

Title: Winning the Cybathlon BCI race: Successful longitudinal training and control by two tetraplegic users

Authors: S. Perdikis^{1*}, L. Tonin¹, S. Saeedi¹, C. Schneider¹, J. d. R. Millán^{1*}

Affiliations:

¹Defitech Chair in Brain-Machine Interface (CNBI), École Polytechnique Fédérale de Lausanne (EPFL), Chemin des Mines 9, CH-1202, Geneva, Switzerland

*Corresponding authors, email: serafeim.perdakis@epfl.ch, jose.millan@epfl.ch

One Sentence Summary: Two spinal cord injury individuals in their chronic state have been capable of long-term, high-performing, non-invasive brain-control of a demanding electroencephalography-based, self-paced brain-computer interface application culminating in the gold medal and record setting during the first international Cybathlon competition.

Abstract: This work draws on the experience of the “Brain Tweaksters” team in the brain-computer interface (BCI) race discipline of the Cybathlon event, the first international competition for disabled pilots in control of robotic and other assistive technology, to demonstrate the maturity of self-paced, non-invasive BCI technology. We hereby provide rare hard evidence of successful longitudinal training and use of a non-invasive, self-paced BCI application in real-world and adverse conditions. Two severely impaired participants, both suffering from chronic spinal cord injury, were trained following a user-centered and mutual learning approach to control their avatar in a virtual race game by means of three mental commands. The maturity and effectiveness of such interfaces are showcased by our pilots’ impressive average and record performances in this testbed application, which compare favorably to “perfect” keyboard control. The superiority of our motor imagery (MI) BCI design and training methodology are substantiated by our team’s prevalence in the competition day including the win of the gold medal and the hold of the BCI race record. Importantly, thanks to longitudinal experimentation we have firmly established with both pilots the hypothesized existence of operant learning effects with respect to the application performances, the BCI command accuracy and the voluntary modulation of brain rhythms. We postulate that the results presented hereby yield solid proof of a considerable current translational impact of MI BCI.

[Main Text:]

Introduction

Brain-computer interface (BCI) technology encompasses systems implementing direct mind-control of devices by circumventing the natural human neuromuscular pathways for communication and control (1). Despite offering the possibility of cognitive and motor enhancement also relevant to able-bodied individuals, BCI has been always primarily envisioned as an assistive technology (AT) for people in paralysis. As such, the main end-user group comprises patients suffering severe motor impairments as a result of traumatic brain or spinal cord injuries (SCI), stroke and neurodegenerative diseases like amyotrophic lateral sclerosis (ALS) and muscular dystrophy (1, 2).

Since the first demonstration of BCI’s profound clinical potential (3), the vast majority of studies have pertained to methodological and technical considerations involving experimentation with able-bodied subjects. While these works can be largely credited with the field’s nowadays widely acknowledged versatility and technological maturity, they carry limited evidence regarding its translational impact. Restricting the scope to the case of BCI for communication and control, the

number of published works involving end-users in the last 20 years remains to date a modest double-digit figure (1).

Besides the overall lack of end-user evaluation and the practical difficulties still faced therein (4), the critical transition from the laboratory to the patients' homes is hampered by additional shortcomings. First, most end-user studies have been so far limited to only a single or a few experimental sessions. However, longitudinal assessment (3, 5–16) is crucial to overcome the proof-of-concept stage in BCI applicability, especially considering the fact that robustness is still one of the main challenges. Secondly, most such experimentation takes place in controlled laboratory or clinical surroundings and relies on the constant supervision of expert personnel (engineers, clinicians), thus hardly accounting for the real-world conditions an end-user and the caregivers will have to cope with in the actual home setting.

On the other hand, end-user evaluation has been adequately distributed across the different established BCI paradigms, a policy to be further pursued by the BCI community (2, 17), since different paradigms exhibit significant advantages in specific contexts and individuals, with none of them emerging as universally optimal. BCI for severe paralysis has initially mostly exploited electroencephalography (EEG)-based slow cortical potentials (SCPs) implementing neurofeedback techniques, a form of operant conditioning relying on brain signals (1, 3). Such interfaces provided the first concrete evidence of BCI's clinical applicability, although they required lengthy training periods. Event-related potential (ERP)-based paradigms like P300 and steady-state visual evoked potential (SSVEP) interfaces have been shown to provide high information transfer rates (ITR) in minimal training times for restoring communication abilities (10, 13, 15, 18–22). However, they usually require gaze control and, most importantly, they impose the timing of the interaction by decoding brain responses elicited by external stimuli. On the contrary, EEG-based interfaces depending on the voluntary modulation of sensorimotor rhythms (SMRs) by means of mental (usually, motor) imagery allow for fully self-paced interaction, shown with a large variety of BCI prototypes (4, 6–8, 11, 23–33). BCI systems employing single- and multi-neuron spiking activity or local field potentials (LFPs) have been early studied in humans (5) and have yielded recent breakthroughs in self-paced brain-controlled prosthetics and spellers (9, 12, 34–39). Invasive BCIs, as well as those based on electrocorticography (ECoG) (16, 40, 41) benefit from the higher signal-to-noise ratio (SNR) offered by the interface's closer proximity to the neural tissue to permit continuous control of several degrees-of-freedom (DOF). Yet, the issues still associated to surgically inserted implants (e.g., medical risk, biocompatibility, biostability) might slow down the wider adaptation of these technologies for long-term use. Last but not least, BCIs that take advantage of classical conditioning and functional near-infrared spectroscopy (fNIRS) signals have been recently proposed (14) for locked-in syndrome patients. Notwithstanding their remarkable success, low temporal resolution renders such systems rather unsuited for other user categories.

The present work demonstrates that SCI-inflicted individuals in their chronic state are capable of long-term motor imagery (MI) BCI control over a demanding application, exhibiting spectacular performances even under adverse conditions like those prevailing at the Cybathlon BCI race event. The Cybathlon has been the first ever international para-Olympics for disabled individuals in control of bionic AT and was held in Zurich, Switzerland on October 8th 2016. It aspired to accelerate the development of advanced AT, promote public awareness and increase user acceptance by bringing funding agencies, technology providers and end-users in contact and, most importantly, by placing the latter in the core of the development process (42). Specifically, two male individuals coded P1 and P2, tetraplegic and wheelchair-bound as a result of accident-inflicted SCI have been trained to operate a MI BCI for the Cybathlon BCI race as pilots of the “Brain Tweakers” team, the franchise that represented the Defitech Chair in Brain-Machine Interface (CNBI) laboratory of the Swiss Federal Institute of Technology (EPFL). The BCI race

consisted of four brain-controlled avatars competing in a virtual race game called “Brain Runners”, where up to three mental commands could be issued on color-coded track segments (“pads”) to accelerate one's avatar (Fig. 1A). Additionally, a fourth type of pad required “idling” to avoid any command delivery. Erroneous and false positives commands would slow down the pilot's course towards the finish line of the track.

Aside from further adding upon the efforts for end-user evaluation and replicability, our work offers a unique account of considerably longitudinal successful use of a SMR-based BCI device under such realistic scenarios. Beyond contributing substantial evidence for the maturity of non-invasive MI BCI in general, we claim that our pilots' performance in the Cybathlon BCI race constitutes a proof of the competency of the state-of-the-art methods and user training approach followed, in particular, mutual learning protocols and user-centered designs. Last but not least, in spite of sporadic evidences in the literature for the existence of learning effects akin to neurofeedback training in this type of interfaces, we argue that this study is the very first to showcase strong learning effects beyond reasonable doubt, at the neuroimaging, BCI output and application level.

Results

Cybathlon BCI race outcomes

The Brain Tweaksters team dominated the BCI race discipline of Cybathlon. After winning the rehearsal event (albeit with an able-bodied pilot on that occasion), our team managed to place both pilots in the final and win the gold medal of the actual competition against ten competitors (Table 1). Given a standard racing track with lower bound (perfect ternary control), upper bound (continuously flawed input) and no-response race completion time of 54 s, 240 s and 162 s, respectively, P1 qualified with 90.1 s, a performance that set the competition record, almost 32 s faster than the second-best time belonging to our second pilot P2 (122.5 s). In the final, the third-best competition record (125.3 s) was made by P2 to capture the gold medal. The closest records belonging to the pilots of competing teams were 132, 135, 136 and 146 s. The ensemble of our competitors' performances averaged 155.5 ± 18.0 s ($N=15$) and including the Brain Tweaksters performances the total average has been 150.6 ± 25.3 s ($N=19$). Isolating the best race of each competing pilot, the average performance further reduces to 146.0 ± 29.4 s ($N=11$), while 140.8 ± 29.9 s ($N=8$) was the average of races belonging to medal winning teams. P1 experienced a momentary loss of BCI control and had to compromise with the 4th place in the final (189.8 s).

