HAL open science

Unifying Speed-Accuracy Trade-Off and Cost-Benefit Trade-Off in Human Reaching Movements

Luka Peternel, Olivier Sigaud, Jan Babič

To cite this version:

Luka Peternel, Olivier Sigaud, Jan Babič. Unifying Speed-Accuracy Trade-Off and Cost-Benefit Trade-Off in Human Reaching Movements. Frontiers in Human Neuroscience, 2017, 11, pp.615. 10.3389/fnhum.2017.00615 . hal-01679624

HAL Id: hal-01679624

https://hal.sorbonne-universite.fr/hal-01679624

Submitted on 10 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

frontiers in Human Neuroscience

ORIGINAL RESEARCH
published: 19 December 2017
doi: 10.3389/fnhum.2017.00615

Unifying Speed-Accuracy Trade-Off and Cost-Benefit Trade-Off in Human Reaching Movements

Luka Peternel[1,2], *Olivier Sigaud*[3*] *and Jan Babič*[2*]

[1] HRII Lab, Advanced Robotics, Istituto Italiano di Technologia, Genoa, Italy, [2] Department for Automation, Biocybernetics and Robotics, Jožef Stefan Institute, Ljubljana, Slovenia, [3] Sorbonne Universités, UPMC Univ Paris 06, CNRS UMR 7222, Institut des Systèmes Intelligents et de Robotique, Paris, France

Two basic trade-offs interact while our brain decides how to move our body. First, with the cost-benefit trade-off, the brain trades between the importance of moving faster toward a target that is more rewarding and the increased muscular cost resulting from a faster movement. Second, with the speed-accuracy trade-off, the brain trades between how accurate the movement needs to be and the time it takes to achieve such accuracy. So far, these two trade-offs have been well studied in isolation, despite their obvious interdependence. To overcome this limitation, we propose a new model that is able to simultaneously account for both trade-offs. The model assumes that the central nervous system maximizes the expected utility resulting from the potential reward and the cost over the repetition of many movements, taking into account the probability to miss the target. The resulting model is able to account for both the speed-accuracy and the cost-benefit trade-offs. To validate the proposed hypothesis, we confront the properties of the computational model to data from an experimental study where subjects have to reach for targets by performing arm movements in a horizontal plane. The results qualitatively show that the proposed model successfully accounts for both cost-benefit and speed-accuracy trade-offs.

Keywords: expected utility, hit dispersion, cost-benefit, speed-accuracy, arm reaching

OPEN ACCESS

Edited by:
Mikhail Lebedev,
Duke University, United States

Reviewed by:
Digby Elliott,
Liverpool John Moores University, United Kingdom
Bastien Berret,
Université Paris-Sud, Université Paris-Saclay, France

***Correspondence:**
Olivier Sigaud
olivier.sigaud@upmc.fr
Jan Babič
jan.babic@ijs.si

Received: *29 September 2017*
Accepted: *05 December 2017*
Published: *19 December 2017*

Citation:
Peternel L, Sigaud O and Babič J (2017) Unifying Speed-Accuracy Trade-Off and Cost-Benefit Trade-Off in Human Reaching Movements. Front. Hum. Neurosci. 11:615. doi: 10.3389/fnhum.2017.00615

1. INTRODUCTION

There has been a recent progress in motor control research on understanding how the time of a movement is chosen. In particular, two models proposed an optimization criterion that involves a trade-off between the muscular effort and the subjective value of getting the reward, hence a *cost-benefit trade-off* (CBT) (Shadmehr et al., 2010; Rigoux and Guigon, 2012). On one hand, getting a reward faster requires a larger muscular effort (Young and Bilby, 1993). On the other hand, the subjective value of getting a reward decreases as the time needed to do so is increased (Green and Myerson, 2004). As a result, the net utility consisting of the subjective value minus the muscular effort is optimal for a certain time, as illustrated in **Figure 1A**. However, these models do not account directly for basic facts about the relation between movement difficulty and movement duration as captured more than 50 years ago by Fitts' law (Fitts, 1954). According to this law, the smaller a target, the slower the reaching movement. This is well explained by the so-called *speed-accuracy trade-off* (SAT) stating that, the faster a movement, the less accurate it is, hence the higher the probability to miss the target. So a subject reaching too

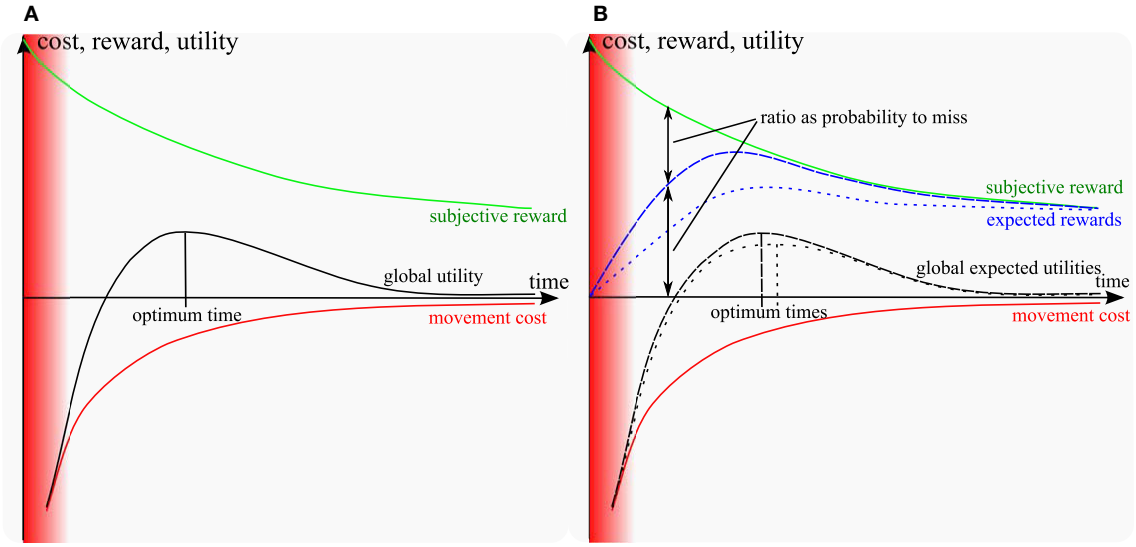

FIGURE 1 | Sketch of models: **(A)** cost-benefit trade-off models (Shadmehr et al., 2010; Rigoux and Guigon, 2012); **(B)** including the probability to miss (proposed model). Red area: unfeasible short times. Green: subjective reward of movement; red: muscular energy cost. Blue and black lines: expected rewards and utilities respectively (dashed for a large target, dotted for a small target). In the standard and proposed models, the subjective utility of getting the reward decreases over time, while slow hitting movements require less muscular effort. As a result, the sum of the reward and the (negative) effort cost reaches a maximum for a certain time (optimum time). The expected reward is the probability-weighted average of the subjective reward over infinitely many repetitions. It converges to the blue line as the number of repetitions increase. The same for the global expected utility. In the proposed model, the probability to hit is null for fast movements and converges to 1 as movements slow down. Furthermore, as movement time increases, the probability to hit increases faster for a large target than for a small one. Therefore, the expected reward matches the subjective reward faster for large targets than for small targets and the optimum time (time at the maximum of global expected utility) is shifted toward longer time for smaller targets, as a result of the speed-accuracy trade-off. Besides, at the optimum time, the probability to miss the target is larger for small targets than for large targets. Note that if the target becomes considerably large compared to the length of the movement, it could affect the cost characteristics. Having a larger target width might decrease the expected effort, because of the biomechanics and anisotropy of inertia of the arm. However, in this study we considered relatively small targets compared to the movement length.

fast may not get the subjective value associated to reaching and should slow down. Several studies in the past developed various theories on SAT (Keele, 1968; Schmidt et al., 1979; Crossman and Goodeve, 1983; Meyer et al., 1988; Elliott et al., 2001). Later on, some of the missing aspects were covered by the model of Dean et al. (2007). The key difference to CBT models (Shadmehr et al., 2010; Rigoux and Guigon, 2012) is that, instead of maximizing a reward, this model maximizes a *reward expectation*, i.e., the reward times the probability to get it. However, this model is abstract and it looks for an optimal trade-off between an externally decayed reward and a parametric SAT diagram that relates the probability of missing to movement time. As such, it accounts neither for movement execution, nor for the choice of a motor trajectory and its impact on the cost of movement.

To overcome the limitations of the previously proposed models, we present a new motor control model that is able to account for both CBT and SAT by maximizing the *expected utility* of reaching movements. Unifying both trade-offs allows our model to be the first to account simultaneously for several motor control phenomena related to movement trajectory, velocity profiles and hit dispersion, highlighting the crucial importance of stochastic optimization in the presence of motor noise. Though the model is general, we illustrate its properties in the context of reaching with the arm.

The faster movements tend to be less accurate due to the presence of single-dependant noise in the human motor control system (Schmidt et al., 1979; Meyer et al., 1988; Harris and Wolpert, 1998; Todorov and Jordan, 2002). Our model builds on the stochastic optimal control view of motor control (Meyer et al., 1988; Harris and Wolpert, 1998; Todorov and Jordan, 2002; Todorov, 2004, 2005; Li, 2006), which considers variability as a key ingredient of human movements. In addition, the model adopts an infinite-horizon formulation, like several other models from the literature (Rigoux and Guigon, 2012; Qian et al., 2013).

