EEG Hybrid Brain-Computer Interfaces: a scoping review seeking techniques and protocols feasible for children

# Introduction

Children with very limited motor abilities may benefit from the use of brain-computer interfaces (BCI) to access play and learning activities, but there is very little research in the area. BCI are devices that use brain signals processed via computational operations to control machines for various purposes, from rehabilitation to gaming. Despite the long list of interfaces created to this day, most traditional BCI still face challenges in achieving the desired performance needed for reliably controlling assistive devices. Hybrid brain-computer interfaces (hBCI) may be able to address the limitations of traditional BCI (cite Kinney-Lang 2020, Orlandi 2021). The main goal of hBCI is to improve BCI performance (effectiveness and efficiency) through multi-modal signal inputs, e.g. combinations of different brain signals, BCI paradigms and/or other external devices (cite Wolpaw 2012).

In general, studies with traditional and hybrid-BCI have mainly been with adults. BCI for children is harder to implement than for adults. There might be difficulties in identifying signal features (E. Mikołajewska, 2014), recognizing oscillatory brain signals (J. Ehlers, 2012), and instructing young participants to perform the desired self-regulating mental task (J. Z. Zhang, 2019). During experiments, external factors such as lab environment or the presence of the caregiver can distract children and negatively influence the recorded signal (cite Gavin and Davis 2009, de Haan 2009 and Richards 2003).

There have been several reviews on hBCI: Shramila (2020) provided an overview on the types of hBCI for wheelchair-based systems; Li et al. (2019) reviewed the types of hBCI, categorized according to Multiple Brain Patterns, Multisensory and Multiple Signals; de Neeling and Van Huelle (2019) focused on multi-input hybrids and their applications; Sadeghi and Maleki (2018) compared accuracy and information transfer rate (ITR) across systems; Hong and Khan (2017) discussed the combination of brain signals and their application for both clinical and non-clinical scenarios; Choi et al. (2017) did a systematic review and proposed a taxonomy classification system for hBCI systems; Banville and Falk (2016) did a systematic review and discussed experimental protocols, signal processing and study rational; Muller-Putz et al. (2015) compared hBCI applications that had participants with and without disabilities; Amiri et al. (2013) reviewed mutli-brain signal hBCIs; and Allison et al. (2012) summarized the initial efforts in hybridization and the perspectives of hBCI. However, it is difficult to determine from the reviews what might be appropriate to develop for use with children.

The purpose of this review was to examine the literature around hBCI with regards to clinical applications, especially applied to control of devices and communication, with a lens for potential use in the peadiatric population. User requirements were examined as potential indications of ease of use. Acceptable performance levels, according to the known requirements of adults with spinal cord injury were also examined (Huggins, 2011).

# Review Objectives

Two main questions guided this study:

**Question 1 (Q1). What are the existing approaches for hBCI systems that are focused on control of devices, which could be used clinically?**

**Question 2** (

**Q2). What are the least complex and most efficient solutions available?**

articles and that could be used to control assistive devices for reviewFrom the gathered articles, we examined features of the interface design such as how many steps required to achieve the goal, how many activities happening at once, how many devices participants had to focus on, and the type of feedback. From these, we make inferences about ease of use from a child’s perspective.

# Search Methodology

*The keyword search consisted of four parts to specify the hybrid modality, the BCI system, the application type and the acquisition source: (Hybrid\* OR Multiinput\* OR Multi-Sensor\* OR Multi-Device\* OR Multimodal)* ***AND*** *(*“*brain computer interface” OR BCI\* OR hBCI OR* “*human machine interface” OR HMI)* ***AND*** *(Activit\* OR Task\* OR Step\* OR Assignment\* OR Exercise OR Test\* OR Execut\*)* ***AND*** *(((EEG) OR (electroencephalogra\*)) OR (non-invasive)).*

As the definition of a hybrid BCI (hBCI) can be broad, to narrow down our scope, we defined some minimal requirements for a hBCI system to be considered valid for our analysis. **Firstly**, we only considered systems that included signal acquisition via EEG. EEG-based BCI are the most popular system compared to other non-invasive methods (such as NIRS, fNIRS or MEG) and they have the highest information transfer rates (ITR), which puts them in an advantageous position in terms of performance compared to other methods. **Secondly**, the BCI component must have had a primary role in the overall system. Counterexamples of this requirement would be hBCI that used brain signals only for target selection confirmation or hBCI that used brain signals only as a mechanism to switch between non-BCI input modes. **Thirdly**, the multiplicity of inputs or paradigms had to work synergistically to achieve improved results. Our main interest was in systems that combine different paradigms, inputs or sensory pathways attempting to improve traditional BCI. Systems that implemented two paradigms to execute completely unrelated tasks, although they happen to be accessed through the same interface, were not considered for this review. Systems that ed were not considered For example, a system that used one BCI paradigm to move a wheelchair and another to select items on a shelf were rejected. Likewise, systems that integrated an on/off switch mechanism to a previously standalone BCI were also rejected. Even though such systems used valid paradigms or secondary inputs, they were not included in this review as we understood that they did not significantly contribute to the improvement of the system’s performance but rather with its controllability and asynchronous capabilities.