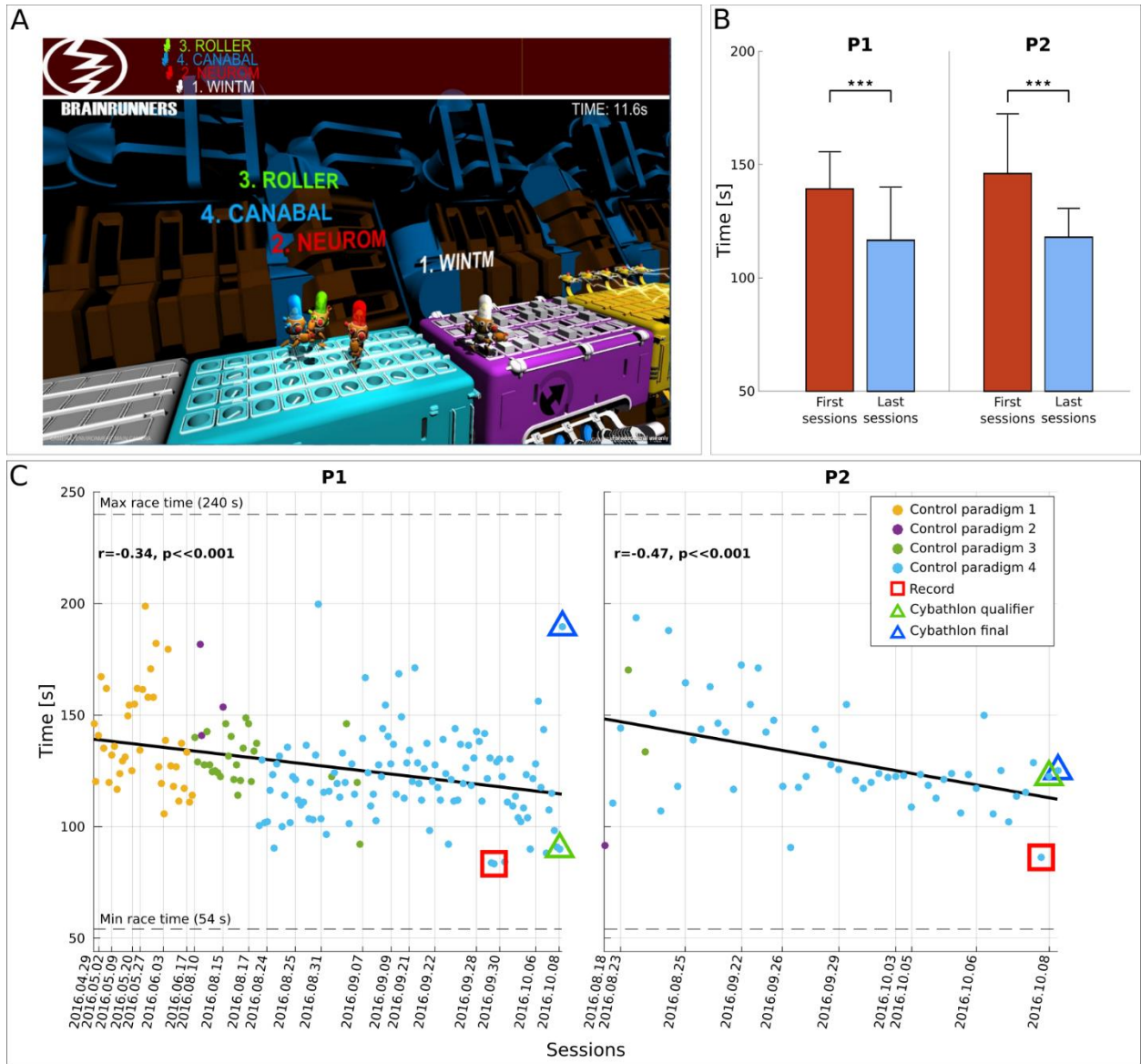


Fig. 1. Cybathlon BCI race track and race completion time. **(A)** Standard race track of Cybathlon's Brain Runners game graphical user interface. Pilots need to deliver the proper command in each color pad (cyan, magenta, yellow) in order to accelerate their own avatar. **(B)** Average and standard deviation of race completion time (s) for pilots P1 and P2 in the first (red) and last (blue) four training sessions including the competition day. Statistically significant differences are shown with two-sided Wilcoxon ranksum tests, (***): $p < .001$. **(C)** Race completion times (s) achieved by pilots P1 and P2 throughout training. The corresponding linear fits and Pearson correlation coefficients (significance extracted with Student's t-distribution) demonstrate training effects. Dashed horizontal lines illustrate the minimum and maximum race completion bounds of Cybathlon's BCI race standard track (perfect control and continuously flawed commands, respectively). Vertical lines indicate the date of each racing session. Marker colors show the control paradigm employed (see section User-centered BCI design and Materials and Methods). Record performances are highlighted with red squares. The competition performances are highlighted with triangles, green for the qualifier and blue for the final.

Primary outcome

Fig. 1B shows that our training procedure reduced the race completion time of P1 from 139.2 ± 16.1 s ($N=18$, first four racing sessions) to 116.5 ± 23.2 s ($N=34$, last four racing sessions

including the competition day) and similarly for P2 from 145.9 ± 26.1 s (N=22) to 117.9 ± 12.5 s (N=21). Both these improvements are statistically significant ($p < .001$, two-sided Wilcoxon ranksum tests). The race completion times of our pilots throughout training (Fig. 1C) averaged 126.9 ± 21.3 (N=182) for P1 and 130.3 ± 22.9 (N=57) s for P2, with all-time records of 83.3 and 86.3 s, respectively. Significant negative Pearson correlations between race time and (chronological) race index establish the existence of a significant training effect on race time (Fig. 1C, P1: $r = -0.34$, $p < .001$, N=182 P2: $r = -0.47$, $p < .001$, N=57). P1 achieved slightly higher average and record performances, while P2 exhibited a superior performance/stability trade-off through training, having race time standard deviation of 12.9 s in the last 5 sessions (N=28), as opposed to 20.6 s for P1 (N=50).

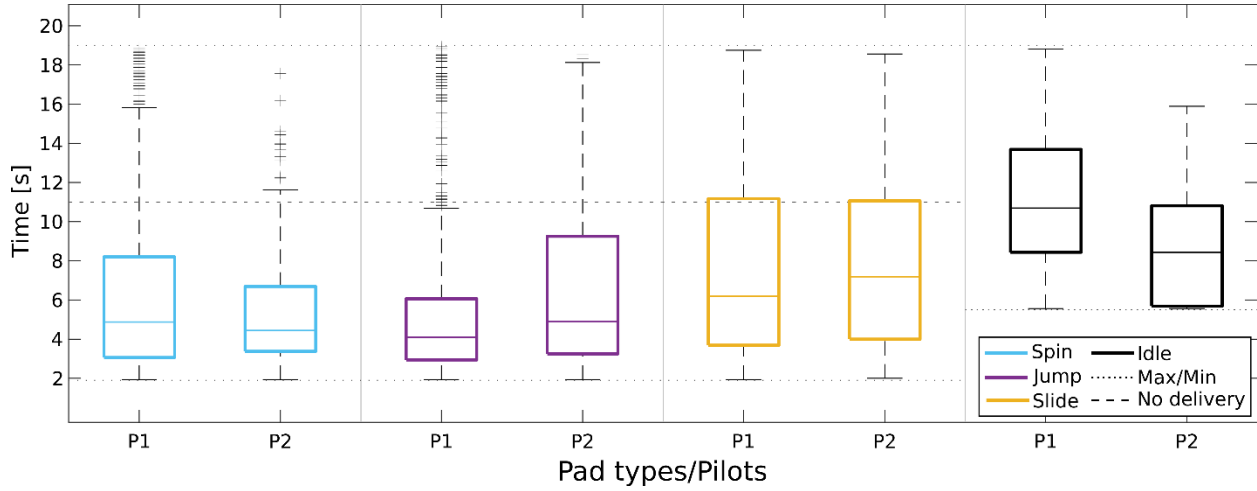


Fig. 2. Pad crossing time throughout training. Boxplots of pad crossing time (s, time spent on each pad) for pilots P1 and P2 and all types of pads (cyan for spin, magenta for jump, yellow for slide, black for idling). The box edges signify the 75th (top) and 25th (bottom) percentiles and the colored horizontal line the median of the corresponding distribution. The whiskers extend to the largest and smallest non-outlier values. Outliers are marked with black crosses. Dotted horizontal lines illustrate the minimum (accurate and precise BCI input), maximum (continuously erroneous BCI input) and no-delivery (unresponsive BCI, avatar goes at “base” speed) crossing times for the different pad types. The dashed line corresponds to the no-delivery time in the spin, jump and yellow pads.

BCI performances

Fig. 2 illustrates that the high-yielding application performances come as a result of our pilots' ability to adequately master all four individual sub-tasks required by the application: the intentional control (IC) ability to deliver the correct command on the action pads (spin, jump, slide) and the intentional non-control (INC) ability to “rest/idle” on the white pads (23, 33, 43). The illustrated median performances (for P1/P2) were 4.9 s (N=853)/4.4 s (N=205) for spin, 4.1 s (N=766)/4.9 s (N=198) for jump and 6.2 s (N=463)/7.2 s (N=196) for slide, which compare favorably to the lower bound (2 s), while lying far away from this metric's imposed upper bounds (11 s if no mental command is forwarded, 19 s for continuously erroneous command delivery) for all command types. Remarkably, a similar argument can be made for the INC ability. The median crossing time of white pads was 10.7 s and 8.4 s for P1 (N=510) and P2 (N=151), respectively, far below the worst case scenario of 19 s and approaching the optimum of 5.5 s. The race completion time amelioration can be justified by improvements in these sub-tasks.

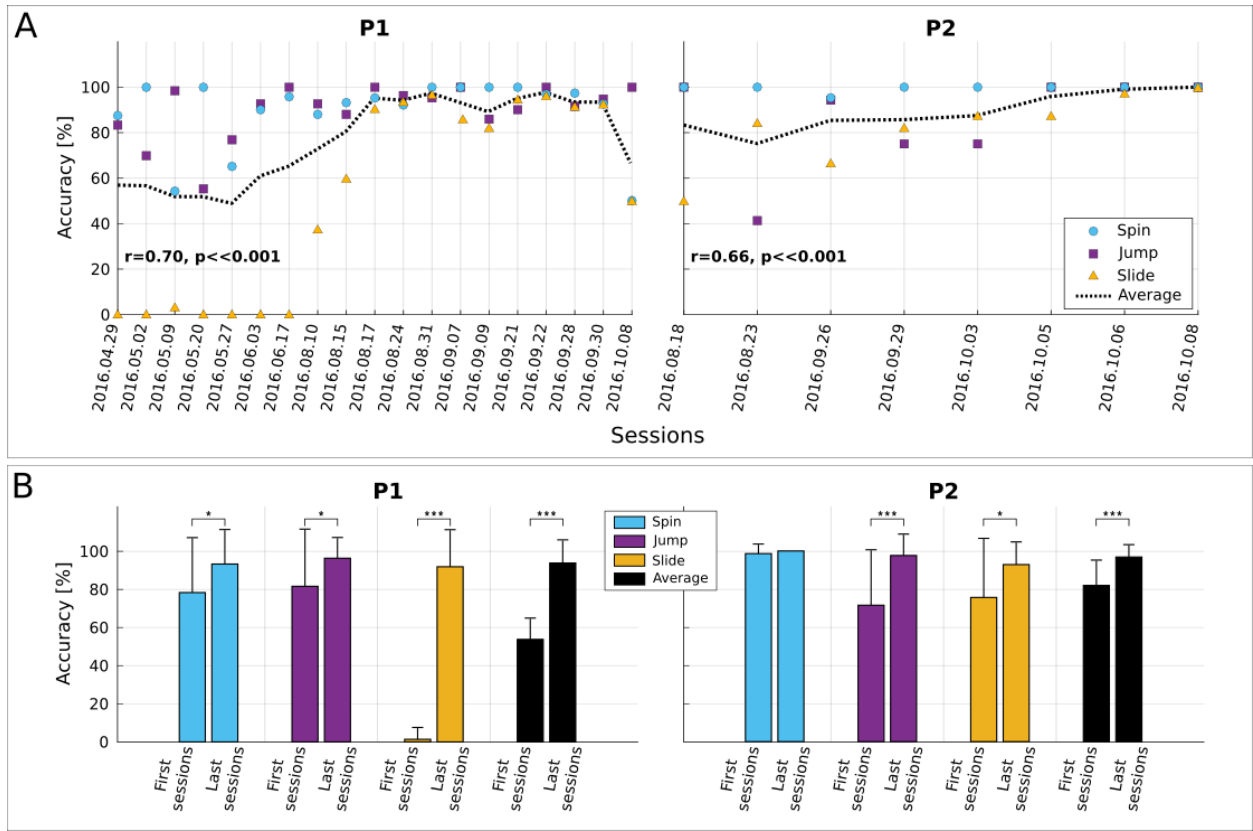


Fig. 3. BCI command accuracy. (A) Average within-session BCI command accuracy (%) for pilots P1 and P2. Spin command accuracy shown in cyan, jump in magenta and slide in yellow. The dashed black line shows the overall accuracy (average of individual command accuracies) in a session. The Pearson correlation between the overall command accuracy and the chronological race index is also reported (significance tested with Student's t-distribution). (B) Average and standard deviation of BCI command accuracy (%) for pilots P1 and P2, for all command types (cyan for spin, magenta for jump, yellow for slide) and overall (black) in the first and last four training sessions including the competition day. Statistically significant differences are shown with two-sided Wilcoxon ranksum tests, (*): $p < .05$, (***): $p < .001$.