The general intuition is illustrated in **Figure 1B**. In the previous CBT models, the target was given as a single point and the movement was considered as always reaching it, irrespective of any target size constraint. In order to fully account for Fitts' law, it is necessary to incorporate the intrinsic dispersion of reaching movements toward a target and the effect of sensory and muscular noise on this dispersion (e.g., Harris and Wolpert, 1998, see Faisal et al., 2008 for a review). As a consequence of noise, movements reaching faster should suffer from a higher dispersion and thus get a lower *probability* of reaching a small target. Thus, instead of optimizing a utility as the sum of a reward and a cost terms as in (1) from Rigoux and Guigon (2012):

$$J(\boldsymbol{u}) = \int_0^\infty e^{-t/\gamma}\left[\rho R(\boldsymbol{s}_t) - \upsilon L(\boldsymbol{u}_t)\right]dt, \quad (1)$$

we hypothesize that human motor control optimizes an *expected utility*, that is the above utility times the probability to get it, as captured in (2):

$$J(\boldsymbol{u}) = \mathbb{E}(\boldsymbol{s}, \boldsymbol{u})\left(\int_0^{\infty} \left[e^{-t/\gamma}\rho R(\boldsymbol{s}_t) - \upsilon L(\boldsymbol{u}_t)\right]dt\right), \qquad (2)$$

where $\mathbb{E}(\boldsymbol{s}, \boldsymbol{u})[...]$ denotes the expectation over the random variables $\boldsymbol{s}$ and $\boldsymbol{u}$, which is the probability-weighted average of its argument over infinitely many repetitions. In (1) and (2), t is time, $\boldsymbol{s}_t$ is the current state, $\boldsymbol{u}_t$ is the current muscular activation vector, $R(\boldsymbol{s}_t)$ is the immediate reward function that equals 1 when reaching the target and is null everywhere else, and $L(\boldsymbol{u}_t)$ is the movement cost. As in Rigoux and Guigon (2012) and many other motor control models (e.g., Guigon et al., 2008b), we take $L(\boldsymbol{u}_t) = \|\boldsymbol{u}_t\|^2$. The meta-parameters of the model, γ, ρ and υ, are explained in section 4.2.1 and given in **Table 3**.

The probability to reach the target can be assumed null for infinitely fast movements with a null duration, and goes to 1 (100%) as movements gets slower. Furthermore, it increases faster for large targets than for small ones. A sketch of the resulting expected reward as a function of movement time is depicted in **Figure 1B** for a small and a large target. The probability to miss can be inferred in **Figure 1B** as the ratio between the subjective and the expected rewards, since the expected reward would be equal to the subjective reward if this probability was null.

As can be seen in **Figure 1B**, if the target is smaller, then the probability to reach it is smaller for a given time, thus the expected reward should itself be smaller. Therefore, the optimum time resulting from the optimal combination of this expected reward with the cost of movement should shift to longer times, which is consistent with Fitts' law.

Furthermore, at the optimum time, the probability to miss the target is larger for small targets than for large targets. This is why subjects miss significantly more often small targets than large targets (see **Figure 4**). Going slower would decrease their probability to miss, but would incur a lower global utility, due to the loss in subjective utility. Explaining this specific phenomenon is a distinguishing property of our model. Besides, other empirical facts resulting from the unification of both trade-offs are studied below.

The proposed computational model was implemented through the simulation of a two degrees-of-freedom (DoFs) planar arm model controlled by 6 muscles, illustrated in **Figure 7**. Critically, the model incorporated delayed feedback and signal-dependent motor noise, accounting for the fact that the motor activation signal descending from the Central Nervous System (CNS) to motoneurons is corrupted with some noise that is proportional to this signal (Selen et al., 2006). An optimization algorithm was used to obtain a controller providing musculars activations to the model. Given stochasticity of the plant and delayed feedback, a state estimation component was required in the control loop (see e.g., Guigon et al., 2008b). We implemented an *ad hoc* state estimator described in section 4.2.3. The goal of the optimized controller was to maximize the cost function given in (4) (see section 4.2.2 for details).

To validate the proposed hypothesis and motor control model, we designed an experimental study where ten subjects had to reach targets displayed on a screen by performing large horizontal arm movements. In order to study the combined effects of the CBT and the SAT on movement trajectory, velocity and hit dispersion, we rewarded the subjects as a function of a number of targets they reached in a limited duration, and we varied the starting point and target size so as to enforce various accuracy constraints in their movements. More precisely, the setup included 15 starting points along three circles at 15, 37.5, and 60% of arm length (see **Figure 5B**) from the target and we used four different target sizes (5, 10, 20, and 40 mm), consistently with the simulation setup. We recorded the reaching hand trajectories of subjects with a dedicated haptic manipulator and their muscular activations through surface electromyography (EMG). To amplify the effect of the cost of motion, the haptic manipulator emulated a viscous media through which the subject had to move the hand. A monitor displayed the current motion from the starting points and a wall where the target was located. From the recorded movement of subjects, we extracted trajectories, movement time, velocity profiles and dispersion of hits on the target. From measured EMG, we obtained muscle activations and calculated the effort related to the arm movement. More details about the experimental study and the computational model are given in section 4.

2. RESULTS

We analyzed three main aspects of reaching movements that are most relevant to CBT and SAT. First, we show the velocity profiles. Second, we show the movement reaching times with respect to Fitts' law. Third, we show the hit dispersion diagram that indicates distribution of reaching movement final position on the wall, where the target was located. Each of the three aspects is shown for both simulation and experimental data. Additional supplementary results are presented in Supplementary Data, which show some other details of the reaching movement.

The subjects were given a limited amount of time per session (100 s) in order to be forced to follow CBT. In the available time, on average each subject performed 71.5 ± 13.7 trials and accumulated reward 62.9 ± 13.0 for 5 mm target, 82.9 ± 13.4 trials and reward 81.9 ± 13.4 for 10 mm target, 116.2 ± 17.3 trials and reward 115.9 ± 17.0 for 20 mm target, and 142.6 ± 10.3 trials and reward 140.6 ± 10.3 for 40 mm target.

2.1. Velocity Profiles

Velocity profiles are depicted in **Figure 2**. One can see that, for subjects and for the model, movement velocity increases with the distance to the target to compensate for longer distance by permitting higher motor cost in order to reach the target in a reasonable time (i.e., maintain reasonable expected utility according to CBT). This relationship is also evident in **Table 1**, where the mean peak velocity is increasing with the reaching distance (see columns). Note that the values are normalized to the first value to facilitate an easier comparison. Furthermore, movement time increases when the target is smaller, and also

TABLE 1 | Peak velocity relationship to target size and distance (normalized to the first value).

		Target [m] - experimental				Target [m] - model			
		0.005	0.01	0.02	0.04	0.005	0.01	0.02	0.04
Distance	Short	1.00	1.08	2.13	2.98	1.00	1.20	1.56	1.74
	Med.	2.69	2.83	3.20	3.90	2.07	2.18	2.43	2.90
	Long	4.20	4.14	4.58	4.93	2.45	2.68	3.36	3.93

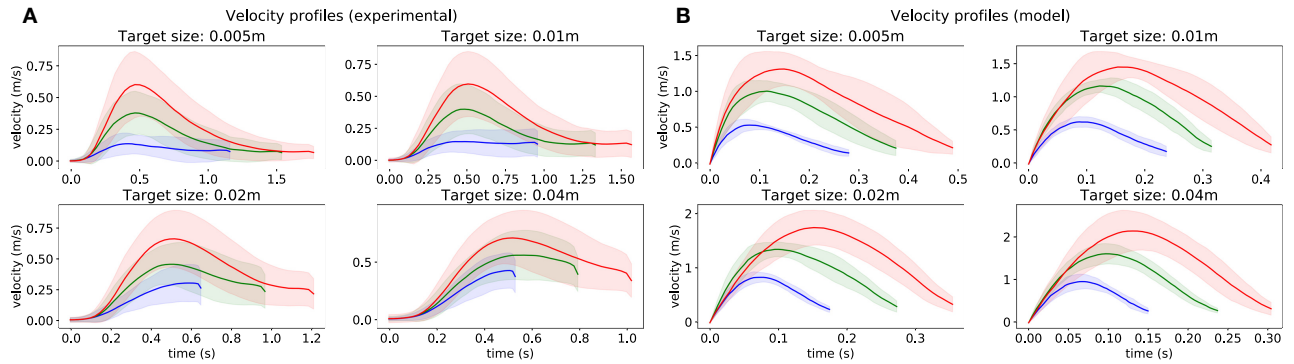

FIGURE 2 | Velocity profiles obtained from 1,500 trajectories starting from 15 points (see **Figure 5B**), for each target size. **(A)** recorded from subjects. **(B)** obtained from the model. The color of lines depends on the distance between the initial point and the target.

increases with the movement distance, consistently with Fitt's law (i.e., following SAT). This relationship is also evident in **Table 2**, where the mean movement time is decreasing with the target size (see rows). One can also observe that these relationships are relatively similar between the experimental data and the simulation data.

In addition, these profiles resemble asymmetric bell-shaped trajectories with the peak velocity lying early in the movement, which corresponds to the established studies from the literature (Plamondon, 1991). The time of this peak occurs earlier when the target is smaller, and later for longer movements. Additionally, the endpoint velocity does not differ much depending on the length of the movement, but is higher for larger targets.

There are also some discrepancies between subjects and the model. The most obvious is that the model is much faster than the subjects who also slow down the motion sooner than the model, resulting in a more pronounced peak. These discrepancies are further discussed in section 3.4.1.