The scoping review methodology proposed by (cite Arksey and O’Malley (2005)) was implemented. Articles from Web of Science, PubMed, Scopus and IEEE Xplore databases were extracted as they focus on medical and engineering topics. The criteria for inclusion and exclusion were delimited per filtering phase, following the scoping review methodology. The exclusion criteria of the previous phases were kept for the next phases in case the article did not explicitly mention an exclusive term in the previous phase. All databases were searched on February 23rd of 2021, and articles published before that date were included. Patents, reviews, and other formats of publication that were not articles or conference papers were not included.

**Title inclusion/exclusion criteria.** Article titles to be **included** had to: (1) contain “hybrid BCI” or other terms that indicated hybridization such as multi-input, multi-modality or multiple paradigms, signal acquisition methods or devices. It also had to (2) make reference to control terms (selection, interaction, classification, etc.) or devices (speller, robotic arm, wheelchair, etc.). The titles that were **excluded** were the ones indicating that the paper focused on: (1) non-hBCI systems (e.g., single-input BCI, multi-input devices), (2) estimation applications (e.g., motion trajectory prediction, group decision making), (3) assessment applications (learning performance, affective/emotion state, mental/psychological state, facial expressions or fatigue), (4) imaging and detection applications (e.g., neuroimaging, algorithms to localize best EEG sources, studies on brain signal detection, cortical reorganization, epilepsy detection), (5) other EEG-related algorithms, (e.g., artifact removal algorithms, algorithm for EEG signal simulation), (6) rehabilitation or therapeutic applications, (7) systems including functional electrical stimulation, (8) invasive technologies, and (9) pure headset development.

**Abstract exclusion criteria.** At this stage, the title-included articles were filtered based on their abstracts. Articles were excluded if they were oriented towards: (1) BCI as a complementary input in a multi-modal system, (2) the study of hybrid classifiers for a single BCI input, (3) optimal channel selection algorithms, (4) development of a method or framework for experimentation, (5) signal identification during other activities or (6) if they had no participants (i.e., only used datasets for validation).

**Full article exclusion criteria.** The final filter allowed a more in-depth analysis of each article, avoiding the inclusion of articles with misleading titles or abstracts. At this phase, articles were excluded if they: (1) proposed an invalid hBCI (i.e., brain signal was a secondary function, any of the inputs in a two-system input was only used to keep/turn the system on/off, paradigms or inputs did not work in synergy), (2) did not have a valid performance measurement (accuracy or true positives, true negatives, false positives and false negatives, or any indication of the number of correct trials versus the total) of the paradigms’ efficiency (as opposed to the task accomplishment, which was not the main interest), (3) had online trials but only provided offline performance measurements, (4) did not include performance measurements for the relevant system role of operation, either for each of the inputs or their combination, (5) proposed a system that was not suitable for control applications, (6) had insufficient information for experiment replicability (i.e., lacked information such as but not limited to number of trials, number of participants, number of sessions, number of analyzed samples, number of training/validation datasets, or clarity about experiment protocol).

#### Data extraction - I'm not sure what heading level this should be, but I can only pick Heading 4 or 5, for some reason 1, 2 and 3 disappeared on me.

<Seems like a heading could be useful here - the above is about the search methods, and starting here and for the taxonomy section, it's about the information that was extracted from the papers. Is there an overarching term for that? if not then could "data collection/extraction" work? >

<then some sort of description about the kinds of things examined <to set the reader up for why the basics like participants, ages, equipment, etc. Perhaps...

Descriptive information was extracted from the articles such as study population, size, age, control task, and the tools that were used for the development of hBCI (model, programming languages, toolboxes). The data from the articles was extracted and labelled according to the taxonomy proposed by Choi et al. (2017). The categories used were: 1) diversity of input signal, 2) role of operation, 3) mode of operation, 4) mental strategy, 5) brain signal signature, and 6) stimulus modality. A brief description of the categories is provided here:

* **Diversity of input** is categorized as homogeneous or heterogeneous, depending on whether the input types are brain signals only, or brain signals combined with other inputs, respectively. Homogeneous inputs can have a single-brain signal approach (electroencephalography (EEG) only) or a multi-brain signal approach (e.g., EEG and functional near-infrared spectroscopy (fNIRS)). Heterogeneous inputs combine multi-physiological signals (e.g., EEG and electromyography (EMG)) or external inputs (e.g., EEG and Eye Tracker).
* **Role of operation** refers to the role of each system and how they are chronologically bound together. The role can be simultaneous, where both systems work concurrently in either the same or in different parts of the task. It can also be a sequential-switch, where one system initiates the other system, or a sequential-selector, where one system partially completes the task and the other system confirms or rejects the selection.
* **Mode of operation** is the mode with which the experiment is paced. For synchronous experiments, stimuli are presented within a specific timeframe and cues are used. Asynchronous interfaces are self-paced by the subject, with more flexible timeframes.
* **Mental strategy** is categorized as either selective attention or operant conditioning. Selective attention strategies rely on external stimuli to generate expected brain responses, while operant conditioning strategies (also known as slow cortical potentials) rely on the self-regulation of the subject to generate distinguishable brain responses.
* **Brain signal signature** are defined by the mental paradigm used for the interface, and is directly associated with the mental strategy. For selective attention, the steady-state evoked potential (SSEP), transient event-related potentials (ERP) and motion-onset evoked potential (mVEP) are possible signatures. For operant conditioning, slow cortical potentials (SCP) can be modulated via movement related efforts (sensory-motor rhythms - SMR) or attention levels (µ-rhythm). Other mental tasks involving music and speech imagery were also classified as SCP.
* **Stimulus modality** is the pathway through which the user is stimulated so that the brain can elicit predictable signals. The pathway can be sensorial such as visual, tactile or auditory, or self-induced in the case of operant conditioning, defined as the operant pathway. A further classification can be made in terms of diversity of stimulus modalities within the interface. Single modality uses the same sensory pathway for all inputs and paradigms, and multi-modality uses different sensory pathways for the same brain signature (e.g., SSVEP and SSSEP).

Effectiveness and efficiency were the two main performance metrics considered. As most of the BCI community uses accuracy (or other parameters that allow for accuracy deduction), we used accuracy as our metric for effectiveness. On the other hand, efficiency is measured in a variety of different ways. Most papers that do present efficiency metrics, use information transfer rate (ITR), but execution time or commands per minute are also recurrent. It is important to notice that our main interest was in the metrics referring to the overall hBCI system’s classification effectiveness and efficiency, and not to the task accomplishments.

# Search Results

When all the filters were applied, 45 articles were selected for this scoping review. Initially, the search on all databases yielded 1585 publications, 617 from Web of Science, 225 from PubMed, 489 from Scopus, and 244 from IEEE Xplore. The number of duplicates was 1214, which resulted in 771 unique articles when removed. From those, 302 were included after title-filtering, 149 after abstract-filtering and 45 after article-filtering. No conference papers remained among the 45 final articles, although we did not set a strict exclusion criteria for conference papers.



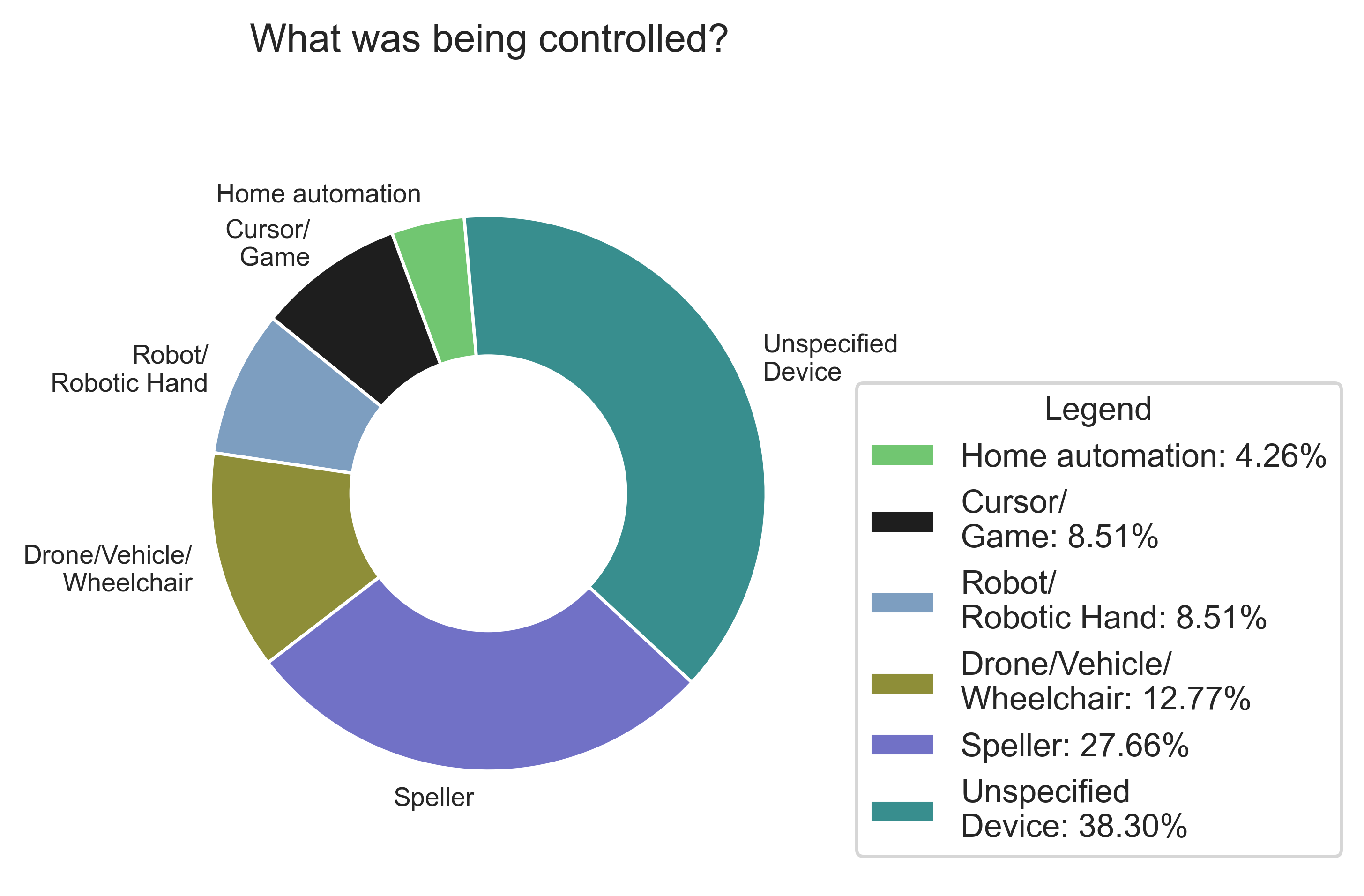
Figure XI shows how many articles were published per year. The transparent bars correspond to all the articles included by abstract. The solid bars correspond to the articles included for the final analysis. For both groups, the overall number of articles per year has been growing. In 2020, 14 articles were written that matched our criteria, over three times more than in the previous six years.