Fig. 3A verifies increasing trends of command accuracy for both pilots and all command types. This can be quantified by significant positive correlations of the overall accuracy to the (chronological) race index (P1: $r=0.70$, $p < .001$, $N=162$ races and P2: $r=0.66$, $p < .001$, $N=45$ races). Fig. 3B showcases that the average total accuracy of P1 improved significantly from 53.8% ($N=18$) to 93.8% ($N=41$) and that of P2 from 81.9% ($N=24$) to 96.8% ($N=21$, $p < .001$ with two-sided Wilcoxon ranksum tests). Both pilots exhibited significant command accuracy increase in all individual tasks (the only exception being the spin command for P2 with stable accuracy). In the same sessions, the percentage of pads crossed without a false positive increased from 19.2% to 29.1% for P1 and slightly deteriorated for P2 (from 34.3% to 31.0%).

Mutual learning approach for user training

Our training approach targeted biweekly sessions and initially involved “offline”, open-loop BCI training, where our pilots performed various MI tasks without observing real-time feedback, so as to identify the optimal MI tasks and calibrate the BCI. This was followed by “online”, closed-loop BCI feedback training allowing the users to gradually optimize the modulation of their brain rhythms (4). Finally, race training allowed our end-users to familiarize with the actual BCI application demands. BCI recalibration was performed only twice per pilot (P1: 30/06/2016 and

14/09/2016, P2: 11/08/2016 and 08/09/2016). Table 2 presents the selected spatio-spectral features (bands and channels).

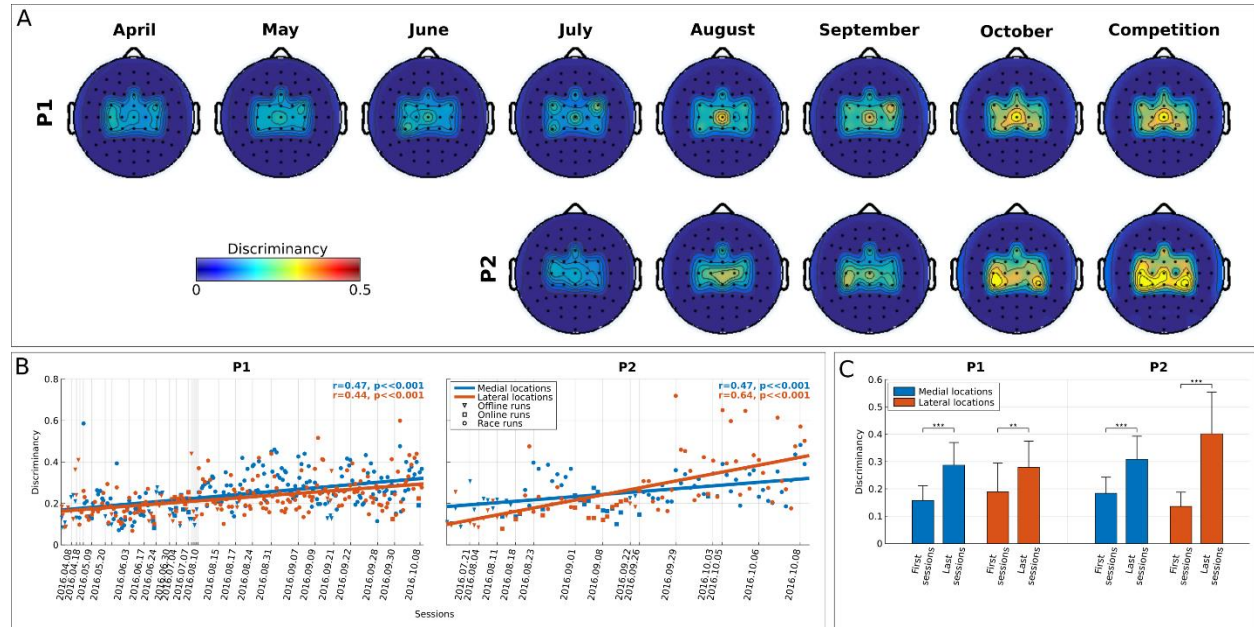


Fig. 4. BCI feature discriminancy. (A) Topographic maps of discriminancy per training month on the 16 EEG channel locations over the sensorimotor cortex monitored. Bright color indicates high discriminancy between Both Hands and Both Feet MI tasks employed by both pilots (P1 top, P2 bottom). The discriminancy of each channel is quantified as the Fisher score of the EEG signal's power spectral density distributions for these two mental classes in the high β -band (22-32 Hz) within each run. Each map illustrates local Fisher scores (with inter-channel interpolation) averaged over all runs within the supertitled month. (B) Average medial (blue, channels: FCz, Cz, CPz) and lateral (red, channels FC3, C3, CP3, FC4, C4, CP4) discriminancy for all performed offline, online and racing runs of pilots P1 and P2. The corresponding linear fits and Pearson correlation coefficients (significance tested with Student's t-distribution) are reported to indicate training effects. Vertical dashed lines indicate the training session where each runs took place. (C) Average and standard deviations of medial region (blue) and lateral region (red) discriminancy within the first and last four runs of training for pilots P1 and P2. Statistically significant differences are shown with two-sided Wilcoxon ranksum tests, (**): $p < .01$, (***) $p < .001$.

Fig.4A demonstrates that this incremental mutual learning procedure has been very effective in bringing up an emerging SMR pattern (high β -band, 22-32 Hz) for both pilots, coherent with both hands MI (lateral, electrodes FC3, C3, CP3, FC4, C4, CP4 of the 10-20 EEG system) and both feet MI (medial, electrodes FCz, Cz, CPz) locations of the sensorimotor cortex. Fig. 4B further substantiates a significant enhancement trend of these patterns' discriminancy over runs (P1, N=214: $r=0.47, p < .001$ for medial and $r=0.44, p < .001$ for lateral locations, P2, N=79: $r=0.47, p < .001$ for medial and $r=0.64, p < .001$ for lateral locations), accounting for considerable, statistically significant increase for both pilots and locations between the first and last four sessions (Fig. 1C). The overall discriminancy of our pilots' SMRs (average of medial and lateral locations) correlates well (P1: $r=-0.42, p < .001, N=162$ and P2: $r=-0.31, p < .001, N=45$) with the average pad crossing time (BCI performance). The latter naturally also correlates with the primary outcome of race time (P1: $r=0.79, p < .001, N=162$ and P2: $r=0.92, p < .001, N=45$). Hence, increased SMR modulation (discriminancy) seems to be crucial for enhanced BCI and application performances.

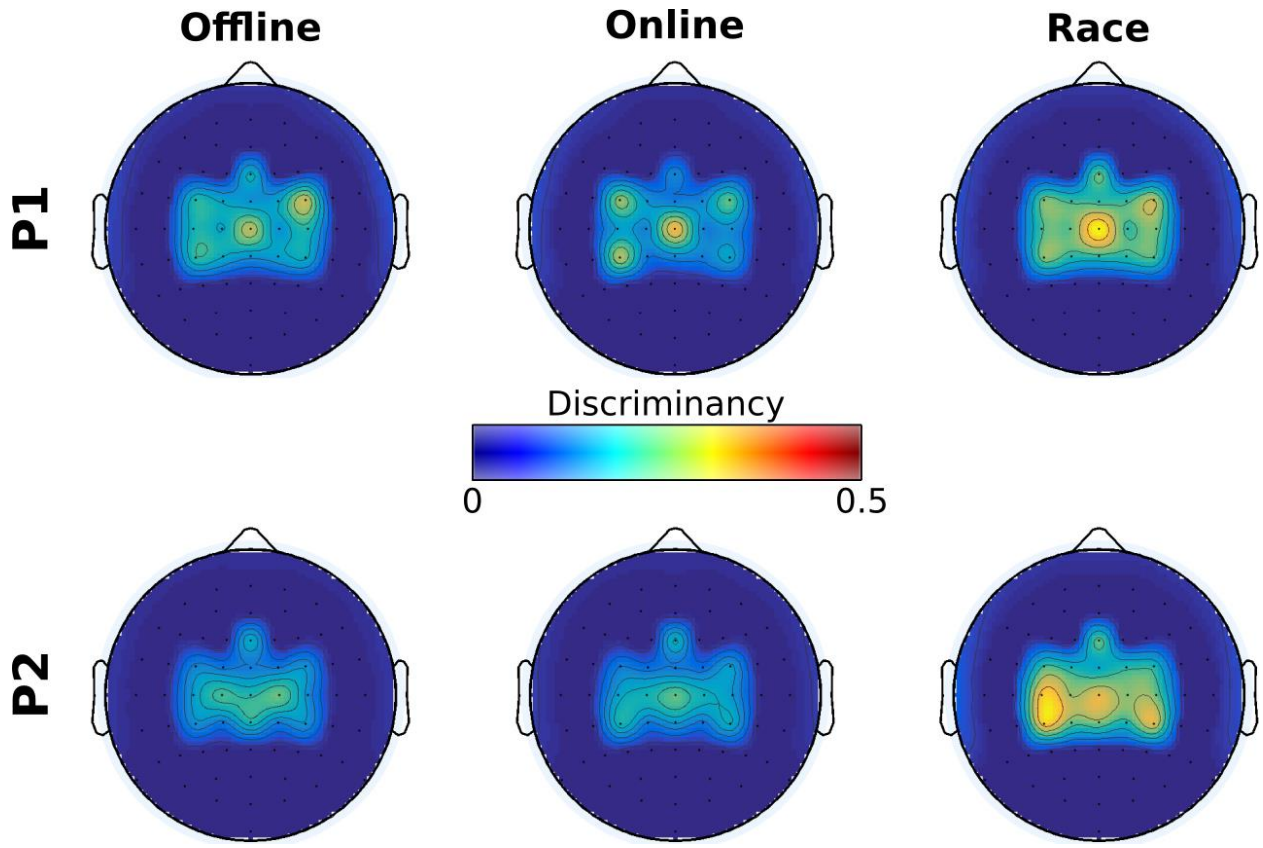


Fig. 5. BCI feature discriminancy per training modality. Topographic maps of discriminancy per training modality on the 16 EEG channel locations over the sensorimotor cortex monitored. Bright color indicates high discriminancy between Both Hands and Both Feet motor imagery tasks employed by both pilots (P1 top, P2 bottom). The discriminancy of each channel is quantified as the Fisher score of the EEG signal's power spectral density distributions for these two mental classes in the high β band (22-32 Hz), on this channel. Each map illustrates local Fisher scores (with inter-channel interpolation) averaged over all runs of the supertitled modality.

Fig. 5 reveals a profound similarity of the SMR patterns for the different training modalities performed, suggesting benefits have been derived by the incremental structure of our mutual learning protocol (17, 44). On the other hand, Fig. 6 sheds light on the neurophysiological basis of P1's poor performance in the final. It can be seen that P1's inability in this particular race to deliver any command associated to the Both Hands MI task (spin, slide) has been accompanied by the disappearance of this task's identified EEG correlates, namely the β -band SMR discriminancy in locations contralateral to the dominant right hand, selected for the classifier used in the competition (CP3, Table 2).