2.2. Fitts' Law

Fitts' law states that the movement time (MT) is linear in its index of difficulty (ID), this index being larger for longer movements and smaller targets. The equation that describes the Fitts' law is

$$\mathrm{MT} = a.\underbrace{\log_2\left(\frac{D}{W}\right)}_{\mathrm{ID}} + b, \tag{3}$$

where D is the length of the movement (denoted with A for amplitude in other papers), W is the width of the target and a and b are the linear coefficients. This law was initially studied for one dimensional movements, and then extended for many other contexts (Soechting, 1984; Bootsma et al., 1994; Laurent, 1994; Plamondon and Alimi, 1997; Smyrnis et al., 2000; Bootsma et al., 2004).

From the trajectory data, we computed ID values for different distances D and target widths W using (3). **Figure 3** shows the resulting movement time MT over ID for the subject and for the model. One can see that the experimental data are strongly consistent with Fitts' law ($r^2 > 0.9$) which also holds for the computational model ($r^2 > 0.7$).

The obtained values of b are quite similar between the model and the experimental data, corresponding to a close-to-null offset. However, the values of a do not match. First, as outlined in section 2.1, the movements from the computational model are faster than those of subjects (see section 3.4.1 for discussion). Besides, according to the literature, the values of a vary widely across subjects (Crossman and Goodeve, 1983; Scott MacKenzie, 1989).

2.3. Hit Dispersion

The hit dispersion resulting from our study is shown in **Figure 4**. One can observe a good match between experimental hit dispersion and the one obtained from the model. The Kullback–Leibler divergence was: 0.032, 0.029, 0.069, and 0.043, respectively for each target. This good match is a distinguishing feature of our model as, to our knowledge, it is the first that can reproduce this property of human reaching movements. Nevertheless, the positive correlation between hit dispersion and target width is coherent with the minimal intervention principle (Todorov and Jordan, 2002).

TABLE 2 | Movement time relationship to target size and distance (normalized to the first value).

		Target [m] - experimental				Target [m] - model			
		0.005	0.01	0.02	0.04	0.005	0.01	0.02	0.04
Distance	Short	1.00	0.83	0.56	0.46	1.00	0.82	0.62	0.54
	Med.	1.32	1.15	0.83	0.69	1.24	1.12	0.97	0.86
	Long	1.53	1.35	1.04	0.88	1.66	1.49	1.22	1.09

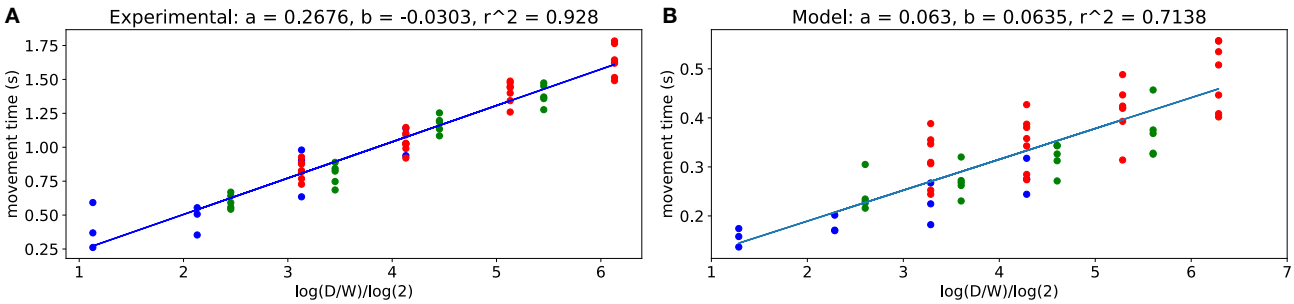

FIGURE 3 | Fitts' law. **(A)** obtained from all subjects trajectories. **(B)** obtained from the model. Each dot corresponds to an average for each target size and target distance pair, obtained either over all subjects or over 1,500 model trajectories.

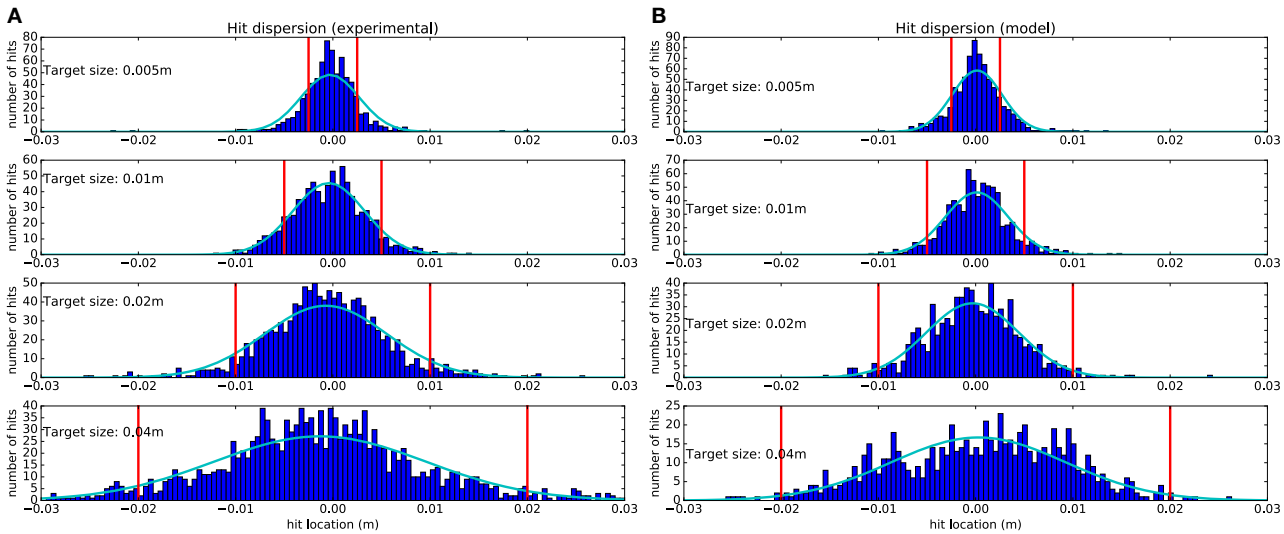

FIGURE 4 | Hit dispersion diagram obtained from 1,500 trajectories starting from 15 points (see **Figure 5B**), for each target size. **(A)** Recorded from subjects. **(B)** Obtained from the model. The vertical red lines denote the lateral target boundaries. Dark blue histograms are obtained by counting all the trajectories that hit the target within a 0.5 millimeter range. Light blue Gaussians are fitted to the histograms.

One can observe that subjects and the model tend to hit more in the center of target, which is in first approximation a good way to maximize the expected utility. The mean location of the fitted Gaussians' peaks with respect to the target center and normalized to the target size for experimental data is at $-5.4 \pm 2.5\%$ and for model is at $0.63 \pm 1.5\%$. Furthermore, when the target is smaller, dispersion is reduced to increase the probability of reaching the target successfully. This reduced dispersion is obtained at the price of a longer movement time, as illustrated in **Figure 3**, as a result of SAT.

However, the probability to miss the target is never null, neither for the subjects nor for the model. Since the experimental task included a limitation on the available time to perform the task, it was more optimal to pay the price of a few failed movements than to move slow enough to succeed every time. This resulted in CBT, as moving slowly would mean spending

too much of the limited time and consequently missing an opportunity to get a larger cumulative reward.

3. DISCUSSION

As outlined in the introduction, the model presented in this paper assumes that the CNS optimizes the *expected utility* of reaching movements. As Equation (2) shows, this expected utility is a function of three factors: the discounted reward resulting from reaching, which is itself a decreasing function of movement time, the probability to get this reward, which decreases with hit dispersion, and the cost of movement, which depends on the movement trajectory and timing. In order to optimize the expected utility, the CNS must find an optimal trade-off between these competing factors, by adjusting the instantaneous muscular activations.

We qualitatively analyzed three main aspects of reaching movements that are most relevant to CBT and SAT. First, the results show that the proposed model accounts for CBT as the velocity of movement is higher for more distant targets. This suggests that higher cost is permitted by the model/CNS to compensate for longer distances in order to reach the target in a reasonable time. Second, the results show that our model accounts for SAT by following Fitts' law, as the reaching time increases while the target size decreases to compensate for the required higher accuracy. Third, the results show that the model accounts for stochasticity of movement due to motor noise, as the target is sometimes missed and the hit position frequency on the target follows Gaussian distribution. Thus, even if our model does not quantitatively account for the reaching time of subject (see section 3.4.1 for discussion), the good qualitative match between experimental results and the behavior of our model strongly suggests that subjects do optimize their reaching time with respect to the global expected utility of their movement, i.e., taking the probability to miss into account. Additional results are in Supplementary Data and which point out that the proposed model can account for several other phenomena observed in the experimental data, such as: asymmetric muscle effort cost with respect to the initial point, and tendency to hit the target perpendicularly.

3.1. Our Model Simultaneously Accounts for Both CBT and SAT

The models presented in Shadmehr et al. (2010) and Rigoux and Guigon (2012) only related discounted reward and movement cost to explain the time of movement. In particular, motor control results in Rigoux and Guigon (2012) were obtained in the absence of sensory and motor noise. This model could explain several well-established motor control facts, as expected from the close relationship to previous optimal control models (Gordon et al., 1994; Shadmehr and Mussa-Ivaldi, 1994; Todorov and Jordan, 2002; Guigon et al., 2007, 2008a; Liu and Todorov, 2007), as well as several phenomena specific to CBT tasks (Watanabe et al., 2003; Rudebeck et al., 2006). However, the model in Rigoux and Guigon (2012) could not provide a direct account of phenomena relying on the stochasticity of the motor system, such as Fitts' law. Actually, their model provided an indirect account of Fitts' law (see, Rigoux and Guigon, 2012). For obtaining these results, the authors had to estimate dispersion as a function of velocity considering a constant velocity over the movement, and they reconstructed the relationship between the index of difficulty and the movement time based on a single starting point and the size of a unique target that would match this estimated dispersion (personal communication).

