significantly worse for contrast-based (CS condition) than for luminance-based stimuli (LS condition), and this both in single task and dual task settings. More importantly, we expected the dual task setting to have a much stronger impact on discrimination performance for contrast-based stimuli (CD condition) than for luminance-based stimuli (LD). In the following, we first describe the posterior distributions of our model for each combination of conditions, before testing our different hypotheses by comparing them using both Bayes Factors and Mann–Whitney U-tests.

3.1. Descriptive Analysis

Fig. 3 shows the posterior distribution of the mean value $\mu_{s,t}$ and standard deviation $\sigma_{s,t}$ of the detection threshold $\tau(s,t)$ for each experimental condition (LS, CS, LD and CD). The threshold $\tau(s,t)$ represents the smallest duration of exposition corresponding to a success rate of at least 50%. The mean value as well as the 95% high density interval (HDI) of these posterior distributions are summarized in Table 2. Note that while the Bayesian model was slightly worse at modeling performance for luminance-based than for contrast-based stimuli, the goodness of fit of both models was sufficient for the analysis presented in this section.

The same pattern of results was observed with all participants: a longer stimulus duration was necessary to discriminate contrast-defined patterns than to discriminate luminance-defined stimuli. The addition of a dual task further increased this duration, but much more so for contrast-based than for luminance-based patches. As reported in Table 2, participants required around 443 ms of exposition to identify the outward moving pattern for luminance-defined stimuli, while the required duration was around 1300 ms for contrast-defined patterns. Interestingly, inter-individual variability was low in the LS condition ($\sigma_{LS} = 57$ ms), and much higher in the CS condition ($\sigma_{LS} = 1237$ ms). This might indicate that while all participants were able to easily detect the odd patch in the LS condition, performance in the CS condition was more dependent on the use of a 'good' strategy by the user, thereby increasing the inter-individual variability. This latter point is in line with the observations of Ashida and colleagues (Ashida et al., 2001). Table 2 shows that participants required on average 724 ms of exposition in the LD condition, which is slightly higher than in the LS experiment (see analysis below). Meanwhile, the standard deviation increased to 654 ms, which might also indicate a plurality of strategies used by the participants to tackle the challenging N-Back task, as well as different subjective experiences (see discussion). Finally, note that discrimination performance drastically decreased in the CD experiment, with a mean exposition duration of 2856 ms.

Table 2
Characteristics of the Posterior distributions of the different parameters of the Bayesian model. Mean value, and the 95 % High Density Interval (HDI) are reported for each parameter.

Stimulus	Task	Parameter	Mean	95 % HDI
Luminance	Single	μ_{LS}	0.443	[0.420,0.465]
		σ_{LS}	0.057	[0.042,0.074]
	Dual	μ_{LD}	0.724	[0.477,0.986]
		σ_{LD}	0.654	[0.482,0.852]
Contrast	Single	μ_{CS}	1.326	[0.832,1.800]
		σ_{CS}	1.237	[0.896,1.602]
	Dual	μ_{CD}	2.856	[2.177,3.522]
		σ_{CD}	1.761	[1.284,2.298]

3.2. Comparative Analysis

Fig. 4 shows the participants' performance (50% detection threshold) for each following pair of experimental conditions: LS and LD, CS and CD, LS and CS, and LD and CD. Table 3 presents the Bayes Factor values and p-values resulting of the Bayesian analysis of the difference between each pair of conditions.

In the first comparison, LS vs LD (i.e., impact of the dual task on the detection of first-order motion), it can be noted that most participants achieved similar results in both conditions (almost all dots are in the bottom left quadrant of the chart, i.e., $\tau_{LD} \approx \tau_{LS}$) – with the exception of three participants whose performance decreased in LD. This result can also be seen in the results of the analysis of $\mu_{LD} - \mu_{LS}$, whose expected value was found to be slightly positive (0.283). However, it should be noted that while there is evidence for H_1 (Bayes Factor 44.6 and p-value 0.013), these values are significantly smaller than for the other comparisons (see below). This was also found in the analysis of the relative increase of the threshold μ_{LD}/μ_{LS}: the expected value was slightly bigger than one (≈ 1.643) with some evidence for H_1 (Bayes Factor 26.0 and p-value 0.012).