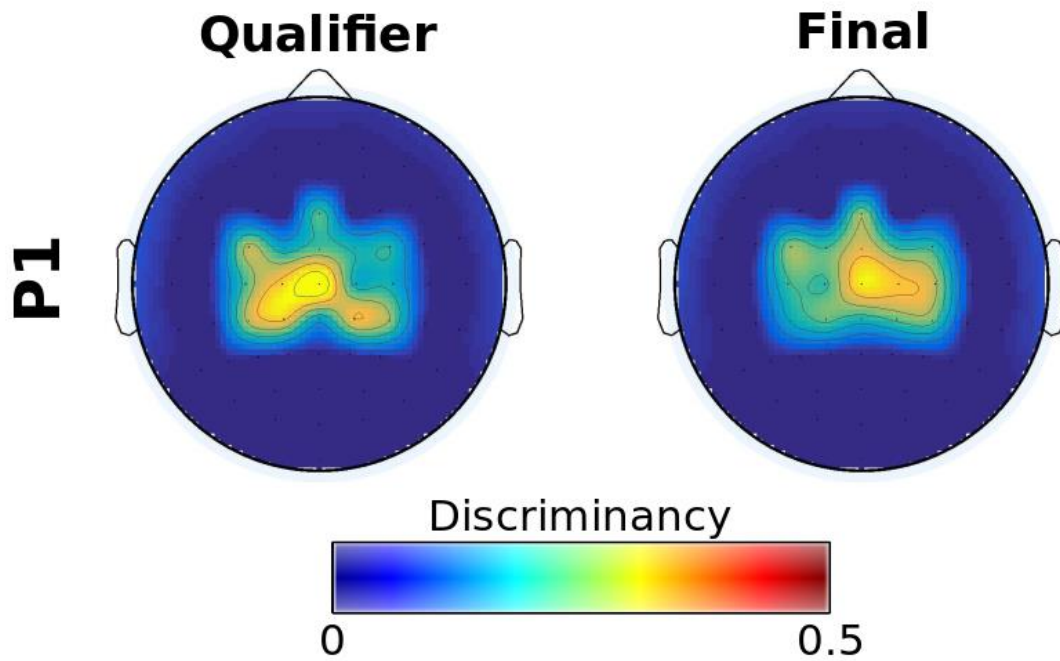


Fig. 6. BCI feature discriminancy for pilot P1 in the Cybathlon. Topographic maps of discriminancy per Cybathlon race on the 16 EEG channel locations over the sensorimotor cortex monitored. Bright color indicates high discriminancy between Both Hands and Both Feet motor imagery tasks employed by pilot P1. The discriminancy of each channel is quantified as the Fisher score of the EEG signal's power spectral density distributions for these two mental classes in the high β band (22-32 Hz), on this channel. Each map illustrates local Fisher scores (with inter-channel interpolation) in the supertitled race.

User-centered BCI design

The BCI's reconfiguration and the application control paradigm have substantially benefited from our pilot's input, following a user-centered approach in BCI design. As shown in Fig.1C, early attempts with a 3-class BCI (paradigm 1) severely compromised the total command accuracy, which is reflected in the high race completion times in this period. Supporting only two commands (paradigm 2) was clearly suboptimal. Thus, while the two separable motor imagery tasks (kinesthetic both hands and feet MI for both our subjects) were directly mapped to the spin and jump avatar actions, two different solutions were evaluated for the slide command: paradigm 3 would make the avatar slide after a configurable inactivity period. Paradigm 4, would trigger sliding when two commands of different type were forwarded within a configurable timeout. The latter protocol has been shown to be significantly superior for P1 (who executed enough races with each control paradigm) in terms of the median time spent on yellow pads (Fig. 7A) that reduced significantly ($p < .001$, two-sided Wilcoxon ranksum test) from 12.4 s (N=83) with paradigm 3 to only 5.1 s (N=363) with paradigm 4. Simultaneously, the slide command accuracy increased significantly (Fig. 7B, $p=0.0019$, two-sided Wilcoxon ranksum test) from $67.2 \pm 37.8\%$ (N=26) to $91.2 \pm 17.0\%$ (N=94). This naturally led to important reduction of the race completion time with paradigm 4 (Fig. 7C, 121.2 ± 20.1 s, N=114 against 129.5 ± 12.4 s, N=26, $p=0.0039$, two-sided Wilcoxon ranksum test), which was naturally selected for the competition.

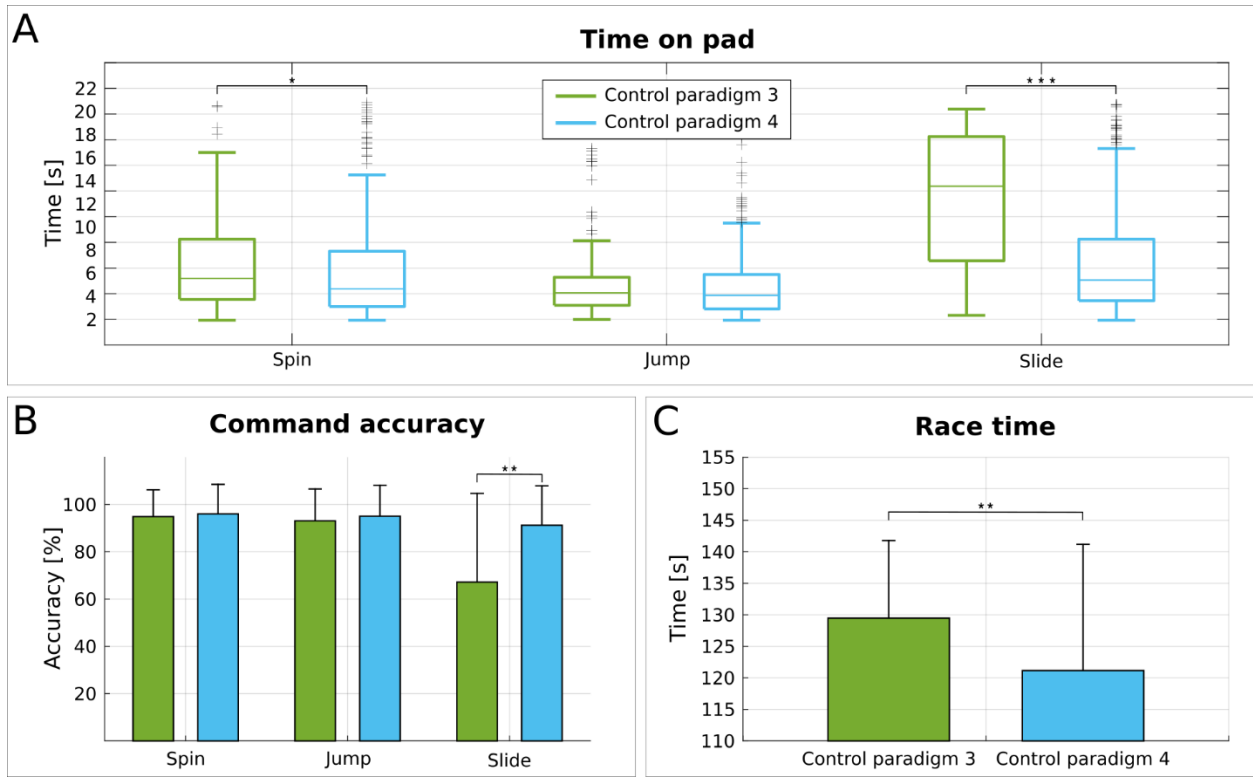


Fig. 7. Effects of the control paradigm. **(A)** Boxplots of pad crossing time (s) for pilot P1 and all pad types (spin, jump, slide) and control paradigms 3 (green) and 4 (cyan). The box edges signify the 75th (top) and 25th (bottom) percentiles and the colored horizontal line the median of the corresponding distribution. The whiskers extend to the largest and smallest non-outlier values. Outliers are marked with black crosses. **(B)** Average and standard deviation of BCI command accuracy (%) for pilot P1 and all command types (spin, jump, slide) and control paradigms 3 (green) and 4 (cyan). **(C)** Average and standard deviation of race completion time (s) for pilot P1 and control paradigms 3 (green) and 4 (blue). Statistically significant differences are shown with two-sided Wilcoxon ranksum tests, (*): $p < .05$, (**): $p < .01$, (***): $p < .001$.

Discussion

The most important conclusion drawn by the outcomes of the Cybathlon 2016 BCI race and our longitudinal training study is the profound translational potential of state-of-the-art self-paced, non-invasive BCI. The Brain Tweaklers along with another ten international teams produced competitive BCI systems driven by the technology's intended end-users despite harsh logistical constraints and conditions that go beyond realistic to being rather adverse: on-demand BCI operation twice in a day, need for swift system setup and pilot transport in hectic circumstances, as well as racing under extreme noise and psychological strain for the pilots (who were facing a crowded arena in a live televised event). For the Brain Tweaklers team, user training took place under uncontrolled environments at the pilots' homes. Besides further improvements along these lines are needed, the employed diverse hardware and software already proved to be portable, lightweight, robust, safe, unobtrusive and user-friendly enough to comply with such requirements. Furthermore, usability has not been traded-off with performance. The average pilot performances in the competition (Table 1) and especially the Brain Tweaklers' averages and records (Fig. 1B-C, Fig. 2) compare favorably with the optimal ternary input on the same task. It is noteworthy that our end-users managed to replicate their average training performances at the actual event, showing that the Cybathlon performances are sufficiently representative. These performances were achieved with decoders trained several weeks before the competition, without

any recalibration on the competition day. Hence, there is strong evidence that MI BCI is nowadays a competitive AT solution.

The Brain Tweakers produced the most effective BCI for the task at hand. We thus find it useful to share our (unavoidably, subjective and to some extent conjectural) opinion on the factors we think have been crucial to this superiority. The signal processing and artificial intelligence (AI) embedded into our system has certainly been an integral part of our Cybathlon success and has been previously validated in end-user studies (4, 29, 32). However, it can be safely assumed (and anecdotally verified for some of our competitors) that most if not all contestants have exploited for the most part equivalent state-of-the-art methodologies. Hence, we are convinced that the determining factors have rather been the user-centered design and the mutual learning approach principles pursued by our team.

Regarding the former, we claim that our pilots have substantially benefited from a system design approach putting the user in the center of the development loop, in two different aspects: the design of the control paradigm and the personalization through re-configurability of the BCI. Our team had prevailed in the rehearsal event in the summer of 2015 with a “native” 3-class BCI (three different MI tasks mapped to one game command each). Nevertheless, this straightforward design (paradigm 1) was rapidly abandoned as soon as it was clear that the recruited end-users could not easily deliver an additional mental command without extensive training, compromising not only the slide command and the total BCI accuracies (Fig. 3B, initial sessions), but also the race time (Fig. 1). Attempting to acquire a third reliable mental command through training was judged impractically time-consuming given the competition deadline. Relinquishing support of a third command (paradigm 2) could not lead to a competitive system. We thus incorporated human-computer interaction (HCI) principles, which we had previously adopted (29) exploiting timing-based features (paradigm 3) and binary interaction techniques (paradigm 4). After shaping, parameterizing and evaluating them based on our pilots’ suggestions, performance improved by almost 10 s (Fig. 7C). Notwithstanding a statistically significant but, comparatively, small reduction of the median time on cyan pads (spin commands, Fig. 7A) that should be the consequence of our mutual learning procedure (see below), Fig. 7 shows that substantial improvements in this period are derived only for the yellow pads (slide command). Thus, total race time improvement should be fundamentally attributed to the control paradigm optimization rather than the underlying learning effects. Besides this critical adjustment to our end-users’ BCI skills, the configuration options offered by our BCI’s evidence accumulation module (see “Materials and Methods, BCI implementation”) allowed for further personalization of the BCI that avoided large performance fluctuations and reduced the need for recalibration of the classifier.