On the other hand, the model of Dean et al. only related discounted reward and accuracy, without consideration for movement cost (Dean et al., 2007). An abstract SAT model was directly fitted to human movement data, taking Fitts' law as a prior rather than explaining it. As such, the model could not account for several motor control phenomena related to the cost of movement.

In section 2.2, we have shown that our model accounts for Fitts' law. In contrast with the model of Dean et al. (2007), in our model the hit dispersion is measured as an effect of motor noise and imperfect state estimation along optimized trajectories, rather than inferred based on a given SAT model. Thus, one of the contributions is a model that addresses the more global inter-relationship between the discounted reward, the probability to reach, and the movement cost through an optimality criterion that accounts for the motor strategy of human subjects in this multi-dimensional choice space. Furthermore, simultaneously taking into account the three factors above endows our model with further properties, resulting in the contributions highlighted below.

3.2. Our Model Accounts for Hit Dispersion

In **Figure 4**, we observe that hit dispersion is smaller for smaller targets, even if there are still some failed movements. Our model puts forward two explanations on how the CNS might do so. First, the hit velocity is lower for smaller targets. Up to a certain level, motor noise can be reduced without slowing down by generating less co-contraction. However, below a certain threshold, less muscular activation implies a slower movement. Thus, the CNS achieves higher accuracy just by arriving slower at the target. Second, we observe in **Figure 2** that subjects start slowing down earlier when the target is smaller. By doing so, they give more time to the state estimation process to accurately estimate the end effector position, which is another way to reduce hit dispersion. It is quite likely that both mechanisms contribute to the necessary reduction in hit dispersion.

Recently, a study by Wang et al. (2016) empirically observed that human subjects preferred to execute the reaching movement with a time higher than the time at which the best endpoint variability was achieved. Based on this, they hypothesized that the human CNS tends to minimize both effort and endpoint variability. While their study excluded the visual feedback, their results are generally in line with the results of our study. However, Wang et al. (2016) did not devise any mathematical model to replicate their hypothesis.

3.3. Our Model Accounts for Asymmetric Bell-Shaped Velocity Profiles

The hit dispersion observed in **Figure 4** results from motor noise and imperfect state estimation. As a consequence of motor noise being proportional to muscular activation, one way to decrease motor noise is to decrease muscular activation. In the case of the minimum intervention principle (Todorov and Jordan, 2002), it results in the fact that the brain will not correct an error which is not related to task achievement. In the case of reaching movements, it will also result in a tendency to move fast earlier to be accurate later, when accuracy matters. Additionally, under the assumption that more accurate state estimation takes more time, a slower movement provides a better opportunity for state estimation to compensate for delayed feedback about the current position of the end effector. Taken together, both phenomena drive controller optimization toward generating less velocity by the end of the movement for a smaller target. So, one way to make sure to hit a small target would be to perform a slow reaching movement.

However, a slower movement results in a more discounted reward, thus the movement should nevertheless be as fast as possible. As a consequence, the best way to optimize reaching accuracy under temporal constraints is to be faster in the beginning of the movement and slower in the end. Such asymmetry was also observed in the literature (MacKenzie et al., 1987; Jean and Berret, 2017). This phenomenon can be attributed to the use of visual feedback, as it has been shown that the peak in velocity profiles occurs earlier in the movement when the visual feedback is available compared to when no visual feedback is available (Hansen et al., 2006; Burkitt et al., 2013). Thus, velocity profiles should be bell-shaped and asymmetric, as visible in **Figure 2**. By contrast, the model of Dean et al. assumes constant velocity (Dean et al., 2007) and the one from Rigoux and Guigon are bell-shaped, but not asymmetric (Rigoux and Guigon, 2012).

3.4. Potential Limitations

3.4.1. Discrepancy in Movement Times between Experimental and Simulation Data

The movement time discrepancy may be due to various factors:

(1) The presence of the haptic manipulator and an additional external damping, which significantly slows down the motion of the subjects and which is not accounted for by the model,
(2) Inaccuracies in the arm model, which does not account for the effects of muscular friction that slows down the arm,
(3) Inaccuracies in the state estimation process. As for the latter, if estimation is too good in the model, the simulated arm may go faster without generating too much inaccuracy.

However, the purpose of this study was not to precisely predict the movement times of the subjects, but to show that the model can account for both CBT and SAT. For clarity and generality, the model was simplified and did not include the external damping of the haptic manipulator. Therefore, the movement times of the model do not correspond to the movement times of the subjects, as the model could move faster due to the less resistance. In particular, any significant addition external damping increases the movement time due the increased effort (Tanaka et al., 2006), therefore such time movement discrepancy is in accordance with the literature.

Including more detailed or better matching models at various stages/aspects could indeed make the data look closer to each other. However, the primary focus of this study was to validate the hypothesis that the model is able to account for both SAT and CBT. In addition, a general model should be able to account for SAT and CBT independently of conditions. Therefore, trying to precisely match the simulations to any particular experimental condition (e.g., external damping) would not help to show such generality, nor would provide any additional validation of unification hypothesis.

3.4.2. Various Discounting of Reward through Time

The existing study did not explore all design choices of how the model should discount the reward through time that has been studied in the literature. The debate between diverse discounting approaches is a long standing one (see e.g., Green and Myerson, 1996; Berret and Jean, 2016). The model of Rigoux and Guigon (2012) uses an exponential discounting of the reward through time. Alternative models suggest linear (Hoff, 1994), quadratic (Shadmehr et al., 2010), or hyperbolic (Shadmehr et al., 2010) discounting approaches. The latter is also in line with studies of many other authors (e.g., Prévost et al., 2010). More recently, using an inverse optimal control approach, it was determined that the experimental "cost of time" for reaching movements would rather follow a sigmoidal function (Berret and Jean, 2016).

3.4.3. Two-Component Reaching Strategy

Several past studies observed that human reaching movements tend to follow two distinct phases (Woodworth, 1899; Meyer et al., 1988; Elliott et al., 2001, 2010). The initial phase, which usually constitutes most of the movement, is rapid and relatively predictable. In the second phase, when the target is approached, the movement is slowed down and the time-displacement profiles often have discontinuities, which reflect modifications to the trajectory (Elliott et al., 2001). While we observed a similar behavior in our subjects, the presented model is not able to explicitly account for such two-component reaching strategy.

3.4.4. Flat Target

In the classic Fitts's law, the target surface is assumed to be oriented perpendicularly with respect to the movement. In our study the target was flat and the hits could be done from different angles. Some insights about this aspect are presented in Supplementary Material. The experimental data showed that subjects hit the target from different angles, but the majority of hits were closer to the perpendicular direction. To account for this observed aspect, we extended the proposed model to include the perpendicularity cost [see (4) in section 4.2.1 for details].

3.4.5. Reliance on Selected Parameters and Optimization Process

The resolution of the optimization problem to acquire the controller relies on extensive numerical experiments with several tunable parameters. The optimization process can be time-consuming and selecting the suitable parameters is essential in acquiring a good solution. This can be viewed as one of the disadvantages of the proposed method.

3.5. Future Work

Beyond better fitting the data and studying the model properties through systematic variations of the meta-parameters, the main line in our research agenda consists in shifting from a motor control perspective to a motor learning perspective by focusing on the optimization process itself. Under a motor learning perspective, we might for instance study the evolution of trajectories, co-contraction, velocity profiles, etc. along the training process. We might also check whether our model accounts for already published motor learning studies such as the work of Izawa et al. (2008) or Diedrichsen et al. (2010). We might explore different discounting of the reward through time from the literature (Berret and Jean, 2016). In addition, we will try to extend the existing model to incorporate multiple-component reaching strategy (Elliott et al., 2017).

4. METHODS

4.1. Experimental Methods

4.1.1. Participants

Ten healthy male volunteers participated in the study. Their average age was 23.0 years (SD = 2.7 years), height 178.8 cm (SD = 4.1 cm) and body mass 77.4 kg (SD = 5.8 kg). All subjects were right-handed. Exclusion criteria were neurological, vestibular, locomotion, visual disorders and recent limb injuries (self-reported). This study was carried out in accordance with the recommendations of National Medical Ethics Committee Slovenia (NO. 112/06/13) with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the National Medical Ethics Committee Slovenia (NO. 112/06/13).

4.1.2. Apparatus

In most experiments where a subject has to reach a target, either in monkeys (e.g., Kitazawa et al., 1998) or humans (Trommershäuser et al., 2003, 2009; Battaglia and Schrater, 2007; Dean et al., 2007; Hudson et al., 2008), the target is displayed on a vertical screen and the subject performs a trajectory in the coronal or horizontal plane that the screen intercepts. Our experimental apparatus reproduces such a scene in the horizontal plane.

As depicted in **Figure 5A**, subjects sat on a chair in front of a TV screen and a 3 axes haptic manipulator (HM) (Haptic master Mk2, MOOG, Nieuw-Vennep, The Netherlands). The TV screen was located 2 m in front of the chair backrest. The subject's right wrist was immobilized and connected to the HM by means of the gimbal mechanism (ADL gimbal mechanism, MOOG, Nieuw-Vennep, The Netherlands). HM constrained the motion of the subject's right hand (hand from here on) to the horizontal plane at the height of the subject's shoulders. The subject's right elbow was suspended from the ceiling by a long string in order to compensate for the effect of gravity and to restrain the motion of the arm to the horizontal plane. In order to fix the position of the shoulders, the trunk of the subject was immobilized by tying it to the backrest of the chair.