Meanwhile, in the second comparison between CS and CD (i.e., impact of the dual task on the detection of second-order motion), the performance of most participants decreased significantly (most points are in the top left corner, i.e., $\tau_{CD} > \tau_{CS}$). This observation is confirmed by the statistical analysis, as the expected value of $\mu_{CD} - \mu_{CS}$ was 1.530 – i.e., more than 1500 ms, with extremely strong evidence for H_1 (Bayes Factor > 10000 and p-value < 0.001). Note that a Bayes Factor of 10000 indicates that the likelihood of H_1 is 10000 times higher than the likelihood of H_0. Therefore, these results hint at a much stronger impact of the N-Back task on the perception of second-order motion. This was also reflected in the analysis of the relative increase μ_{CD}/μ_{CS}, as the expected value was significantly larger than one (≈ 2.794) with very strong evidence (Bayes Factor > 1000 and p-value < 0.001).

Similarly, it can be seen that the difference between the luminance-based and contrast-based stimuli were significant in the single task setting (LS and CS), in which all participants achieved lower performance with contrast-based than with luminance-based stimuli (all points are above the diagonal line, i.e., $\tau_{CS} > \tau_{LS}$). This is reflected in the analysis, were the expected value of $\mu_{CS} - \mu_{LS}$ was 0.880, with very strong evidence in favor of H_1 (Bayes Factor ≈ 1400 and p-value < 0.001). This difference between the two conditions was even larger in the dual task setting (LD and CD), (data points are concentrated in the top left corner of the chart, i.e., $\tau_{CD} > \tau_{LD}$). Here the expected value of $\mu_{CD} - \mu_{LD}$ was 2.132 – more than 2 s – with extremely strong evidence for H_1 (Bayes Factor > 100000 and p-value < 0.001).

The last result hints at the fact that motion discrimination was significantly more impacted by the dual task for contrast-based than for luminance-based stimuli. This was confirmed by the last analyses,

$$(\mu_{CD} - \mu_{CS}) - (\mu_{LD} - \mu_{LS}),$$

$$(\mu_{CD}/\mu_{CS}) - (\mu_{LD}/\mu_{LS}),$$

which quantifies the aforementioned difference of impact. Indeed, their expected values were respectively 1.254 – i.e., the impact on the contrast-based stimuli was on average more than 1 s larger than the impact on the luminance-based stimulus– and 1.212, with very strong evidence in favor of H_1 in both cases (Bayes Factor resp. ≈ 3900 and 400 and p-value resp. ≈ 0.001 and ≈ 0.007).

Finally, note that none of the correlation between respectively μ_{CD} and μ_{CS}, μ_{LD} and μ_{LS}, μ_{CD} and μ_{CS}, μ_{LD} and μ_{LS} were found to be statistically significant (all p-values > 0.1). This is also shown by Fig. 4, where no correlation can be observed.