As for mutual learning, our results clearly demonstrate that both pilots were able to gradually increase the discriminancy of both MI tasks employed (Fig. 4), thus facilitating their reliable recognition. It is mainly to this “subject learning” effect, coupled with a parallel “machine learning” process of occasionally recalibrating the classifier’s parameters to better capture the evolving SMR patterns (twice per subject, Table 2), that the established improvement trends of BCI command accuracy (Fig. 3) and, in turn, race completion time (Fig. 1B-C) should be mainly attributed. The reported correlations between discriminancy, BCI performance and race time further establish this mechanism. Given that around 25 s were gained overall throughout BCI training (Fig. 1B), of which approximately 10 s should be attributed to optimizing the control paradigm (Fig. 7C), it seems that longitudinal subject learning can be credited with more than 50% of the improvement. In other words, our mutual learning strategy considerably helped the pilots improve the discriminancy of the SMR patterns related to the two employed MI tasks (Fig. 4C), thus turning two originally rather “average” BCI users into champions.

A choice we view as critical towards the success of operant learning realized here is the infrequent calibration of the BCI (Table 2). The fact that the selected features (i.e., those whose activity is implicitly fed back to the user) were the ones where the discriminancy has improved through training (Fig. 4) advocates for the instrumental nature of the learning observed in our study. Related co-adaptive approaches (31, 45, 46) might assist towards greater system performance, but the extent to which continuous or frequent recalibration might harm the user's learning process should be carefully assessed (46). Although machine learning is particularly critical for identifying the initial discriminant brain patterns that subjects can naturally modulate in a stable manner, machine learning alone will unlikely solve the problem of universal BCI accessibility altogether. Indeed, this problem is typically associated with subjects unable to elicit such stable discriminant brain patterns. As our results show, see also Carmena's work with monkeys (47, 48) BCI is primarily "a skill to be learned" that mutual learning approaches greatly facilitate.

The importance of the successful and longitudinal mutual learning realized in this study needs to be further highlighted with respect to the relevant literature. Owing to the analogy to neurofeedback training (49, 50) upon which the first BCIs relied (3, 6, 7), the view that similar instrumental type of learning takes place during online BCI training has been pervasive in (especially, self-paced) BCI literature. However, this analogy is not straightforward. Unlike neurofeedback studies, BCI's are complex pattern recognition systems, feeding back some transformation of multivariate brain activity. There is also no guarantee that operant conditioning will be possible with any type of brain signal, underlying task or user category (51–53), what has been a limitation even for classical neurofeedback approaches (1). Moreover, whenever the presence of BCI learning has been claimed, it was mostly on the grounds of improved BCI classification accuracy (8, 25, 28, 32, 45) or application performances (54) only, which are indirect measures of improved brain signal modulation (improvements could be due to better calibrated decoders and re-parameterizations of the BCI and the application, among other factors). Often, MI BCI learning is claimed, but hardly substantiated (6, 7, 24). Direct evidences on the BCI feature level are in fact rare (55), derived in able-bodied populations and not longitudinal (30, 31, 45, 55, 56). Additionally, many studies argue in favor of complete avoidance of longitudinal mutual learning, investigating the possibility of machine learning able to decode any type spontaneous modulation of brain activity (26, 30). Hence, our results constitute unique evidence of the existence and efficacy of operant mutual learning in self-paced MI BCI training, as learning is significant and substantiated at all levels of the interface (brain rhythm modulation, command accuracy, application primary outcome), on both chronic SCI pilots. Importantly, these learning effects were achieved under uncontrolled circumstances at the subjects' home with minimal expert personnel intervention, while the learned outcome was replicated at a demanding international competition under adverse circumstances.

Having claimed that this work demonstrates the translational impact of self-paced, non-invasive BCI, it is necessary to justify the extent to which the Cybathlon BCI race application has served as a suitable and adequate testbed in this respect. Although Brain Runners might appear a simple video game (a choice motivated by the need to appeal to a live audience and adhere to a competitive spirit), it has been explicitly designed to assess all critical skills required for real BCI applications, being in fact rather demanding in this regard. Concerning IC skills, it required control over three mental commands, where literature has shown that one or two are enough for most applications (4, 6, 11, 15, 23, 25–30, 32, 33), while sufficiently rewarding both command accuracy and delivery speed. Most importantly, it penalized considerably the lack of INC skills, an essential feature of self-paced interaction, yet, so far rather neglected (4, 23, 33, 43).

The present study suffers certain limitations, the main one being reporting on only two individuals. Still, the fact that both participants exhibited the same training effects and

comparable performances makes us confident that our conclusions should generalize, at least to populations with similar medical profile. Other limitations, related to logistic constraints of preparing for a competition, include the inability to compare against other BCI paradigms, neural interfaces or algorithms. The main shortcoming of our BCI system is unsatisfactory robustness. Lack of robustness is a well-known issue of all BCI paradigms and has been associated to the non-stationarity of brain signals. As shown, although P1 showcased better average performance, he also exhibited higher variability than P2. This effect, also reflected in our pilot's competition outcomes where P1 set an impressive record but was unable to defend it a few hours later, suggests that stability (robustness) is at least as crucial as performance (effectiveness) for optimal BCI control. We have shown that loss of control for P1 in the final was the result of the disappearance of the SMR modulations normally induced by his Both Hands MI (Fig. 7). Various psychological factors (such as motivation, attention, stress) have been implicated in these negative effects (1, 17, 44) which unfortunately are quite frequent in MI BCI operation. Yet, the underlying causes remain largely a mystery. Our future work on translational BCI applications will focus on further end-user evaluation in larger populations and the investigation of ways to increase robustness, a field where we have already recently contributed (46).

Materials and Methods

Study design

The objective of this study was to train two end-users with severe motor impairments to optimally control the Brain Runners BCI application and participate in the Cybathlon BCI race. Towards this goal the Brain Tweakers have applied the ensemble of BCI methods, control paradigms and training protocols developed in CNBI. The competition and logistical constraints have imposed the nature of this study as an uncontrolled, longitudinal, observational two-case study.

Our inclusion criteria necessarily coincided with those of the Cybathlon BCI race: minimum age of 18, sufficient cognitive and communication abilities to understand the discipline's rules and tetraplegia or tetraparesia as a result of SCI, ALS or other lesion, quantified with a score above (including) "C" in the American Spinal Injury Association (ASIA) impairment scale. The exclusion criteria consisted of cardiac pacemakers, cyber-sickness and epilepsy. All EEG and race time data collected have been included into our statistical analysis and no outliers have been defined.

The race completion time is naturally the study's primary outcome. Each individual training or competitive race is an evaluation end-point. We additionally define a number of secondary outcomes which correlate with the primary outcome, but are characteristic of specific aspects of a pilot's performance: the time spent on each pad type (crossing time), the BCI command delivery accuracy, as well as the SMR brain pattern discriminancy.

Pilots

Both our pilots, 48-year-old P1 injured in December 1989 and 30-year-old P2 injured in May 2003, have sustained complete lesions at level C5-C6 and have scored "A" (Complete injury-No motor or sensory function is preserved in the sacral segments S4 or S5) in the ASIA scale. Both end-users were under medication for the treatment of spasms and other symptoms related to their medical condition. Certified confirmation by their medical doctor of safety to participate in the Cybathlon event was requested and signed for both pilots and insurance against accidents and injuries was taken, as per Cybathlon's regulations. A safety and eligibility check was also conducted by the organizers the day before the competition.

Both participants maintain no control of the lower limbs and only limited of the upper limbs. They are both able to stabilize their neck and head, but only P2 can also stabilize his trunk. None

of our pilots carries pacemakers or other implants, suffers epilepsy, cyber-sickness or needs respiratory assistance. They both use other advanced AT in their daily lives, like driving aids and speech-to-text software. P1 had previously participated in the MI BCI studies reported in (4, 29, 32), while P2 was BCI naive at the onset of his Cybathlon training. Informed consents have been signed in accordance with the declaration of Helsinki and their participation in the training sessions as well as in the competition has been approved by the Swiss committees for ethics in human research.

Cybathlon BCI race

The Cybathlon competition comprised six different disciplines, each concerning a different type of AT (functional electrical stimulation, powered arm and leg prostheses, exoskeletons, wheelchairs and BCI). The BCI race consisted of (up to) four brain-controlled avatars, each actuated by a disabled pilot by means of mental commands, so as to reach the finishing line of a virtual race game called “Brain Runners” ahead of its opponents. The BCI pilot should be able to forward three mental commands to his/her avatar (spin, jump over prickles, slide under electrical rays), each of which would accelerate it only when issued while the avatar was traversing the corresponding color-coded track segment called “pad” (spin on cyan, jump on magenta and slide on yellow pads). The acceleration effect would last until the avatar reached the beginning of the next pad, or upon reception of a following erroneous command overriding the user’s correct command (whichever happened first). In addition to these three “action” pads, a fourth type (white pads) required “idling” to avoid any command delivery. A misplaced command, including false positives on the white pads, would slow down the pilot’s course towards the finish line of the track for 4 s (this timer would reset if another erroneous command or false positive is received in the meantime), until the beginning of the next pad or a following correct command overriding the erroneous one (whichever happened first). Besides the accelerating/slowing down behavior of the avatar, a thunder of the corresponding color briefly appearing over the avatar’s head would inform the pilot of the command currently sent. Support of at least one mental command was required to participate in the competition.

The standard track was composed of sixteen pads (four of each type) randomly arranged, so that the order of pads was not known to the competitors beforehand and was different for every race. The starting and finishing lines were situated on two additional white pads, so that the total distance to be covered by the pilots’ avatars was 500 virtual meters. The lower bound of race completion time on this track (i.e., the one achieved with an ideal keyboard input) is 54 s. The corresponding upper bound (continuous erroneous delivery) is 327 s, although only times below 240 s were considered valid in the actual competition. Since the avatars would proceed by default forward at a low “base” speed, the race completion time in case of no input whatsoever would be 162 s. The equivalent minimum, no-response and maximum crossing times for the action pads were 2 s, 11 s and 19 s, respectively. Hence, 11-19 s is the time frame within which a user is required to forward a correct command, with delivery speed being equally important to command accuracy. The minimum and maximum crossing times for the white pads were 5.5 s and 19 s, respectively. The corresponding times for the starting white pad were 5 s and 13 s, while for the ending white pad 3 s and 10 s.

The exclusion criteria for the technology provider dictated the use of non-invasive interfaces, while visual, tactile or any kind of BCI feedback other than the one provided directly by the Brain Runners graphical user interface was prohibited, effectively excluding synchronous, stimuli-driven BCI paradigms like P300 and SSVEP. Besides the Brain Tweakers, another twelve franchises of some of the most prestigious BCI research groups worldwide have contested this unique trophy. Two teams withdrew and one participated “out of competition” due to pilot ineligibility. The BCI race tournament format involved, initially, four qualification races

(morning). The pilots marking the best-four race completion times would qualify to Final A (afternoon) and compete for one of the three medals (gold, silver, bronze), while the second-best-four times would compete for places 5-8 in Final B. The event took place in a crowded, sold-out arena in front of a loud audience of roughly 4600 spectators. A mock-up “rehearsal” event was held in July 2015 to ensure the best possible preparation for both the teams and the organizers.

Training modalities, periods and locations

Our mutual (user and BCI algorithm) learning approach involved three different training modalities aiming to establish, on the one hand, the end-users’ best possible control over spontaneous modulation of their SMRs by means of MI tasks and, on the other hand, their fast and accurate recognition on the part of the trained MI BCI algorithm. MI is defined as the mental rehearsal of a movement without overt motor output (57). For MI tasks related to completely paralyzed limbs (legs for both P1 and P2) our pilots were instructed to attempt the corresponding movement, otherwise imagination suppressing any overt motor act was requested. During the competition, violations of the latter prerequisite were controlled by the judges.