Details about the performed trajectories are shown in **Figure 5B**. Subject's were allowed to move the hand from 15 initial circular areas with 10 mm diameter spread in front of them to a single target that was located symmetrically to their right shoulder on a virtual wall at a distance equal to 95% of their arm length (defined as the distance between shoulder and wrist). The initial areas were spread on three arcs with their center on the target on the virtual wall. The radius of the first arc was equal to 15% of the arm length and included 3 initial areas, the radius of the second arc was equal to 37.5% of the arm length and included 5 initial areas, and the radius of the third arc was equal to 60% of the arm length and included 7 initial areas. The initial areas on each arc were placed symmetrically with respect to the target on the virtual wall and were spread within the $\pm 45^{\circ}$. The initial areas were spread in the region well within the area of arm motion where the passive joint torques are negligible.

The size of the target on the virtual wall was either 5, 10, 20, or 40 mm wide. The position of the current initial area, the position of the target and the position of the subject's hand were drawn in 2D perspective on the screen in real time.

The experiment was divided into five parts corresponding to the four randomly selected sizes of the target, preceded by a familiarization trial with the target size of 20 mm. In each part, the randomly selected initial position from where the subject had to perform the motion was indicated in red. When the subject reached and stood within the initial position for more than 1 s, the color of the patch changed from red to green. This allowed the execution of the complete behavioral neuro-cognitive processes responsible for the SAT (Perri et al., 2014). When the patch turned green, the subject was indicated that the motion could be started. The time of the motion from the moment when the hand left the initial circular patch to the moment when the hand touched the virtual wall was recorded and subtracted from the total time of 100 s. The remaining time was displayed at the bottom of the screen. Success in hitting the target on the virtual wall was clearly indicated by changing the color of the virtual wall from gray to green for 1*s*. Besides, the subject obtained a money award of 2.5 euro cents. If the target was missed, the virtual wall turned red and no reward was given to the subject. The cumulative reward was displayed on top of the screen using large bold fonts.

The task of the subject was to obtain as high reward as possible in the given time. The cost of movement for each starting point was calculated by using muscle activity measurements obtained from EMG. The results regarding the movement cost are presented in Supplementary Material. To amplify the effect of the cost of motion, the haptic manipulator emulated a viscous media through which the subject had to move the hand.

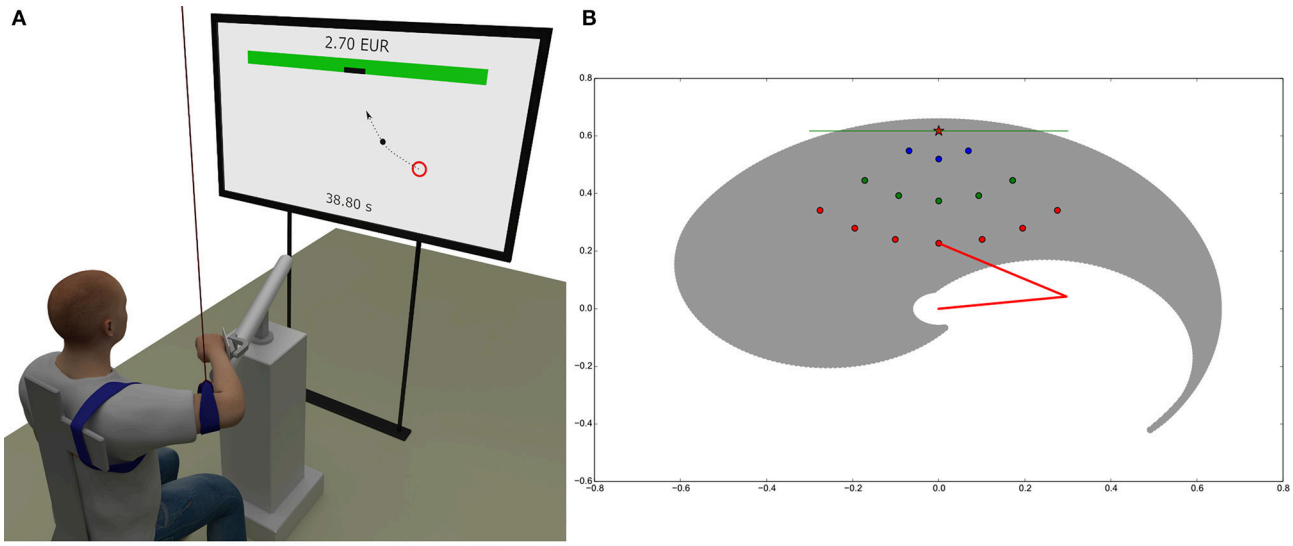

FIGURE 5 | Experimental and simulation setup. **(A)** Subjects sat on a chair in front of a TV screen with the right wrist connected to the haptic manipulator that constrained the motion of the right hand to the horizontal plane at the shoulder height. The screen displayed the current position of the subject's hand, the randomly occurring initial circular areas, the target on the virtual wall and gave feedback about the remaining time and the cumulative reward. **(B)** Simulated arm workspace and distribution of the 15 initial circular areas relative to the target on the wall. The reachable space is delimited by a spiral-shaped envelope (gray). The two segments of the arm are represented by red lines. The screen is represented as a green line positioned at $y = 0.6175m$ and the target center as a red star. The origin of the arm is at $x = 0.0, y = 0.0$. Initial areas are organized into three sets at different distances to the target, corresponding to 15% (3 blue circles), 37.5% (5 green circles), and 60% (7 red circles). The color code of the dots depending on the distance to the target is reused throughout the paper (in Fitts' law and velocity profiles diagrams).

The coefficient of viscous friction was set to $30Nm^{-1}s$. In the analysis, the muscular activations necessary to compensate for the friction of the haptic manipulator at the end-effector were removed by estimating them through the arm model described in section 4.2.5 and a model of the haptic manipulator friction.

4.2. Computational Methods

The computational model consists of a simulation set-up, an arm model, a state estimator, a set of controllers and a way to optimize these controllers. All the meta-parameters of these various components are summarized in **Table 3**, apart from arm model parameters which are given in section 4.2.5. Though the model is formally described using continuous time, in all the computational methods time is discretized with a time step $\delta_t = 0.002s$.

4.2.1. Specific Computational Model

The model described in the introduction was general. Here, we give a more specific account of the computational model of arm reaching movement which is actually used in the simulations.

Instead of (2), the complete cost function that we optimize is

$$J(\boldsymbol{u}) = \mathbb{E}(\boldsymbol{s}, \boldsymbol{u})\left(\int_0^\infty \left[e^{-t/\gamma}\rho R(\boldsymbol{s}_t) - \upsilon L(\boldsymbol{u}_t)\right]dt + C(\boldsymbol{s}_f)\right), \quad (4)$$

where the perpendicularity cost $C(\boldsymbol{s}_f)$ accounts for the tendency of subjects to hit the target perpendicularly and penalizes the scalar product between a vector colinear to the target and the Cartesian velocity of the end effector when the arm hits the screen. The necessity of this additional term is discussed in Supplementary Material.

TABLE 3 | Meta-parameters of the computational model.

Cost function	ρ	Immediate reward factor	3,000
	υ	Effort term factor	1
	γ	Discount factor	0.6
CMA-ES	σ	Initial covariance factor	0.01
	max_iter	Maximum number of iterations	5,000
	popsize	Population size	30
	repet	Number of repetitions	50
State estimation	κ	Multiplicative muscular noise	0.3
	Δ	Sensory delay in time steps	10
	k_1	Open-loop term factor in state estimation	0.2
	k_2	Closed-loop term factor in state estimation	1.0
Controller	N_h	Number of neurons in hidden layer	10

There are three meta-parameters in (4): the continuous-time discount factor γ accounts for the "greediness" of the controller, i.e., the smaller γ, the more the agent is focused on short term rewards, ρ is the weight of the reward term and υ is the weight of the effort term. Their value is given in **Table 3**.

4.2.2. Controllers and Optimization

In order to evaluate the criterion defined in (4), we need a controller which optimizes it. Actually, the optimal control

problem arising from a cost function including an expectation cannot be solved analytically. The utility expectation itself as defined in (4) must be estimated empirically as an average of the return over a set of attempts to hit the target (these attempts are called "rollouts" hereafter). The more rollouts, the better the estimate but we have to limit their number because computation takes time.

We thus rely on a stochastic optimization process, where a controller is represented as a parametric function whose parameters are tuned to optimize (4). The class of functions considered here consists of multi-layer neural networks with one input layer, one output layer and a hidden layer. In order to facilitate optimization by decoupling the parameter optimization problems for each trajectory, we define one such network for each starting point.

These controllers take the state of the system as input and provide muscular activations for all muscles in the arm model. As described in section 4.2.4, states are 4D and muscular activations are 6D, thus we have $4 + N_h + 6$ neurons in the networks, where the meta-parameter N_h is the number of neurons in the hidden layer.

All networks are initialized randomly with all weights and biases taken in $[0, 0.1]$. They are then optimized with respect to the approximated utility expectation described above using a state-of-the-art black-box optimization tool named CMA-ES (Hansen et al., 2003). Given an initial random controller, CMA-ES optimizes its parameters with local stochastic search. New rollouts are performed with varying parameters for all parameters around those of the current controllers, and the parameters that give rise to a better performance with respect to the cost function (4) are retained in the new current controller. In practice, the parameters are the weights and biases of each neuron in the networks.