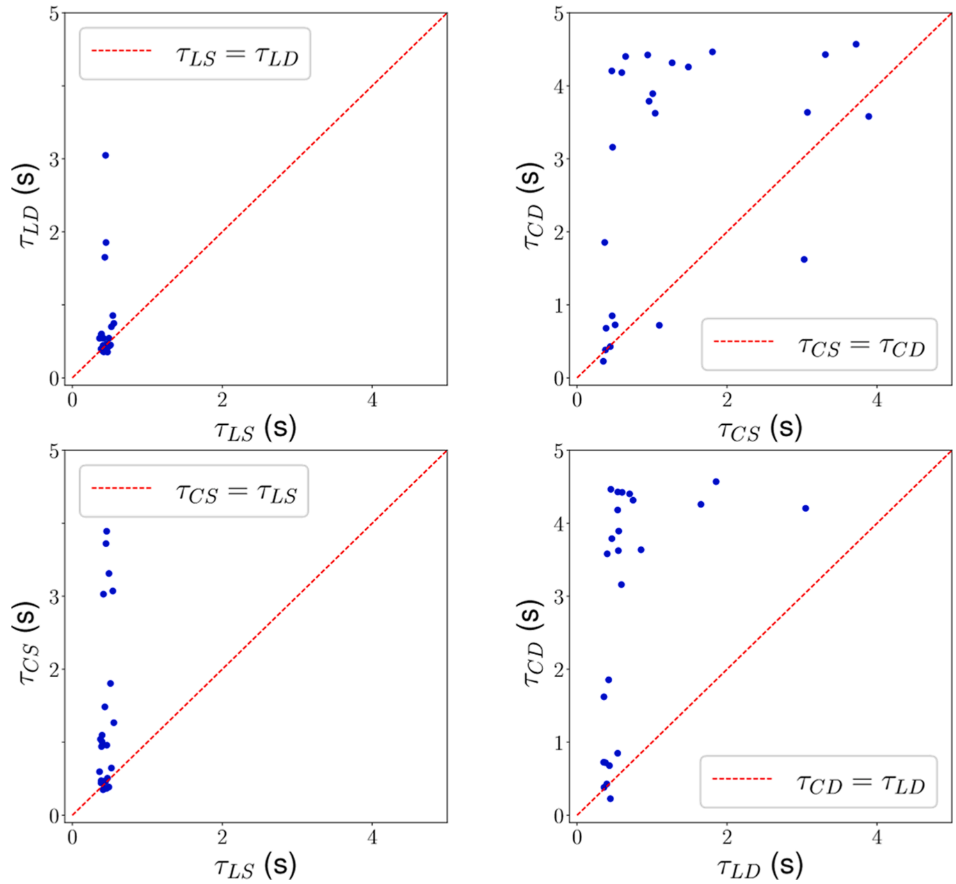

Fig. 4. Pairwise performance (detection threshold) of the 24 observers, for any two combination of conditions: (Top Left) Luminance Single Task and Luminance Dual Task, (Top Right) Contrast Single Task and Contrast Dual Task, (Bottom Left) Luminance Single Task and Contrast Single Task, (Bottom Right) Luminance Dual Task and Contrast Dual Task. The red line denotes the equality line for easier reading.

Table 3
Characteristics of the Posterior distributions of the difference $\mathscr{D}$ between average threshold values for different conditions. $\mathbb{E}(\mathscr{D})$ reports the mean of the difference, the Bayes Factor is the quotient of likelihood between the assumptions, and the p-value are obtained using anon parametric Mann–Whitney U-test.

Variable of interest $\mathscr{D}$	H_0	H_1	Mean Value $\mathbb{E}(\mathscr{D})$	Bayes Factor H_1/H_0	p–value
$\mu_{LD} - \mu_{LS}$	$\mathscr{D}\leqslant 0$	$\mathscr{D} > 0$	0.283 s	44.6	0.013
$\mu_{CD} - \mu_{CS}$	$\mathscr{D}\leqslant 0$	$\mathscr{D} > 0$	1.530 s	1.5×10^4	<0.001
$\mu_{CS} - \mu_{LS}$	$\mathscr{D}\leqslant 0$	$\mathscr{D} > 0$	0.880 s	1.4×10^3	<0.001
$\mu_{CD} - \mu_{LD}$	$\mathscr{D}\leqslant 0$	$\mathscr{D} > 0$	2.132 s	2.4×10^6	<0.001
μ_{CD}/μ_{CS}	$\mathscr{D}\leqslant 1$	$\mathscr{D} > 1$	2.794	1.1×10^3	< 0.001
μ_{LD}/μ_{LS}	$\mathscr{D}\leqslant 1$	$\mathscr{D} > 1$	1.643	26.0	0.012
$(\mu_{CD} - \mu_{CS}) - (\mu_{CD}/\mu_{CS})$	$\mathscr{D}\leqslant 0$	$\mathscr{D} > 0$	1.254 s	3.9×10^3	0.001
$(\mu_{CD}/\mu_{CS}) - (\mu_{LD}/\mu_{LS})$	$\mathscr{D}\leqslant 0$	$\mathscr{D} > 0$	1.212	400.75	0.007

4. Discussion

Motion sensors and Feature tracking. In our study, the exposition time required to identify the outward moving patch was consistently and significantly larger for contrast-based than for luminance-based stimuli. This was true both in single task and dual task settings. This result is in line with previous works (Allen and Derrington, 2000; Ashida et al., 2001) which observed that motion discrimination based on feature tracking is significantly slower than motion discrimination based on motion sensors. Importantly, we show here for the first time that adding a dual task designed to increase mental workload affects significantly more motion discrimination for contrast-based stimuli than for luminance-based stimuli. This suggests that mechanisms underlying second-order motion detection rely more heavily on cognitive resources that are not necessarily 'perception-related', supporting the idea of the existence of the feature tracking mechanism, as well as its reliance of cognitive resources.