Initially, “offline” training (MI without real-time feedback) was applied, in order to exploit and calibrate the BCI on spontaneous SMR modulations the subjects could already elicit for the tested MI tasks. In this phase, we have mainly explored the existence of distinct brain patterns corresponding to right hand, left hand, both hands, both feet MI and rest. Subsequently, offline runs were limited to both hands and both feet MI, which both our pilots were found to optimally modulate, so as to collect “clean” data for updating the BCI algorithm’s parameters. P1 has also unsuccessfully tried imagination of tongue movement, as well as “word” and “mathematical association” mental tasks.

“Online” sessions followed, where the pilots proceed with real-time BCI control of a conventional, continuous visual feedback cursor targeting the enhancement of the patterns’ discriminancy in an operant conditioning fashion, while the BCI parameters were later re-calibrated to better reflect the evolving brain patterns with the derived EEG data. Online runs were mainly conducted using the discriminant (coincidentally, for both our subjects) both hands and both feet MI tasks (2-class). P1 attempted to operate a 3-class online modality (left hand, right hand, feet MI) for a few runs. More details on the visual interface of these two modalities and exactly how the BCI feedback cursor is driven by the BCI algorithm, can be found in Appendix A of (4).

The third and latest stage was dominated by actual racing with the training version of the Brain Runners game delivered to the contestants, so that our pilots could get accustomed to the real application’s demands, where one had to rely solely on the discrete feedback embedded into the cluttered graphics of the game itself. Offline, online and racing runs were often interleaved (Table S1) in order to make these transitions smoother. For the first racing runs only, we allowed our pilots to also observe the visual BCI feedback. During race training, our pilots would generally compete against the “bot” avatar option provided by the game. The “skill” of this bot competitor would be gradually increased to increasingly challenge our pilots. The racing track was randomized for each race, simulating the actual Cybathlon conditions.

Prior to (and including) the competition day, P1 received 35 training sessions within the period April-October 2016, while P2 underwent 16 sessions within July-October 2016, both in an individualised and flexible (approximately) bi-weekly schedule, which was intensified as the competition day was approaching. AN15VE executed in total 40 offline, 12 online and 182 race runs, while P2 did 15, 19 and 57 runs, respectively (Table S1). All training sessions took place at the pilots’ homes under the supervision of one or two BCI engineers, except for two distinct sessions accommodated in the laboratory, where our two pilots competed against each other in

the presence of a crowd of spectators, so as to simulate and get used to the special conditions they would cope with on the competition day.

BCI implementation

The Brain Tweaksters participation in the Cybathlon BCI race relied on the EEG-based MI BCI design previously developed in CNBI, which had already been shown to allow end-users to successfully operate a number of BCI prototypes (4). For both user training and competitive racing, EEG was acquired with a lightweight 16-channel g.USBamp amplifier (g.Tec medical engineering, Schiedelberg, Austria). The experimental setup during training additionally consisted of a laptop running the BCI algorithms and another one running the Brain Runners game. In the actual competition, the latter was substituted by the competition's dedicated monitor displaying the race from each pilot's individual viewpoint.

The EEG signal was recorded at 512 Hz sampling rate, band-passed filtered within 0.1 and 100 Hz and notch filtered at 50 Hz. The monitored EEG channels were selected so as to adequately cover the sensorimotor cortex (Fig. S1A). The signal was spatially filtered with a Laplacian derivation and the power spectral density (Welch periodogram) of each channel was computed with 2 Hz resolution in 1 s-long windows sliding every 62.5 ms. Feature selection was performed by ranking the candidate spatio-spectral features according to discriminant power, calculated through canonical variate analysis, eventually manually selecting the most discriminant and neurophysiologically relevant ones. A Gaussian classifier outputting a probability distribution over two MI tasks was used to classify the consecutive feature vectors in real time. The Gaussian classifier was trained with a gradient-descent supervised learning approach using the labeled MI datasets resulting from the aforementioned training protocols. The samples with "uncertain" probability distributions (where the maximum probability does not exceed a certain threshold) were rejected, while the remaining ones were fed to an evidence accumulation module smoothing the classifier output by means of a leaky integrator (exponential smoothing). A final decision is emitted by the BCI system once the pilot is able to push the integrated probabilities of some mental class to reach a configurable decision threshold by consistently performing the corresponding MI task, thus forwarding the associated command to his Brain Runners avatar. Upon delivery of a BCI command, the integrated probabilities are reset to the uniform distribution so as to start an unbiased new trial. A refractory period of 1 s was set in between consecutive commands. An artifact rejection scheme would block the BCI output once ocular and facial muscle artifacts were detected. A more detailed description of all the above methods is provided in Appendix A of (4) and the references therein.

Artifact rejection scheme

Under the Cybathlon BCI race regulations, all teams should embed an artifact removal or rejection framework into their BCI system, ensuring that the their pilot's avatar is actuated by means of brain signals only, without interference from other signals originating from muscle activity or at the level of the peripheral nervous system. Thus, the Brain Tweaksters artifact control scheme targeted the detection of electrooculogram (EOG) and facial electromyogram (EMG) signals, upon which the BCI output was blocked for a configurable interval preventing any outgoing command towards the pilot's BrainRunners avatar. Respecting the need for a minimally obtrusive setup, only 4 electrode/sensor pairs are employed to extract two bipolar EOG channels, by means of a second synced g.USBamp device. One sensor is placed on either eye canthus, a third one on the pilot'sinion bone, while the last sensor acts as the reference and is placed on the pilot's forehead (Fig. S1B). In sync with the EEG acquisition, EOG signals are acquired at 512 Hz in frames of 62.5 ms. Artifact detection is performed separately on each consecutive frame, resulting in very fluid and responsive detection of artifact onset and offset.

For each frame, the original channels EOG_i , $i \in [1, 4]$, are combined to form a horizontal, $EOG_h = EOG_1 - EOG_3$, and a vertical $EOG_v = EOG_2 - (EOG_1 + EOG_3)/2$ channel, specializing in capturing horizontal and vertical eye movements, respectively. The average of all channels was also extracted and monitored as it is particularly sensitive to eye blinks and intense facial muscle flexions. All channels were band-pass filtered between 1 and 10 Hz with a second-order Butterworth filter and rectified. Finally, the processed channel frames are compared against a common configurable threshold. The individual frame decision was 1 when any of the processed samples within the current frame exceeds the threshold and 0 otherwise. The final artifact detection module would communicate an artifact onset event to the game controller upon a frame decision transition from 0 to 1, signaling the blocking of the BCI output. An artifact offset event lifting the BCI command blocking was issued after a configurable timeout since the latest artifact onset detection.

Game control paradigm

The game control paradigm defines the way the pilot's motor imaginations translate into avatar actions through the emitted BCI commands. Several control paradigms have been designed and tested throughout the training period in close cooperation between the Brain Tweakers researchers and pilots. Initially we explored the straightforward option of a 3-class BCI (paradigm 1) employing right hand, left hand and both feet MI. Thereby, each BCI command was directly mapped to a certain avatar action (right hand MI to spin, both feet MI to jump and left hand MI to slide). A 2-class BCI (paradigm 2) preserving the previous mapping but leaving the slide command unsupported was also tested. Given the unsatisfactory outcome of these two approaches, another two paradigms were designed, both investigating well-known human-computer interaction principles for supporting all three avatar commands given only a binary input. Specifically, the two separable MI tasks (both hands and feet MI for both our subjects) were again directly mapped to the spin and jump avatar actions. Additionally, paradigm 3 would make the avatar slide after a configurable period of INC. Paradigm 4, on the other hand, would trigger sliding when two consecutive commands of different type (i.e., a spin/jump or jump/spin pair) were forwarded within a configurable interval. Paradigm 4 was adopted for the competition.

In all four tested control paradigms, "idling" is achieved through the "resting" mental task, where the subject is deliberately not engaging in any MI task (INC). Since the BCI classifier is continuously (every 62.5 ms, i.e., at 16 Hz) outputting a probability distribution over the MI mental classes (not including the resting state), INC is achieved through a statistical approach, where, thanks to the evidence accumulation module and the BCI's optimized parametrization (decision and sample rejection thresholds, smoothing parameter) a BCI command is only forwarded when the subject is consistently performing the associated MI. Otherwise the integrated probabilities will tend to fluctuate below the decision thresholds avoiding any command forwarding (4).

Evaluation metrics and data

The race completion and the pad crossing times are measured in seconds (s). BCI performance is quantified through BCI command accuracy, which is the percentage of pads in a race where the correct command has been delivered within the given time frame. Pad crossing times are reported to simultaneously evaluate BCI command accuracy and delivery speed. The total BCI command accuracy in a race is computed as the average per-command accuracies (class-specific true positive rates). For the white pads, an equivalent accuracy metric (true negative rate) is calculated as the percentage of white pads in the race that the pilot managed to cross without delivering any command. Finally, discriminancy of a given spatio-spectral EEG feature

(corresponding to a certain EEG channel and a frequency band) for two mental classes is quantified through Fisher score as $FS = |\mu_1 - \mu_2| / \sqrt{s_1^2 + s_2^2}$, where μ_1, μ_2 are the means and s_1, s_2 are the standard deviations of this feature's sample values for mental class 1 and 2, respectively. Discriminancy over larger topographic or spectral regions is computed as the average Fisher score of all features corresponding to the channels and frequency bands of the regions in question. Table S1 presents the list of sessions executed and the type of data acquired.

Statistical analysis

Point estimates are reported using averages or medians and dispersions as standard deviation or 25th and 75th percentiles, when the underlying distribution is normal or skewed, respectively. Training effects are shown by reporting Pearson correlation coefficients and their significance at the 95% confidence interval through Student's t-distribution. Additionally, the first and last four sessions are compared and tested for significant differences at the 95% confidence interval using unpaired, two-sided Wilcoxon non-parametric rank-sum tests.

Supplementary Materials

Fig. S1. Electrode configurations

Table S1. Training session information.

Movie S1. Typical race training session of pilot P1.