CMA-ES comes with four meta-parameters: the initial size σ of a covariance matrix used for exploration, the size of the population *popsize*, the number of repetitions *repet* for each trajectory to get a decent estimation of the reward expectation and the maximum number of iterations *max_iter*. The values of these meta-parameters are given in **Table 3**.

4.2.3. State Estimation

A widely accepted overview of the human motor control loop is depicted in **Figure 6A**. According to this view, human movements are performed in closed loop. However, the sensory feedback being delayed, the control loop has to rely on state estimation to be stable (see **Figure 6B**).

Our implementation of the state estimator contains an estimate of the current state $\tilde{x}_t$ of the arm. The initial estimated state $\tilde{x}_0$ is the initial state of the arm x_0. Then, at each time step t, this estimated state is updated by the combination of two terms $\tilde{x}_t^1$ and $\tilde{x}_t^2$.

The first term corresponds to an open-loop estimation. At each time step, it simply updates the estimated state by applying the forward model of the arm to the previous estimated state, given the efferent copy of the motor command sent to the arm (before adding motor noise), i.e.,

$$\tilde{x}_t^1 = FM(\tilde{x}_{t-1}^1, u_{t-1}).$$

This term being open-loop and the motor command being different from the noisy command that has actually been applied to the arm, it generates an estimated state trajectory that may eventually drift away from the actual trajectory of the end effector.

The second term is in charge of closing the estimation loop by making profit of the delayed sensory feedback. The actual state of the arm $x_{t-\Delta}$ is available after a sensory delay of Δ time steps.

A first, naive approach to this second term is the following. Given this known state $x_{t-\Delta}$ and the Δ efferent copy of the commands $u_{t-\Delta}$ to u_{t-1} sent to the arm in between, one can infer an estimate of the current state $\tilde{x}_t^2$ by simply applying the forward model of the arm Δ times starting from $x_{t-\Delta}$ given the sequence of commands, using

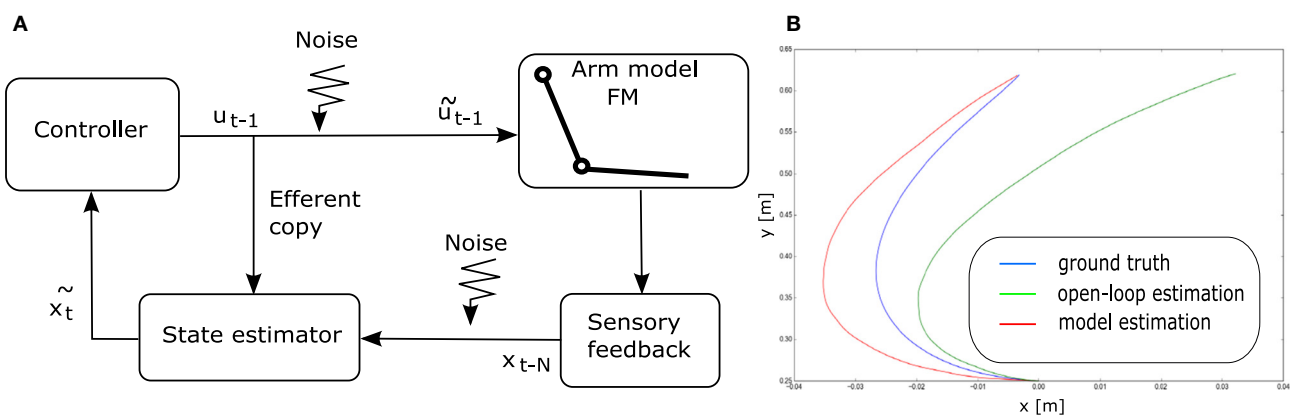

FIGURE 6 | (A) A standard schematic view of human motor control (see e.g., Scott, 2004). **(B)** Illustration of the state estimation dynamics. The open-loop term (in green) diverges from the perturbation-free "ground truth" trajectory (in blue). The model state estimation (in red) diverges in the accelerating phase of the movement, but the closed-loop term makes it converge again toward ground truth when the end-effector decelerates.

$$\begin{aligned}\tilde{x}^2_{t-\Delta-1} &= FM(x_{t-\Delta}, u_{t-\Delta}),\\ \tilde{x}^2_{t-\Delta-2} &= FM(\tilde{x}^2_{t-\Delta-1}, u_{t-\Delta-1}),\\ &\dots\\ \tilde{x}^2_{t} &= FM(\tilde{x}^2_{t-1}, u_{t-1}).\end{aligned}$$

Again, the obtained state estimation $\tilde{x}^2_t$ is not perfectly accurate given that the available efferent copy of the commands do not incorporate the applied motor noise. However, this term counterbalances the drifting tendency of the open-loop term because it updates the current estimate from some delayed ground truth. This approach neglects sensory noise. As a result, it is simpler than the standard state estimation model using an extended Kalman filter, as described in Guigon et al. (2008b).

However, this simpler approach still suffers from its computational cost, because it requires Δ iterations of the forward model of the arm. In order to improve the computational efficiency of the model, we call upon a neural network to replace these Δ iterations by a single function call. The neural network learns to predict the current estimated step given the delayed ground truth $x_{t-\Delta}$ and the Δ efferent copies that were received in between.

We made sure through dedicated simulations (not shown) that the learned neural network provides a reasonably accurate estimate of the current state, corresponding to what Δ iterations of the forward model would have inferred, at a much lower computational cost and with an increased biological implementation plausibility.

Finally, we combine the open-loop and the closed loop term by a weighted summation, using

$$\tilde{x}_t = (k_1 \tilde{x}_t{}^1 + k_2 \tilde{x}_t{}^2)/(k_1 + k_2).$$

The delay in number of time steps Δ as well as coefficients k_1 and k_2 are meta-parameters described in **Table 3**. They are tuned so that the estimated state can significantly differ from the true state when the arm is moving fast, but can become more accurate again when the arm slows down, so as to give a chance to hit the target. This tuning process was performed empirically over a small set of trajectories before starting to optimize the controller parameters.

4.2.4. Simulation Set-Up

The state-space consists of the current articular position $\mathbf{q}$ of the arm and its current articular speed $\dot{\boldsymbol{q}}$. The state $\boldsymbol{s} = (\boldsymbol{q}, \dot{\boldsymbol{q}})$ has a total of 4 dimensions. The initial state is defined by null speed and a variable initial position. The positions are bounded to represent the reachable space of a standard human arm, with $q_1 \in [2.6, -0.6]$ and $q_2 \in [3.0, -0.2]$, as shown in **Figure 5B**. The action-space consists of an activation signal for each muscle, resulting in a total of 6 dimensions.

The target is defined as an interval of varying length around $(x = 0, y = 0.6175m)$. The movement is stopped once the line $y = 0.6175m$ has been crossed, and the intersect between the trajectory and this line is computed to determine whether the target was hit. The reward for immediately hitting the target without taking incurred costs into account depends on the meta-parameter ρ (see **Table 3**).

4.2.5. Arm Model

The plant is a two degrees-of-freedom (DoFs) planar arm controlled by 6 muscles, illustrated in **Figure 7**. There are several such models in the literature. The model described in Kambara et al. (2009) lies in the coronal plane so it takes the gravity force into account. Most other models are defined in the horizontal plane and ignore gravity effects. They all combine a simple

TABLE 4 | Arm parameters.

m_1	Arm mass (*kg*)	1.4
m_2	Forearm mass (*kg*)	1.1
l_1	Arm length (*m*)	0.3
l_2	Forearm length (*m*)	0.35
I_1	Arm inertia ($kg.m^2$)	0.11
I_2	Forearm inertia ($kg.m^2$)	0.16
s_1	Distance from the center of segment 1 to its center of mass (*m*)	0.025
s_2	Distance from the center of segment 2 to its center of mass (*m*)	0.045

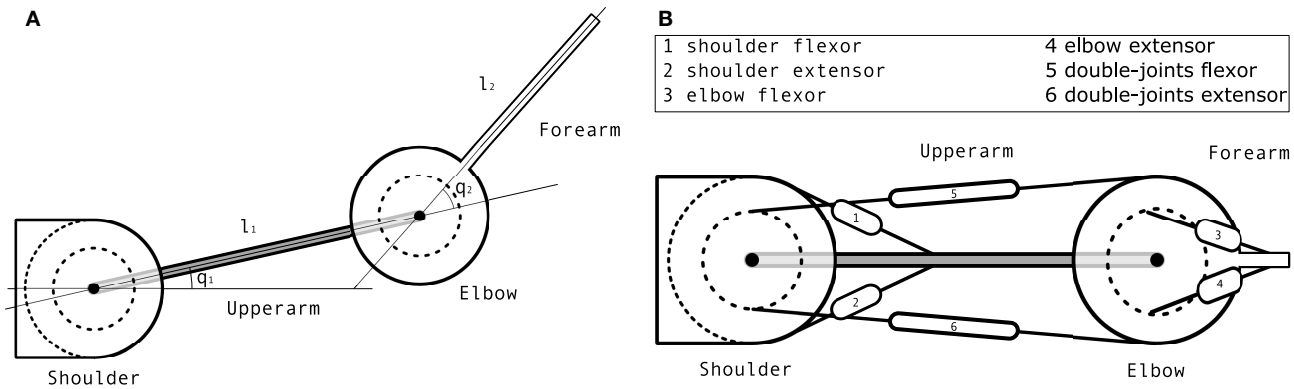

FIGURE 7 | Arm model. **(A)** Schematic view of the arm mechanics. **(B)** Schematic view of the muscular actuation of the arm, where each number represents a muscle whose name is in the box.