Impact of MWL during dual tasks. Our results show that adding a dual task resulted in a poorer discrimination performance and an increased variance for both types of stimuli (LD and CD). While this effect was observed with both luminance-based and contrast-based stimuli, it was significantly larger for contrast-based stimuli, both in term of absolute increase and of relative increase. On the one hand, the increase in variance can be explained by the subjective nature of MWL. In particular, previous research suggests that the MWL associated to a task can be decomposed in a task load, which is inherent to the task (here N-Back), and a subjective experience, which depends on many factors, such as past experience, availability of cognitive resources, or emotional load (Pickup et al., 2005; Oviatt et al., 2018). Therefore, it is likely that each participant experienced the N-Back task differently, thus increasing the spread of the results. This can also be observed in the absence of statistically significant correlations between the different thresholds. On the other hand, the decrease in performance (i.e., higher time of exposition required for stimulus detection) was consistent across participants and significantly higher for the contrast-defined stimuli. This suggests that in the CD condition, both tasks were competing for the same cognitive resources, inducing poorer discrimination performance. As opposed to that, the visual task in the LD condition was likely less demanding in terms of mental resources. This would be in line with our initial assumption. Interestingly, the observed effects were large, despite the fact that the N-Back task relied on auditory perception, which has been shown to mobilize different resources than visual perception, thereby limiting the competition for perceptive resources (Wickens, 2008) (see discussion below).

MWL and increased reaction time. It has long been recognized that the manipulation of MWL affects reaction time. As a matter of fact, reaction time is frequently used as an indirect measurement of MWL (see e.g., (Just and Carpenter, 1992; Braver et al., 1997)). In line with this, one might wonder whether the increased time of exposition required to properly detect odd patches in the dual task setting were not 'merely' a consequence of an increased reaction time. We argue against this interpretation for two important reasons. First, the increase of threshold measured in the CD condition significantly larger than that observed in the LD condition, and this even though the secondary task was strictly identical in both cases (the N-Back task). Second, while the exposition time was limited for each stimulus in order to assess the performance of the participants, the response time was completely free of any constraint. As a consequence, an increased reaction time should not have directly impacted motion discrimination performance. For these reasons, we believe that the effect of an increased MWL on motion perception cannot be explained solely by an increased reaction time.