References and Notes:

1. U. Chaudhary, N. Birbaumer, A. Ramos-Murguialday, Brain-computer interfaces for communication and rehabilitation, *Nat. Rev. Neurol.* **12**, 513–525 (2016).
2. C. Brunner, N. Birbaumer, B. Blankertz, C. Guger, A. Kübler, D. Mattia, J. d. R. Millán, F. Miralles, A. Nijholt, E. Opisso, N. Ramsey, P. Salomon, G. R. Müller-Putz, BNCI Horizon 2020: towards a roadmap for the BCI community, *Brain-Comput. Interfaces* **2**, 1–10 (2015).
3. N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, H. Flor, A spelling device for the paralysed, *Nature* **398**, 297–298 (1999).
4. R. Leeb, S. Perdikis, L. Tonin, A. Biasucci, M. Tavella, M. Creatura, A. Molina, A. Al-Khodairy, T. Carlson, J. d. R. Millán, Transferring brain-computer interfaces beyond the laboratory: Successful application control for motor-disabled users, *Artif. Intell. Med.* **59**, 121–132 (2013).
5. P. R. Kennedy, R. A. E. Bakay, M. M. Moore, K. Adams, J. Goldwaithe, Direct control of a computer from the human central nervous system, *IEEE Trans. Rehabil. Eng.* **8**, 198–202 (2000).
6. G. Pfurtscheller, G. R. Müller, J. Pfurtscheller, H. J. Gerner, R. Rupp, “Thought” – control of functional electrical stimulation to restore hand grasp in a patient with tetraplegia, *Neurosci. Lett.* **351**, 33–36 (2003).
7. J. R. Wolpaw, D. J. McFarland, Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans, *Proc. Natl. Acad. Sci.* **101**, 17849–17854 (2004).
8. A. Kübler, F. Nijboer, J. Mellinger, T. M. Vaughan, H. Pawelzik, G. Schalk, D. J. McFarland, N. Birbaumer, J. R. Wolpaw, Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface, *Neurology* **64**, 1775–1777 (2005).
9. L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, J. P. Donoghue, Neuronal ensemble control of prosthetic devices by a human with tetraplegia, *Nature* **442**, 164–171 (2006).
10. E. W. Sellers, T. M. Vaughan, J. R. Wolpaw, A brain-computer interface for long-term independent home use, *Amyotroph. Lateral Scler.* **11**, 449–455 (2010).
11. G. Onose, C. Grozea, A. Angheliescu, C. Daia, C. J. Sinescu, A. V. Ciurea, T. Spircu, A. Mirea, I. Andone, A. Spânu, C. Popescu, A. S. Mihăescu, S. Fazli, M. Danóczy, F. Popescu, On the feasibility of using motor imagery EEG-based brain-computer interface in chronic tetraplegics for assistive robotic arm control: a clinical test and long-term post-trial follow-up, *Spinal Cord* **50**, 599–608 (2012).
12. J. L. Collinger, B. Wodlinger, J. E. Downey, W. Wang, E. C. Tyler-Kabara, D. J. Weber, A. J. McMorland, M. Velliste, M. L. Boninger, A. B. Schwartz, High-performance neuroprosthetic control by an individual with tetraplegia, *The Lancet* **381**, 557–564 (2013).
13. E. W. Sellers, D. B. Ryan, C. K. Hauser, Noninvasive brain-computer interface enables communication after brainstem stroke, *Sci. Transl. Med.* **6**, 257re7–257re7 (2014).

14. U. Chaudhary, B. Xia, S. Silvoni, L. G. Cohen, N. Birbaumer, Brain–Computer Interface–Based Communication in the Completely Locked-In State, *PLoS Biol.* **15**, e1002593 (2017).
15. E. M. Holz, L. Botrel, T. Kaufmann, A. Kübler, Long-Term Independent Brain-Computer Interface Home Use Improves Quality of Life of a Patient in the Locked-In State: A Case Study, *Arch. Phys. Med. Rehabil.* **96**, S16–S26 (2015).
16. M. J. Vansteensel, E. G. M. Pels, M. G. Bleichner, M. P. Branco, T. Denison, Z. V. Freudenburg, P. Gosselaar, S. Leinders, T. H. Ottens, M. A. Van Den Boom, P. C. Van Rijen, E. J. Aarnoutse, N. F. Ramsey, Fully Implanted Brain–Computer Interface in a Locked-In Patient with ALS, *N. Engl. J. Med.* **375**, 2060–2066 (2016).
17. R. Chavarriaga, M. Fried-Oken, S. Kleih, F. Lotte, R. Scherer, Heading for new shores! Overcoming pitfalls in BCI design, *Brain-Comput. Interfaces*, 1–14 (2016).
18. F. Piccione, F. Giorgi, P. Tonin, K. Priftis, S. Giove, S. Silvoni, G. Palmas, F. Beverina, P300-based brain computer interface: Reliability and performance in healthy and paralysed participants, *Clin. Neurophysiol.* **117**, 531–537 (2006).
19. E. W. Sellers, E. Donchin, A P300-based brain–computer interface: Initial tests by ALS patients, *Clin. Neurophysiol.* **117**, 538–548 (2006).
20. A. Kübler, A. Furdea, S. Halder, E. M. Hammer, F. Nijboer, B. Kotchoubey, A Brain-Computer Interface Controlled Auditory Event-Related Potential (P300) Spelling System for Locked-In Patients, *Ann. N. Y. Acad. Sci.* **1157**, 90–100 (2009).
21. A. Combaz, C. Chatelle, A. Robben, G. Vanhoof, A. Goeleven, V. Thijs, M. M. Van Hulle, S. Laureys, A Comparison of Two Spelling Brain-Computer Interfaces Based on Visual P3 and SSVEP in Locked-In Syndrome, *PLoS ONE* **8**, e73691 (2013).
22. D. Lesenfants, D. Habbal, Z. Lugo, M. Lebeau, P. Horki, E. Amico, C. Pokorny, F. Gómez, A. Soddu, G. Müller-Putz, S. Laureys, Q. Noirhomme, An independent SSVEP-based brain–computer interface in locked-in syndrome, *J. Neural Eng.* **11**, 035002 (2014).
23. G. E. Birch, Z. Bozorgzadeh, S. G. Mason, Initial on-line evaluations of the LF-ASD brain-computer interface with able-bodied and spinal-cord subjects using imagined voluntary motor potentials, *IEEE Trans. Neural Syst. Rehabil. Eng.* **10**, 219–224 (2002).
24. J. d. R. Millán, F. Renkens, J. Mouriño, W. Gerstner, Brain-actuated interaction, *Artif. Intell.* **159**, 241–259 (2004).
25. G. R. Müller-Putz, R. Scherer, G. Pfurtscheller, R. Rupp, EEG-based neuroprosthesis control: A step towards clinical practice, *Neurosci. Lett.* **382**, 169–174 (2005).
26. O. Bai, P. Lin, D. Huang, D. Y. Fei, M. K. Floeter, Towards a user-friendly brain–computer interface: Initial tests in ALS and PLS patients, *Clin. Neurophysiol.* **121**, 1293–1303 (2010).
27. R. Leeb, D. Friedman, G. R. Müller-Putz, R. Scherer, M. Slater, G. Pfurtscheller, Self-Paced (Asynchronous) BCI Control of a Wheelchair in Virtual Environments: A Case Study with a Tetraplegic, *Comput. Intell. Neurosci.* **2007**, 1–8 (2007).
28. E. M. Holz, J. Höhne, P. Staiger-Sälzer, M. Tangermann, A. Kübler, Brain–computer interface controlled gaming: Evaluation of usability by severely motor restricted end-users, *Artif. Intell. Med.* **59**, 111–120 (2013).
29. S. Perdikis, R. Leeb, J. Williamson, A. Ramsay, M. Tavella, L. Desideri, E.-J. Hoogerwerf, A. Al-Khodairy, R. Murray-Smith, J. d. R. Millán, Clinical evaluation of BrainTree, a motor imagery hybrid BCI speller, *J. Neural Eng.* **11**, 036003 (2014).
30. J. Höhne, E. Holz, P. Staiger-Sälzer, K.-R. Müller, A. Kübler, M. Tangermann, Motor Imagery for Severely Motor-Impaired Patients: Evidence for Brain-Computer Interfacing as Superior Control Solution, *PLoS ONE* **9**, e104854 (2014).
31. J. Faller, R. Scherer, U. Costa, E. Opisso, J. Medina, G. R. Müller-Putz, A Co-Adaptive Brain-Computer Interface for End Users with Severe Motor Impairment, *PLoS ONE* **9**, e101168 (2014).
32. R. Leeb, L. Tonin, M. Rohm, L. Desideri, T. Carlson, J. d. R. Millán, Towards Independence: A BCI Telepresence Robot for People With Severe Motor Disabilities, *Proc. IEEE* **103**, 969–982 (2015).
33. R. Rupp, M. Rohm, M. Schneiders, A. Kreiling, G. R. Müller-Putz, Functional Rehabilitation of the Paralyzed Upper Extremity After Spinal Cord Injury by Noninvasive Hybrid Neuroprostheses, *Proc. IEEE* **103**, 954–968 (2015).
34. L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel, S. Haddadin, J. Liu, S. S. Cash, P. van der Smagt, J. P. Donoghue, Reach and grasp by people with tetraplegia using a neurally controlled robotic arm, *Nature* **485**, 372–375 (2012).
35. D. P. McMullen, G. Hotson, K. D. Katyal, B. A. Wester, M. S. Fifer, T. G. McGee, A. Harris, M. S. Johannes, R. J. Vogelstein, A. D. Ravitz, W. S. Anderson, N. V. Thakor, N. E. Crone, Demonstration of a Semi-Autonomous Hybrid Brain–Machine Interface Using Human Intracranial EEG, Eye Tracking, and Computer Vision to Control a Robotic Upper Limb Prosthetic, *IEEE Trans. Neural Syst. Rehabil. Eng.* **22**, 784–796 (2014).
36. M. S. Fifer, G. Hotson, B. A. Wester, D. P. McMullen, Y. Wang, M. S. Johannes, K. D. Katyal, J. B. Helder, M. P. Para, R. J. Vogelstein, W. S. Anderson, N. V. Thakor, N. E. Crone, Simultaneous Neural Control of Simple