TABLE 5 | Nomenclature of the arm model parameters.

m_i	Mass of segment *i* (*kg*)
l_i	Length of segment *i* (*m*)
I_i	Inertia of segment *i* ($kg.m^2$)
s_i	Distance from the center of Segment *i* to its center of mass (*m*)
$\boldsymbol{A}$	Moment arm matrix ($\in \mathbb{R}^{6\times 2}$)
$\boldsymbol{f_{max}}$	Maximum muscular tension ($\in \mathbb{R}^6$)
$\boldsymbol{M}$	Inertia matrix ($\in \mathbb{R}^{2\times 2}$)
$\boldsymbol{C}$	Coriolis force ($N.m \in \mathbb{R}^2$)
$\boldsymbol{\tau}$	Segments torque ($N.m \in \mathbb{R}^2$)
$\boldsymbol{B}$	Damping term ($N.m \in \mathbb{R}^2$)
$\boldsymbol{u}$	Raw muscular activation (action) ($\in [0,1]^6$)
$\boldsymbol{\kappa}$	Multiplicative muscular noise ($\in \mathbb{R}$)
$\boldsymbol{\tilde{u}}$	Noisy muscular activation ($\in [0,1]^6$)
$\boldsymbol{q}^*$	Target articular position ($rad \in [0, 2\pi[^2$)
$\boldsymbol{q}$	Current articular position ($rad \in [0, 2\pi[^2$)
$\dot{\boldsymbol{q}}$	Current articular speed ($rad.s^{-1}$)
$\ddot{\boldsymbol{q}}$	Current articular acceleration ($rad.s^{-2}$)

two DoFs planar rigid-body dynamics model with a muscular actuation model. The differences between models mostly lie in the latter component. Our muscular actuation model is taken from Katayama and Kawato (1993) (pp. 356–357) through (Mitrovic et al., 2008). It is a simplified version of the one described in Li (2006) in the sense that it uses a constant moment arm matrix **A** whereas Li (2006) is computing this matrix as a function of the state of the arm. This arm model is used in simulation through a standard Euler integration method.

The rigid-body dynamics equation of a mechanical system is

$$\ddot{\boldsymbol{q}} = \boldsymbol{M}(\boldsymbol{q})^{-1}(\boldsymbol{\tau} - \boldsymbol{C}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} - \boldsymbol{g}(\boldsymbol{q}) - \boldsymbol{B}\dot{\boldsymbol{q}}) \tag{5}$$

where $\mathbf{q}$ is the current articular position, $\dot{\boldsymbol{q}}$ the current articular speed, $\ddot{\boldsymbol{q}}$ the current articular acceleration, $\mathbf{M}$ the inertia matrix, $\mathbf{C}$ the Coriolis force vector, $\boldsymbol{\tau}$ the segments torque, $\mathbf{g}$ the gravity force vector and $\mathbf{B}$ a damping term that contains all unmodelled effects. Here, $\mathbf{g}$ is ignored since the arm is working in the horizontal plane.

The inertia matrix is computed as $\mathbf{M} = \begin{bmatrix} k_1 + 2k_2\cos(q_2) & k_3 + k_2\cos(q_2) \\ k_3 + k_2\cos(q_2) & k_3 \end{bmatrix}$, with $k_1 = I_1 + I_2 + m_2 l_1^2$, $k_2 = m_2 l_1 s_2$ and $k_3 = I_2$, where m_i is the mass of segment i, l_i the length of segment i, I_i the inertia of segment i and s_i the distance from the center of segment i to its center of mass. The value of all these parameters is given in **Table 4**. They have been taken from the literature.

The Coriolis force vector is $\mathbf{C} = \begin{bmatrix} -\dot{q}_2(2\dot{q}_1 + \dot{q}_2) \\ \dot{q}_1{}^2 \end{bmatrix} k_2 \sin(q_2)$.

The damping matrix $\mathbf{B}$ is defined as $\mathbf{B} = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}\dot{\boldsymbol{q}} = \begin{bmatrix} 0.05 & 0.025 \\ 0.025 & 0.05 \end{bmatrix}\dot{\boldsymbol{q}}$.

The computation of the torque $\boldsymbol{\tau}$ exerted on the system given an input muscular actuation $\mathbf{u}$ is as follows. First, the muscular activation is augmented with Gaussian noise using $\tilde{\mathbf{u}} = \mathbf{u}_t \times (1 + \kappa\mathcal{N}(0, \mathbf{I}))$, where $\times$ refers to the element-wise multiplication and $\mathbf{I}$ is a 6×6 identity matrix. Then, the input torque is computed as $\boldsymbol{\tau} = \mathbf{A}^\top(\mathbf{f_{max}} \times \tilde{\mathbf{u}})$, where the moment arm matrix $\mathbf{A}$ is defined as

$$\mathbf{A}^\top = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} \end{bmatrix} = \begin{bmatrix} 0.04 & -0.04 & 0.0 & 0.0 & 0.028 & -0.035 \\ 0.0 & 0.0 & 0.025 & -0.025 & 0.028 & -0.035 \end{bmatrix},$$

and the matrix of the maximum force exerted by each muscle is defined as

$$\mathbf{f_{max}} = \begin{pmatrix} 700 & 0 & 0 & 0 & 0 & 0 \\ 0 & 382 & 0 & 0 & 0 & 0 \\ 0 & 0 & 572 & 0 & 0 & 0 \\ 0 & 0 & 0 & 445 & 0 & 0 \\ 0 & 0 & 0 & 0 & 159 & 0 \\ 0 & 0 & 0 & 0 & 0 & 318 \end{pmatrix}.$$

The nomenclature of all the parameters and variables of the arm model is given in **Table 5**.

AUTHOR CONTRIBUTIONS

LP, OS, and JB contributed to conception/design of the work, and the acquisition, analysis and interpretation of data. LP, OS, and JB contributed to drafting/writing/revising the paper. LP, OS and JB gave the final approval of the version to be published and agreed to be accountable for all aspects of the work.

FUNDING

This work was partially supported by the European Commission, within the CoDyCo project (FP7-ICT-2011-9, No.600716).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2017.00615/full#supplementary-material

REFERENCES

Battaglia, P. W., and Schrater, P. R. (2007). Humans trade off vieweing time and movement duration to improve visuomotor accuracy in a fast reaching taks. *J. Neurosci.* 27, 6984–6994. doi: 10.1523/JNEUROSCI.1309-07.2007

Berret, B., and Jean, F. (2016). Why don't we move slower? the value of time in the neural control of action. *J. Neurosci.* 36, 1056–1070. doi: 10.1523/JNEUROSCI.1921-15.2016

Bootsma, R. J., Fernandez, L., and Mottet, D. (2004). Behind Fitts law: kinematic patterns in goal-directed movements. *Int. J. Hum. Comput. Stud.* 61, 811–821. doi: 10.1016/j.ijhcs.2004.09.004

Bootsma, R. J., Marteniuk, R. G., MacKenzie, C. L., and Zaal, F. T. (1994). The speed-accuracy trade-off in manual prehension: effects of movement amplitude, object size and object width on kinematic characteristics. *Exp. Brain Res.* 98, 535–541. doi: 10.1007/BF00233990

Burkitt, J. J., Grierson, L. E., Staite, V., Elliott, D., and Lyons, J. (2013). The impact of prior knowledge about visual feedback on motor performance and learning. *Adv. Phys. Educ.* 3:1. doi: 10.4236/ape.2013.31001

Crossman, E. R., and Goodeve, P. J. (1983). Feedback control of hand-movement and fitts' law. *Q. J. Exp. Psychol.* 35, 251–278. doi: 10.1080/14640748308402133

Dean, M., Wu, S.-W., and Maloney, L. T. (2007). Trading off speed and accuracy in rapid, goal-directed movements. *J. Vis.* 7, 1–12. doi: 10.1167/7.5.10

Diedrichsen, J., White, O., Newman, D., and Lally, N. (2010). Use-dependent and error-based learning of motor behaviors. *J. Neurosci.* 30, 5159–5166. doi: 10.1523/JNEUROSCI.5406-09.2010

Elliott, D., Hansen, S., Grierson, L. E., Lyons, J., Bennett, S. J., and Hayes, S. J. (2010). Goal-directed aiming: two components but multiple processes. *Psychol. Bull.* 136:1023. doi: 10.1037/a0020958

Elliott, D., Helsen, W. F., and Chua, R. (2001). A century later: Woodworth's (1899) two-component model of goal-directed aiming. *Psychol. Bull.* 127:342. doi: 10.1037/0033-2909.127.3.342

Elliott, D., Lyons, J., Hayes, S. J., Burkitt, J. J., Roberts, J. W., Grierson, L. E., et al. (2017). The multiple process model of goal-directed reaching revisited. *Neurosci. Biobehav. Rev.* 72, 95–110. doi: 10.1016/j.neubiorev.2016.11.016

Faisal, A. A., Selen, L. P. J., and Wolpert, D. M. (2008). Noise in the nervous system. *Nat. Rev. Neurosci.* 9, 292–303. doi: 10.1038/nrn2258

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* 47, 381–391. doi: 10.1037/h0055392