Choice of the N-Back task. It could be argued that the choice of the *N*-back task with N = 2 as a secondary task forfeit the study of the dose–effect of MWL on motion perception. The rationale behind the choice of this task was threefold. First, as discussed before, the N-Back task is a commonly used dual task to study the effect of MWL, both in neuro-imaging and in applications such as driving. Second, the N-Back task only involves auditory stimuli, which likely mobilize different resources than those mobilized by visual perception (i.e., the modality of the main task). This in turns results in lower competition for cognitive resources, therefore enticing lower MWL and better performance (Wickens, 2002; Wickens, 2008). Therefore, if a significant effect is detected with the N-Back task, as it was the case in our study, one can reasonably expect an even stronger effect with secondary tasks including a visual component (such as looking at a phone while driving). Third, due to the subjective nature of the MWL task, studying the dose–effect of MWL on motion perception would require a proper quantification of the effect evoked by the MWL on each participant (Just and Carpenter, 1992; Pickup et al., 2005). This is a complicated task for which there is no consensus; indeed even tests such as the NASA-TLX (Hart and Staveland, 1988), which are considered gold standard tools for the measurement of MWL, have important limitations (Hart, 2006). Therefore, we chose to focus on the effect of the MWL on motion perception, using the N-Back with N = 2, which has been shown to successfully increase the MWL (see e.g. (Wickens, 2008) and references therein). Future works might try to study the dose–effect phenomenon by first evaluating the MWL response of each participant in a large variety of tasks, before assessing their effect on motion perception. However, this might prove to be an arduous and non-trivial challenge.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Josa a, 2*(2), 284–299.
Allen, H. A., & Derrington, A. M. (2000). Slow discrimination of contrast-defined expansion patterns. *Vision Research, 40*(7), 735–744.
Anstis, S., Verstraten, F. A., & Mather, G. (1998). The motion aftereffect. *Trends in Cognitive Sciences, 2*(3), 111–117.
Ashida, H., Seiffert, A. E., & Osaka, N. (2001). Inefficient visual search for second-order motion. *JOSA A, 18*(9), 2255–2266.
Braddick, O. (1974). A short-range process in apparent motion. *Vision Research, 14*(7), 519–527.
Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage, 5*(1), 49–62.
Cavanagh, P. (1992). Attention-based motion perception. *Science, 257*(5076), 1563–1565.
Cavanagh, P., Mather, G. (1989). Motion: The long and short of it. Spatial vision.
Derrington, A. M., Allen, H. A., & Delicato, L. S. (2004). Visual mechanisms of motion analysis and motion perception. *Annual Review of Psychology, 55*, 181–205.
Derrington, A. M., & Badcock, D. R. (1985). Separate detectors for simple and complex grating patterns? *Vision research, 25*(12), 1869–1878.
Derrington, A. M., Badcock, D. R., & Henning, G. B. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision research, 33*(13), 1785–1794.
Derrington, A. M., Badcock, D. R., & Holroyd, S. A. (1992). Analysis of the motion of 2-dimensional patterns: Evidence for a second-order process. *Vision Research, 32*(4), 699–707.
Georgeson, M. A., & Shackleton, T. M. (1989). Monocular motion sensing, binocular motion perception. *Vision Research, 29*(11), 1511–1523.
Hart, S. G. (2006). Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 50, pp. 904–908). Los Angeles, CA: Sage publications Sage CA.
Hart, S. G., & Staveland, L. E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in Psychology, 52* pp. 139–183). Elsevier.
Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122.
Kantowitz, B.H., 1987. 3. mental workload. In: Advances in psychology. Vol. 47. Elsevier, pp. 81–121.
Kono, T., Sato, Y., & Wada, T. (2019). Model analysis of influence of mental workload on vestibulo-ocular reflex. *IFAC-PapersOnLine, 52*(19), 329–334.
Morrone, M. C., Burr, D. C., & Vaina, L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature, 376*(6540), 507–509.
Obinata, G., Tokuda, S., & Shibata, N. (2008). Mental workloads can be objectively quantified in real-time using vor (vestibulo-ocular reflex). *IFAC Proceedings Volumes, 41*(2), 15094–15099.
Oviatt, S., Schuller, B., Cohen, P., Sonntag, D., Potamianos, G., & Krüger, A. (2018). The handbook of multimodal-multisensor interfaces, Volume 2: Signal processing, architectures, and detection of emotion and cognition. *Morgan & Claypool.*
Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203.
Pickup, L., Wilson, J. R., Sharpies, S., Norris, B., Clarke, T., & Young, M. S. (2005). Fundamental examination of mental workload in the rail industry. *Theoretical Issues in Ergonomics Science, 6*(6), 463–482.
Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in python using pymc3. *PeerJ Computer Science, 2*, e55.
Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin & Review, 24*(4), 1077–1096.
Ullman, S. (1979). The interpretation of visual motion. Massachusetts Inst of Technology Pr.
Watson, A. B. (2017). Quest+: A general multidimensional bayesian adaptive psychometric method. *Journal of Vision, 17*(3), 10.
Watson, A. B., & Pelli, D. G. (1983). Quest: A bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113–120.
Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. fitting, sampling, and goodness of fit. *Perception & Psychophysics, 63*(8), 1293–1313.
Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159–177.
Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors, 50*(3), 449–455.
Wohlgemuth, A. (1911). *On the after-effect of seen movement, No. 1–2*. University Press.
Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: Mental workload in ergonomics. *Ergonomics, 58*(1), 1–17.