No studies included child participants. One study had a participant aged 18 years (Kaongoen & Jo, 2017), but most studies had at least one participant within the range of 20 to 30 years, except for Nann et al. (2020) who worked with tetraplegic participants from X to y years). Only two papers included participants above 40 years of age (Brennan et al., 2020 and Nann et al., 2020). Figure YYa displays the age range of participants in each study with the achieved average accuracy. The most common population size was ten participants (21.3% of the studies), as seen in Figure N.

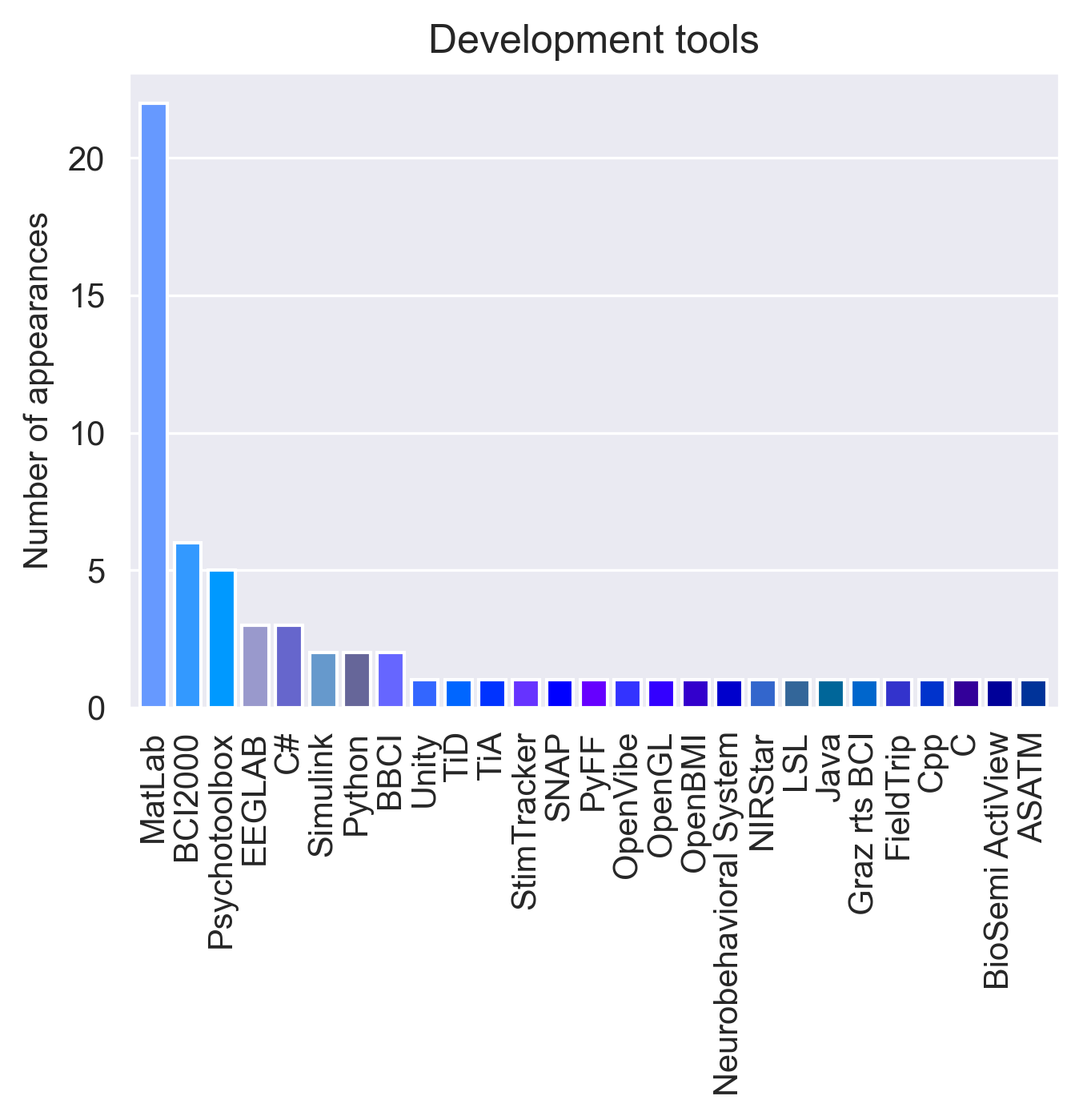
Only three studies included participants with disabilities. Soekadar et al. (2015) tested the system with one participant with flaccid hand paralysis, a 34 year old male. The study reported that he was able to control a robotic hand via motor imagery (MI), even though his accuracy was slightly lower than the average of the other neurotypical participants (76.03% compared to 80.65%). Brennan et al. (2020) included 14 participants with brain injury (although only nine completed the hBCI trials) with an average age range of 41.6±13.9 years. Participants underwent trials with both an SSVEP BCI and an SSVEP Eye Tracker hBCI for comparison. The hBCI trials had higher accuracy than the BCI, with 99.14% compared to 80.26%. Participants with brain injury only did one session of experiments, while neurotypical participants did two. Nann et al. (2020) had four participants with tetraplegia with an average age of 51.8±15.2. The study tested an EEG BCI and an hBCI combining EEG with horizontal oculoversion, increasing the accuracy from 58.68±10.62% to 81.25±5.84%. All participants with tetraplegia rated the system as user-friendly and reliable.