- Reaching and Grasping With the Modular Prosthetic Limb Using Intracranial EEG, *IEEE Trans. Neural Syst. Rehabil. Eng.* **22**, 695–705 (2014).
37. V. Gilja, C. Pandarinath, C. H. Blabe, P. Nuyujukian, J. D. Simeral, A. A. Sarma, B. L. Sorice, J. A. Perge, B. Jarosiewicz, L. R. Hochberg, K. V. Shenoy, J. M. Henderson, Clinical translation of a high-performance neural prosthesis, *Nat. Med.* **21**, 1142–1145 (2015).
 38. B. Jarosiewicz, A. A. Sarma, D. Bacher, N. Y. Masse, J. D. Simeral, B. Sorice, E. M. Oakley, C. Blabe, C. Pandarinath, V. Gilja, S. S. Cash, E. N. Eskandar, G. Friehs, J. M. Henderson, K. V. Shenoy, J. P. Donoghue, L. R. Hochberg, Virtual typing by people with tetraplegia using a self-calibrating intracortical brain-computer interface, *Sci. Transl. Med.* **7**, 313ra179–313ra179 (2015).
 39. T. Aflalo, S. Kellis, C. Klaes, B. Lee, Y. Shi, K. Pejisa, K. Shanfield, S. Hayes-Jackson, M. Aisen, C. Heck, C. Liu, R. A. Andersen, Decoding motor imagery from the posterior parietal cortex of a tetraplegic human, *Science* **348**, 906–910 (2015).
 40. E. C. Leuthardt, G. Schalk, J. R. Wolpaw, J. G. Ojemann, D. W. Moran, A brain–computer interface using electrocorticographic signals in humans, *J. Neural Eng.* **1**, 63–71 (2004).
 41. W. Wang, J. L. Collinger, A. D. Degenhart, E. C. Tyler-Kabara, A. B. Schwartz, D. W. Moran, D. J. Weber, B. Wodlinger, R. K. Vinjamuri, R. C. Ashmore, J. W. Kelly, M. L. Boninger, An Electrocorticographic Brain Interface in an Individual with Tetraplegia, *PLoS ONE* **8**, e55344 (2013).
 42. R. Riener, The Cybathlon promotes the development of assistive technology for people with physical disabilities, *J. NeuroEngineering Rehabil.* **13** (2016) (available at <http://jneuroengrehab.biomedcentral.com/articles/10.1186/s12984-016-0157-2>).
 43. M. Tavella, R. Leeb, R. Rupp, J. d. R. Millán, in *IEEE Engineering in Medicine and Biology Society (EMBC)*, (IEEE, Buenos Aires, Argentina, 2010), pp. 126–129.
 44. F. Lotte, F. Larrue, C. Mühl, Flaws in current human training protocols for spontaneous Brain-Computer Interfaces: lessons learned from instructional design, *Front. Hum. Neurosci.* **7** (2013).
 45. C. Vidaurre, C. Sannelli, K.-R. Müller, B. Blankertz, Machine-Learning-Based Coadaptive Calibration for Brain-Computer Interfaces, *Neural Comput.* **23**, 791–816 (2011).
 46. S. Perdikis, R. Leeb, J. d. R. Millán, Context-aware adaptive spelling in motor imagery BCI, *J. Neural Eng.* **13**, 036018 (2016).
 47. K. Ganguly, J. M. Carmena, Emergence of a Stable Cortical Map for Neuroprosthetic Control, *PLoS Biol.* **7**, e1000153 (2009).
 48. A. L. Orsborn, H. G. Moorman, S. A. Overduin, M. M. Shanechi, D. F. Dimitrov, J. M. Carmena, Closed-Loop Decoder Adaptation Shapes Neural Plasticity for Skillful Neuroprosthetic Control, *Neuron* **82**, 1380–1393 (2014).
 49. N. Birbaumer, Breaking the silence: Brain-computer interfaces (BCI) for communication and motor control, *Psychophysiology* **43**, 517–532 (2006).
 50. N. Birbaumer, S. Ruiz, R. Sitaram, Learned regulation of brain metabolism, *Trends Cogn. Sci.* **17**, 295–302 (2013).
 51. A. Kübler, N. Birbaumer, Brain–computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients?, *Clin. Neurophysiol.* **119**, 2658–2666 (2008).
 52. G. Pfurtscheller, P. Linortner, R. Winkler, G. Korisek, G. Müller-Putz, Discrimination of Motor Imagery-Induced EEG Patterns in Patients with Complete Spinal Cord Injury, *Comput. Intell. Neurosci.* **2009**, 1–6 (2009).
 53. K. Gourab, B. D. Schmit, Changes in movement-related β -band EEG signals in human spinal cord injury, *Clin. Neurophysiol.* **121**, 2017–2023 (2010).
 54. J. Meng, S. Zhang, A. Bekyo, J. Olsoe, B. Baxter, B. He, Noninvasive Electroencephalogram Based Control of a Robotic Arm for Reach and Grasp Tasks, *Sci. Rep.* **6** (2016).
 55. J. D. Wander, T. Blakely, K. J. Miller, K. E. Weaver, L. A. Johnson, J. D. Olson, E. E. Fetz, R. P. N. Rao, J. G. Ojemann, Distributed cortical adaptation during learning of a brain-computer interface task, *Proc. Natl. Acad. Sci.* **110**, 10818–10823 (2013).
 56. J. Faller, C. Vidaurre, T. Solis-Escalante, C. Neuper, R. Scherer, Autocalibration and Recurrent Adaptation: Towards a Plug and Play Online ERD-BCI, *IEEE Trans. Neural Syst. Rehabil. Eng.* **20**, 313–319 (2012).
 57. J. Decety, The neurophysiological basis of motor imagery, *Behav. Brain Res.* **77**, 45–52 (1996).

Acknowledgments: We are grateful to our pilots, Eric A. and Numa P. for their time, effort and dedication. We would also like to thank our colleagues Dr Robert Leeb and Tiffany Corbet for their assistance during the rehearsal event. Special thanks goes to Prof. Robert Riener, the board and organizing committee of Cybathlon, for undertaking the great burden to make the Cybathlon event happen and, especially, Dr Roland Sigrist, Anni Kern and Nicolas Gerig for promptly answering our numerous requests. We show our appreciation to g.Tec medical engineering for their support with hardware equipment.

Funding: The study was funded by the Swiss National Center of Competence in Research (NCCR) Robotics. **Author contributions:** S. P and L.T. have designed and implemented the training protocol and control paradigm, collected and analyzed EEG data, prepared the pilots, operated the BCI during the Cybathlon BCI race and wrote the manuscript. S.S. and C.S. have collected EEG data, prepared the pilots and operated the BCI during the Cybathlon BCI race. J.d.R.M. conceived the idea, managed and led the Brain Tweaksters team and wrote the manuscript. S.P. and L.T. have equally contributed to this work. **Competing interests:** The authors declare no competing interests.

Table 1. Cybathlon BCI race results. The table presents the race completion times of all competing pilots in the qualification and final races of the Cybathlon BCI race event arranged in ascending order. The race where each time has been achieved is indicated by the second column and the pilot's competition ranking in the last column.

Team (pilot)	Race	Completion Time (s)	Rank
Brain Tweaksters (P1)	Qualifier	90	4
Brain Tweaksters (P2)	Qualifier	123	1
Brain Tweaksters (P2)	Final A	125	1
Neurobotics	Final B	132	5
BrainGain	Qualifier	135	2
NeuroCONCISE	Final B	136	6
Athena-Minerva	Final B	146	7
BrainStormers	Qualifier	146	3
Athena-Minerva	Qualifier	148	7
OpenBMI	Final B	149	8
OpenBMI	Qualifier	149	8
BrainGain	Final A	156	2
BrainStormers	Final A	161	3
NeuroCONCISE	Qualifier	165	6
Mahidol BCI	Qualifier	167	9
Ebrainers	Qualifier	186	10
Brain Tweaksters (P1)	Final A	190	4
MIRAGE 91	Qualifier	196	11

Table 2. Features selected for mutual learning. The table presents all the spatio-spectral features selected for the BCI classifiers trained throughout our pilots' mutual learning process. Each feature refers to a specific frequency band (2 Hz resolution) and EEG channel location according to the international 10-20 system.

P1			P2		
Date	Feature		Date	Feature	
	Location	Band (Hz)		Location	Band (Hz)
	C1	22		C1	26
	C1	24		C1	28
	Cz	12		C1	30

30/06/2016	Cz	20	11/08/2011	Cz	30
	Cz	22		C2	26
	Cz	24		C2	28
	C2	18		C2	30
	C2	20		CPz	26
	CP3	20		CPz	28
	CP3	22	08/09/2016	C1	32
	CP3	24		Cz	28
	CPz	18		Cz	30
	CPz	20		Cz	32
14/09/2016	FC4	28		CP3	30
	FC4	32		CP3	32
	Cz	10		CPz	24
	Cz	12		CPz	26
	Cz	20			
	Cz	22			
	Cz	24			
	C4	26			
	C4	28			
	C4	30			
	CP3	22			
	CP3	24			

Supplementary Material

Table S1. Training session information. The table presents the date of all executed training sessions for both pilots and the number and type of runs performed in each session and reported here. Asterisks indicate one or more runs have been lost due to technical failure or bad maintenance.

P1					P2				
Session #	Session Date	Offline	Online	Race	Session #	Session Date	Offline	Online	Race
1	04/04	5	1	0	1	14/07	4	1*	0
2	08/04	3	0	0	2	21/07	2	0	0
3	18/04	2	0	0*	3	04/08	4	0	0
4	29/04	0	0	2*	4	11/08	3	1*	0
5	02/05	0	0	5	5	18/08	0	2*	2
6	09/05	0	0	8	6	23/08	0	1	8
7	20/05	1	0	3	7	25/08	0	0	7*
8	27/05	1	0	9	8	01/09	0	6	0
9	03/06	0	0	9	9	08/09	0	6	0
10	17/06	2	2	3	10	22/09	2	0	5*
11	24/06	5	0	0	11	26/09	0	1*	7
12	30/06	6	1	0	12	29/09	0	1	7
13	04/07	2	5	0	13	03/10	0	0	2
14	07/07	1*	1	0*	14	05/10	0	0	8
15	13/07	1	0	0*	15	06/10	0	0*	9
16	20/07	1	0	0*	16	08/10	0	0	2
17	25/07	0*	0*	0		Total	15	19	57
18	26/07	0*	0*	0					
19	27/07	1	0	0*					
20	03/08	1	0	0*					
21	10/08	1	0	11					
22	15/08	0	0	10					
23	17/08	3	0	7					
24	24/08	0	0	11					
25	25/08	0	0	16*					
26	31/08	0	0	16					
27	07/09	0	0	10					
28	09/09	4	0	7					
29	14/09	0	0	0					
30	21/09	0	0	10					
31	22/09	0	0	16					
32	28/09	0	1	9					
33	30/09	0	1	14					
34	06/10	0	0	9*					
35	08/10	0	0	2					
Total	35	40	12	182					

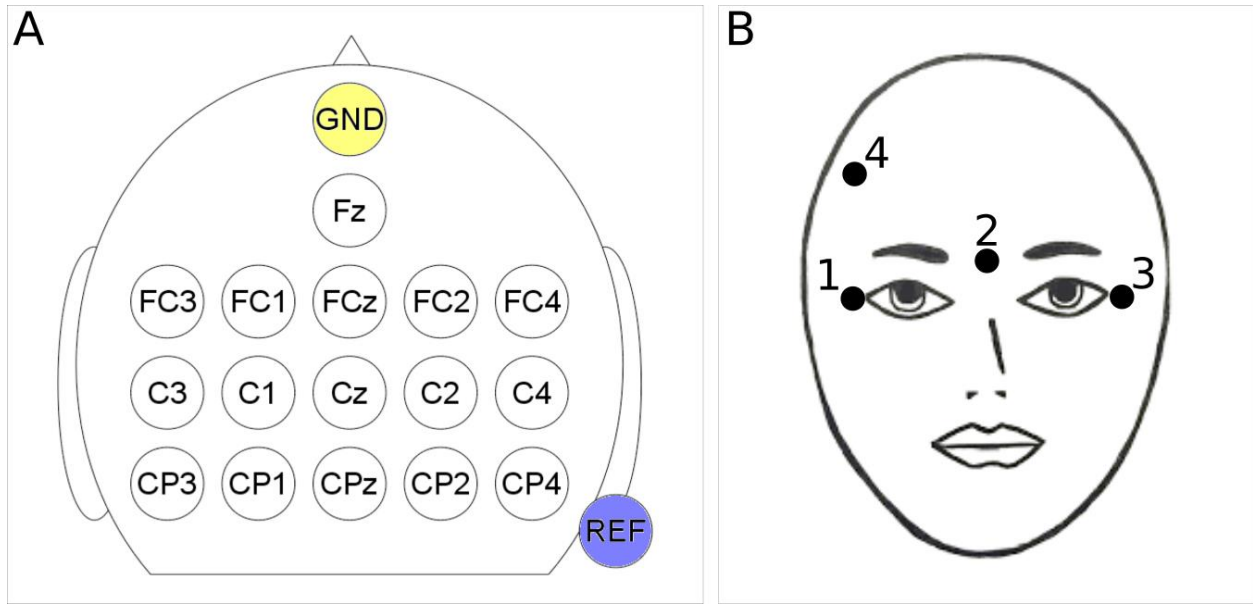


Fig. S1. Electrode configurations. **(A)** EEG channel configuration over 16 locations of the sensorimotor cortex according to the international 10-20 system. **(B)** EOG electrode configuration on the pilot's right and left canthi, nasion and forehead for the detection of ocular and facial muscle artifacts.

Movie S1. Typical race training session of pilot P1. The pilot performs race training at his home a few days before the Cybathlon BCI competition. The video demonstrates the usage setup and the Brain Runners BCI control paradigm.