Gordon, J., Ghilardi, M. F., Cooper, S. E., and Ghez, C. (1994). Accuracy of planar reaching movements. *Exp. Brain Res.* 99, 112–130. doi: 10.1007/BF00241416

Green, L., and Myerson, J. (1996). Exponential versus hyperbolic discounting of delayed outcomes: risk and waiting time. *Am. Zool.* 36, 496–505. doi: 10.1093/icb/36.4.496

Green, L., and Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychol. Bull.* 130:769. doi: 10.1037/0033-2909.130.5.769

Guigon, E., Baraduc, P., and Desmurget, M. (2007). Computational motor control: redudancy and invariance. *J. Neurophysiol.* 97, 331–347. doi: 10.1152/jn.00290.2006

Guigon, E., Baraduc, P., and Desmurget, M. (2008a). Computational motor control: feedback and accuracy. *Eur. J. Neurosci.* 27, 1003–1016. doi: 10.1111/j.1460-9568.2008.06028.x

Guigon, E., Baraduc, P., and Desmurget, M. (2008b). Optimality, stochasticity and variability in motor behavior. *J. Comput. Neurosci.* 24, 57–68. doi: 10.1007/s10827-007-0041-y

Hansen, N., Muller, S., and Koumoutsakos, P. (2003). Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evol. Comput.* 11, 1–18. doi: 10.1162/106365603321828970

Hansen, S., Glazebrook, C. M., Anson, J. G., Weeks, D. J., and Elliott, D. (2006). The influence of advance information about target location and visual feedback on movement planning and execution. *Can. J. Exp. Psychol.* 60:200. doi: 10.1037/cjep2006019

Harris, C. M., and Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature* 394, 780–784. doi: 10.1038/29528

Hoff, B. (1994). A model of duration in normal and perturbed reaching movement. *Biol. Cybernet.* 71, 481–488. doi: 10.1007/BF00198466

Hudson, T. E., Maloney, L. T., and Landy, M. S. (2008). Optimal compensation for temporal uncertainty in movement planning. *PLoS Comput. Biol.* 4:e1000130. doi: 10.1371/journal.pcbi.1000130

Izawa, J., Rane, T., Donchin, O., and Shadmehr, R. (2008). Motor adaptation as a process of reoptimization. *J. Neurosci.* 28, 2883–2891. doi: 10.1523/JNEUROSCI.5359-07.2008

Jean, F., and Berret, B. (2017). "On the duration of human movement: from self-paced to slow/fast reaches up to fitts's law," in *Geometric and Numerical Foundations of Movements*, eds J.-P. Laumond, N. Mansard, and J.-B. Lasserre (Cham: Springer), 43–65.

Kambara, H., Kim, K., Shin, D., Sato, M., and Koike, Y. (2009). Learning and generation of goal-directed arm reaching from scratch. *Neural Netw.* 22, 348–361. doi: 10.1016/j.neunet.2008.11.004

Katayama, M., and Kawato, M. (1993). Virtual trajectory and stiffness ellipse during multijoint arm movement predicted by neural inverse models. *Biol. Cybernet.* 69, 353–362. doi: 10.1007/BF01185407

Keele, S. W. (1968). Movement control in skilled motor performance. *Psychol. Bull.* 70:387. doi: 10.1037/h0026739

Kitazawa, S., Kimura, T., and Yin, P.-B. (1998). Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature* 392, 494–497. doi: 10.1038/33141

Laurent, M. (1994). Fitts' law in two-dimensional task space. *Exp. Brain Res.* 100, 144–148.

Li, W. (2006). *Optimal Control for Biological Movement Systems*. PhD thesis, University of California, San Diego, CA.

Liu, D., and Todorov, E. (2007). Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *J. Neurosci.* 27, 9354–9368. doi: 10.1523/JNEUROSCI.1110-06.2007

MacKenzie, C. L., Marteniuk, R., Dugas, C., Liske, D., and Eickmeier, B. (1987). Three-dimensional movement trajectories in fitts' task: Implications for control. *Q. J. Exp. Psychol.* 39, 629–647. doi: 10.1080/14640748708401806

Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., and Keith Smith, J. (1988). Optimality in human motor performance: ideal control of rapid aimed movements. *Psychol. Rev.* 95:340. doi: 10.1037/0033-295X.95.3.340

Mitrovic, D., Klanke, S., and Vijayakumar, S. (2008). "Adaptive optimal control for redundantly actuated arms," in *Proceedings of the Tenth International Conference on Simulation of Adaptive Behavior* (Osaka), 93–102.

Perri, R. L., Berchicci, M., Spinelli, D., and Di Russo, F. (2014). Individual differences in response speed and accuracy are associated to specific brain activities of two interacting systems. *Front. Behav. Neurosci.* 8:251. doi: 10.3389/fnbeh.2014.00251

Plamondon, R. (1991). On the origin of asymmetric bell-shaped velocity profiles in rapid-aimed movements. *Tutorials Motor Neurosci.* 62, 283–295.

Plamondon, R., and Alimi, A. M. (1997). Speed/accuracy trade-offs in target-directed movements. *Behav. Brain Sci.* 20, 279–303; discussion: 303–349. doi: 10.1017/S0140525X97001441

Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M., and Dreher, J. (2010). Separate valuation subsystems for delay and effort decision costs. *J. Neurosci.* 30, 14080–14090. doi: 10.1523/JNEUROSCI.2752-10.2010

Qian, N., Jiang, Y., Jiang, Z.-P., and Mazzoni, P. (2013). Movement duration, fitts's law, and an infinite-horizon optimal feedback control model for biological motor systems. *Neural Comput.* 25, 697–724. doi: 10.1162/NECO_a_00410

Rigoux, L., and Guigon, E. (2012). A model of reward and effort-based optimal decision making and motor control. *PLoS Comput. Biol.* 8:e1002716. doi: 10.1371/journal.pcbi.1002716

Rudebeck, P. H., Walton, M. E., Smyth, A. N., Bannerman, D. M., and Rushworth, M. F. (2006). Separate neural pathways process different decision costs. *Nat. Neurosci.* 9, 1161–1168. doi: 10.1038/nn1756

Schmidt, R. A., Zelaznik, H., Hawkins, B., Frank, J. S., and Quinn Jr, J. T. (1979). Motor-output variability: A theory for the accuracy of rapid motor acts. *Psychol. Rev.* 86:415. doi: 10.1037/0033-295X.86.5.415

Scott, S. H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nat. Rev. Neurosci.* 5, 532–546. doi: 10.1038/nrn1427

Scott MacKenzie, I. (1989). A note on the information-theoretic basis for fitts law. *J. Motor Behav.* 21, 323–330. doi: 10.1080/00222895.1989.10735486

Selen, L. P., Beek, P. J., and van Dieen, J. H. (2006). Impedance is modulated to meet accuracy demands during goal-directed arm movements. *Exp. Brain Res.* 172, 129–138. doi: 10.1007/s00221-005-0320-7

Shadmehr, R., and Mussa-Ivaldi, F. A. (1994). Adaptive representation of the dynamics during learning of a motor task. *J. Neurosci.* 14, 3208–3324.

Shadmehr, R., Orban de Xivry, J.-J., Xu-Wilson, M., and Shih, T.-Y. (2010). Temporal discounting of reward and the cost of time in motor control. *J. Neurosci.* 30, 10507–10516. doi: 10.1523/JNEUROSCI.1343-10.2010

Smyrnis, N., Evdokimidis, I., Constantinidis, T., and Kastrinakis, G. (2000). Speed-accuracy trade-off in the performance of pointing movements in different directions in two-dimensional space. *Exp. Brain Res.* 134, 21–31. doi: 10.1007/s002210000416

Soechting, J. F. (1984). Effect of target size on spatial and temporal characteristics of a pointing movement in man. *Exp. Brain Res.* 54, 121–132. doi: 10.1007/BF00235824

Tanaka, H., Krakauer, J. W., and Qian, N. (2006). An optimization principle for determining movement duration. *J. Neurophysiol.* 95, 3875–3886. doi: 10.1152/jn.00751.2005

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nat. Neurosci.* 7, 907–915. doi: 10.1038/nn1309

Todorov, E. (2005). Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural Comput.* 17, 1084–1108. doi: 10.1162/0899766053491887

Todorov, E., and Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.* 5, 1226–1235. doi: 10.1038/nn963

Trommershäuser, J., Maloney, L. T., and Landy, M. S. (2003). Statistical decision theory and the selection of rapid, goal-directed movements. *J. Opt. Soc. Am. Opt. Image Sci. Vision* 20, 1419–1433. doi: 10.1364/JOSAA.20.001419

Trommershäuser, J., Maloney, L. T., and Landy, M. S. (2009). The expected utility of movement. *Neuroeconomics*, 95–111. doi: 10.1016/B978-0-12-374176-9.00008-7

Wang, C., Xiao, Y., Burdet, E., Gordon, J., and Schweighofer, N. (2016). The duration of reaching movement is longer than predicted by minimum variance. *J. Neurophysiol.* 116, 2342–2345. doi: 10.1152/jn.00148.2016

Watanabe, K., Lauwereyns, J., and Hikosaka, O. (2003). Effects of motivational conflicts on visually elicited saccades in monkeys. *Exp. Brain Res.* 152, 361–367. doi: 10.1007/s00221-003-1555-9

Woodworth, R. S. (1899). *The Accuracy of Voluntary Movement.* New York, NY: Columbia University.

Young, W. B., and Bilby, G. E. (1993). The effect of voluntary effort to influence speed. *J. Strength Condition. Res.* 7, 172–178.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Peternel, Sigaud and Babič. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.