### Task Categories

The devices controlled are represented in Figure H. Eighteen of the hBCI systems were oriented to control, but only controlled the interface, and did not specify the device. The most common control task was spellers, with 13 systems. Six articles controlled devices that were moved with brain signals such as drones, wheelchairs and other vehicles (physical or simulated). Four controlled robotic devices, four controlled cursors or games, and two focused on home automation systems.



Twelve out of 42 of studies used the g.USBamp, as shown in Figure CC. Most of the papers reported using a fabric cap with Ag/AgCl electrodes or golden cups. The only study that used a headset-style was the Cognionics system, used in Yang et al. (2020). Articles that combined multiple input caps were: Buccino et al. (2016) combining the microEEG with the fNIRS NIRScout; Khalaf et al.(2020) combining the g.USBamp with the SONARA TCD; Chiarelli et al. (2018) combining the Net300 with the Imagent fNIRS; Shin et al.(2018) combining Biosemi with LIGHTNIRS; and Glowinsky et al. (2018) combining BrainAmp with ETG-4000 NIRS.

The stated tools that were used for the development of hBCI are represented in Figure PP. Programming languages, toolboxes, stimuli generators and processing tools were considered. The most used language was MatLab, followed by Python and C#, C and C++. BCI2000, Psychotoolbox and EEGLAB were the most used toolboxes, usually paired with MatLab. 

## Taxonomy report

### Diversity of Input

Considering the diversity of input, 36 out of 42 of the systems (76.6%) were homogeneous and 11 (23.4%) were heterogeneous, as show in figure X. Thirty of the homogeneous systems used EEG only and the reminder used a multi-brain signal input approach: two combined EEG and fNIRS, two combined EEG and NIRS, and two combined EEG and fTCD. All the multi-brain signal input systems only presented offline results. Of the eleven heterogeneous systems, eight were multi-physiological and three made use of external input. The multi-physiological signals were mostly EEG and EOG, but Lin et al. (2016) combined EEG and EMG and Zhang et al. (2019) combined EEG, EOG and EMG. As for the ones with external input, Mannan et al. (2020) and Brennan et al. (2020) used an eye tracker and Saravanakumar & Reddy (2019) used EOG combined with a video-based eye tracker.

### Role of operation

Almost 60% of systems were simultaneous in their role of operation, totaling 29 out of 42 systems. Eighteen were sequential, including five sequential-selectors (Lee et al. 2018, Yang et al. 2020, Fan et al. 2015, Ko et al. 2020 and Long et al. 2012) and two sequential-switching systems (Li et al. 2018 and Yu et al. 2017).

### Mode of operation

Most experiments relied on cues and used the synchronous mode of operation. Ten had asynchronous modalities, where the participant could self-pace his selections. Zhang et al. (2019) was the only study to utlize both synchronous and asynchronous. Due to its multi-input nature, Zhang and colleagues allowed the EOG and EMG to operate asynchronously, and when the EEG mode was selected, the system switched to cue-based operation.

### Mental Strategy & Brain signal signature

Almost half of the systems used selective attention (23 systems, 48.9%), about a quarter used operant conditioning (13 systems, 27.7%), and the reminder combined both (11 systems, 23.4%). The selective attention systems were equally divided in terms of paradigms: six used only SSEP, six used only ERP and eleven combined both. As for the operant conditioning, all used SMR. Most were exclusively SMR (10 systems), two combined SMR with µ-rhythms, and one used SMR combined with SCP. The systems with multiple mental strategies mostly used SSEP, and combined it with SMR (6 systems), SMR and µ-rhythm (1 system) and SCP (1 system). The other three combined ERP and SMR.

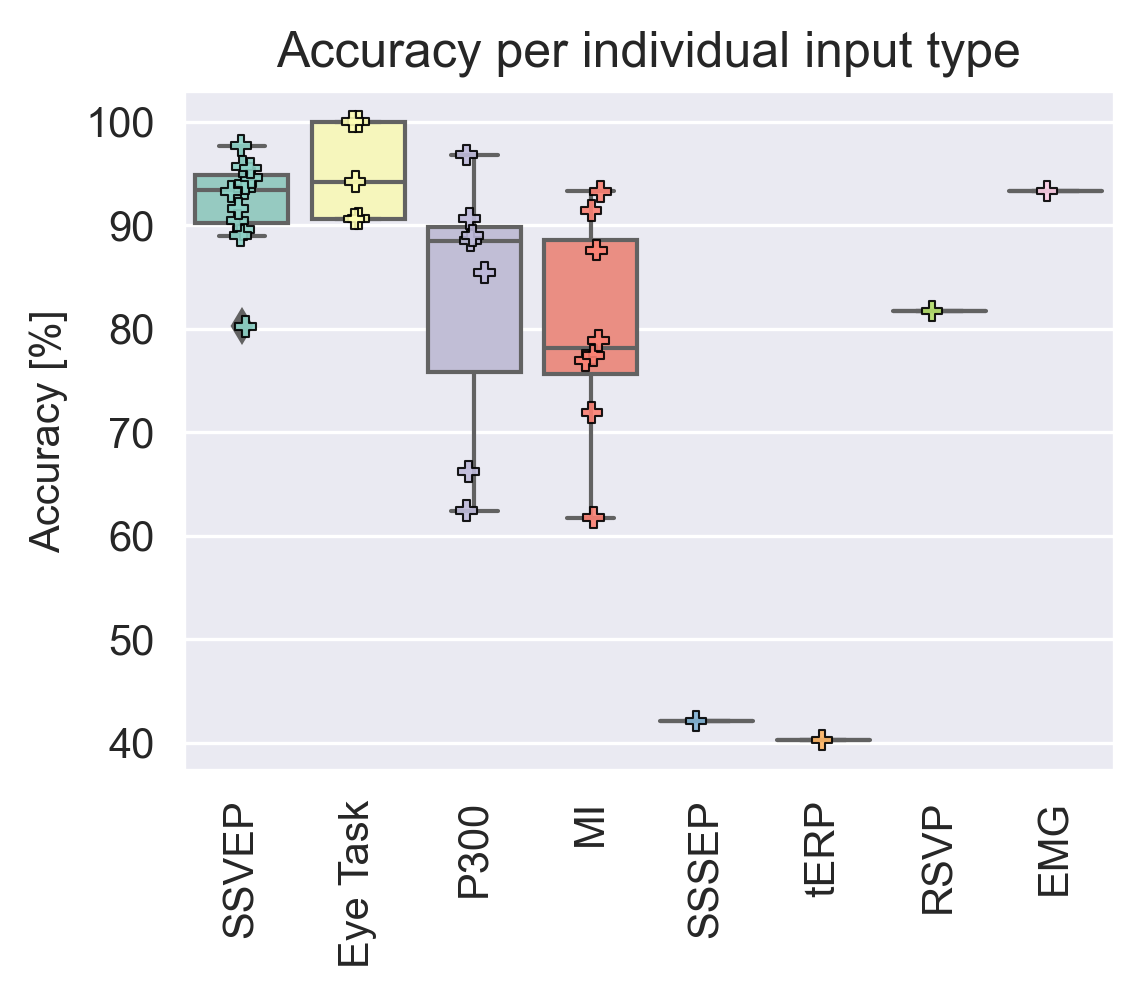
Figure Y shows the percentage of brain signatures that were used. The total number for each brain signal signature is represented by the outermost ring, and the combinations made with each signature are represented by the innermost ring. The other brain signatures were always combined with other signatures. All the systems using µ-rhythms had SMR as well, therefore, the SMR combination is not represented in the figure.

### Stimulus modality

Figure AA shows how the stimulus modalities were distributed for the considered systems. Some papers presented more than one variation of the system. In the 42 papers, 47 systems were presented. Khalaf et al. (2020) presented two systems with different brain signals, one using SMR and one combining SSEP and SCP. Other papers presented synchronous and asynchronous experiments using spellers for cue-based experiments and free-spelling (Lee et al., 2018, Lin et al., 2016 & Xu et al., 2020), and others did both sequential and simultaneous stimuli interfaces (An et al., 2014 & Ahn et al., 2014). Therefore, all statistics were calculated with 47 total systems, unless otherwise stated.

The matching color sections between the inner and outer ring indicate a single stimulus modality. The two most utilized stimulus modalities were visual and operant. Twenty of the 47 systems were purely visual, and 32 systems had visual stimulus combinations. Ten systems were based on operant stimuli and 22 of systems included operant stimuli. The combination of both visual and operant was also used, totaling nine systems using this modality. Only five systems included auditory stimuli (two purely auditory, two combined with visual and one combined with visual and operant stimuli) and three systems included tactile stimuli (one purely tactile and two combined with operant stimulus).



Sixteen of the 42 studies only performed offline experiments. Some articles made available the individual input results while in hybrid mode during online trials. The average accuracy according to input device is represented in Figure Jj. When multiple conditions were tested (e.g. results with different number of runs, with and without correction mechanisms, with more or less samples, etc.), the best results were considered. Preference was given to results from real-world control applications rather than simulations. Eye gaze and EOG activities (blinking, frowning, vertical/horizontal movements, etc) were reported as eye-tasks.

<<TABLE WITH ALL THE ARTICLES>> <<HERE OR IN SUPPLEMENTARY MATERIAL?>>

# Discussion

## Complexity considerations

From the analysed articles, we found some features that might play an important role on the complexity of a hBCI system. The first five were advenient from taxonomic traits and the others were related to the interface. These features are important to be considered when designing hBCI as they can directly impact the workload, appeal and the levels of engagement of children when using the system:

* Taxonomic traits:
  + Number of Brain Signatures
  + Role of Operation: Sequential, Simultaneous
  + Stimulus Modality: Visual, Tactile, Operant, Auditory or multiple
  + Number of Stimulus Modalities
  + Includes external inputs
* Interface traits:
  + Type of targets: still, strobe, goal/spatial, off-screen
  + Number of targets
  + Number of actions before selection

Although there were no studies with children using hBCI, the descriptions and ordering of elements below were elaborated based on the expected cognitive development of children and how easily it could be “adapted to the children’s age, deficits, preferences and needs” (Mikołajewska and Mikołajewska, 2014).

### Taxonomic traits

**Number of Brain Signatures** - Switching between brain signatures or performing multiple brain signatures simultaneously can increase the complexity of the system, especially if the brain signatures belong to different mental strategies. For example, Duan et al. (2015) utilized SSVEP to move a robot, mu-rhythms to switch modes and MI for grasping. It was the only system with more than two Brain Signals for control. The accuracy was lower than the average accuracy of all the included papers (85.64%) with 73.3%. Similarly, but in a simultaneous role of operation, Alisson et al. (2012) developed a system where a ball could be moved in a 2D space utilizing SSVEP and MI for horizontal and vertical movement, respectively. The average accuracy was 60%.

**Role of Operation** - Multi-tasking generally decreases processing speed and increases the amount of information needed to make a decision (cite Howard et al. 2020). Using multiple brain signals simultaneously can decrease participant’s ability to focus or increase mental fatigue. Ahn et al. (2014) designed two experiments combining MI and tactile selective attention, with sequential and simultaneous roles of operation. The results with the sequential experiment yielded 71% accuracy, while the simultaneous reached 60%, thus, multi-tasking reduced accuracy in this task.

**Stimulus Modality** - Most papers chose visual pathways to stimulate the brain. Visual stimulation is the most used, and it is also the least complex modality. Figure K shows the relationship between accuracy and each Stimulus Modality. Although not all modalities had the same sample size, we can see a trend where systems using visual pathways have higher accuracies than the others. *Visual* paradigms, in general, elicit clear signals over the Occipital and Parietal regions, especially when using SSVEP and P300. On the other hand, *operant* modalities require a certain level of training and focus from the participant to generate distinguishable signals (cite Yuan and He, 2014). The *auditory* modality was considered more complex than the operant modality because it requires more attention and has a steeper learning curve than operant modalities (cite Nijboer et al., 2008). Lastly, *tactile* modalities require body awareness, and can become confusing with multiple targets (cite Brouwer and van Erp, 2010).

**Number of Stimulus Modalities** - Switching between stimulus modalities or having to simultaneously focus on more than one sense might also increase the system’s complexity, especially if the different stimulus modalities refer to a different target. An et al. (2014) experimented with two roles of operation and each paradigm individually. In the sequential mode, stimuli alternated between visual and auditory repeatedly with different sequences within 300ms, so that two independent decisions could be made in parallel. The participants reported that the modality had a considerably higher workload than the simultaneous mode or the individual stimulus modalities.

**Includes external inputs** - Having an external input can facilitate selection if the participant does not have severe impairments. Inputs such as eye trackers, joysticks and buttons add reliability to the system and therefore increase its performance. Three of the selected papers used external inputs, Mannan et al. (2020) and Brennan et al. (2020) used NIR eye trackers and Saravanakumar & Reddy (2019) utilized a camera-based eye tracker. All the systems used the eye gaze as a means to narrow down the possible targets. The gaze indicated the most likely region on the screen where the final target should be. Mannan et al. and Saravanakumar and Reddy used the gaze to select the character sub-group and the SSVEP to select the character within the sub-group while Brennan and colleagues used the gaze to select the sub-region on the screen and compare it agains the SSVEP selection for the final decision. Their accuracy results were among the highest (average 96.92%), even when utilized by the nine participants with brain injury from Brennan et al. (2020) (accuracy of 99.14%).

### Interface traits

**Type of targets** - The stimuli that happen on the screen can either elicit a certain brain response or indicate to the participant what self-regulating action to take. There are indications that certain types of targets can cause more fatigue in users (Seo et al., 2019), while other types of targets place all the stimulus responsability on the participant. For example, SSVEP had a higher eye-fatigue level than P300. Based on required effort and likely fatigue caused by the stimulus, we ordered the existing types of target (from the selected papers) from least to most complex.

On-screen targets require visual focus on the stimuli so that the brain can evoke certain signal patterns. The *still targets* flash periodically (usually with less than 6 Hz) with a certain inter-stimuli interval and are usually associated with P300 paradigms and spellers. Those targets generally require counting and focus on a single desired target. The *strobic targets* have flashing with higher frequencies (usually above 6 Hz) incorporated onto them. They are mostly used in SSEP or rapid serial visual presentation (RSVP) paradigms and can change in intensity, color, shape, visuals or position, and targets usually have different frequencies. *Goal or spatial targets* stimulate users to displace objects, cursors or other elements over time by triggering certain threshold of intensity. They are mostly used in SCP paradigms.

There are also off-screen targets, which require greater focus and mental training as they do not present stimuli, nor feedback in some cases (Mahmoudi and Erfanian, 2006). These include *motor/tactile targets*, which require focus on certain motor imageries or tactile stimulations. This approach could be an alternative for people with severe visual impairments. The *mental tasks* measure the blood flow generated in the brain when arithmetic operations, mental gemoetric manipulation or word formation are performed by the participant. They are usually associated with NIRS and fNIRS inputs. Finally, *sound cues* are targets that rely on sound for selection. These targets can be difficult to distinguish and even when the audio tracks are substantially manipulated as in Glowinsky et al. (2018) or An et al. (2014), they yield lower accuracies (77.43% and 66.2% for the auditory component, respectively) and have higher workload for mental demand and effort compared to visual stimuli (Nijboer et al., 2008).

**Number of targets** - Some authors have attempted to increase the number of targets to increase the ITR. Although it might be a good strategy that can give the user more flexibility and a faster system, a greater number of targets could make a user distracted or overwhelmed with many options. A greater number of targets was mostly seen in spellers. For example, Xu et al. (2020) developed a speller with 108 targets. Twelve 3x3 character matrices were presented to participants at once and they underwent a synchronous and asynchronous experiments. Although they had some of the highest ITR (172.5 bits/min for synchronous and 164.7 bits/min for asynchronous) they had the lowest accuracies (81.67% for synchronous and 79.17% for asynchronous) compared to the other spellers, with average of 90.7%.

**Number of actions before selection** - Some systems required multiple sub-tasks before the final selection was completed (i.e., multiple input commands and classifications to make a final selection). Although it appears that having multiple sub-tasks can add to the robustness of the system, it could also increase its complexity. Most systems in the literature had two sub-tasks, at maximum (22 out of 47 articles). Seven systems utilized three sub-tasks and Zhang et al. (2019) was the only one with four sub-tasks. Their system utilized EEG, EOG and EMG modes, each with specific commands. EOG blinking switched modes and a participant could need up to four commands to cycle through all the modes and then make a selection.

## Children considerations

As exposed by different research groups, the implementation of hBCI can be a bigger challenge for children than for adults. We will now discuss further different topics that might be especially relevant for children.

**Multiple headsets and caps** - Some authors have exposed that children report discomfort when using a single cap or headset (cite Zhang et al. 2019 and Jadavji et al. 2022). Systems that combined multiple input caps (Buccino et al. (2016), Khalaf et al.(2020), Chiarelli et al. (2018), Shin et al.(2018) and Glowinsky et al. (2018)) could face the challenge of attaining children’s focus while they face discomfort from overlapped caps/headsets. This challenge might require new approaches to headset designs so they can be more comfortable and sizable for children. Additionally, headsets with a built-in capability of measuring different brain inputs might be the next generation of sought-after hBCI systems for research. The creation of new headsets that integrate multiple inputs like the OpenBCI Galea demonstrate this trend.

**Target types** - Based on qualitative responses and the reported accuracies, we infer that some target types can reduce the mental overload on children using hBCI. The complexity order of targets that we propose, from less to more complex, for children is *strobic targets, still targets, goal or spatial targets, motor/tactile targets, mental tasks*, and *sound cues*.

Most of the qualitative responses were for systems containing SSVEP stimuli. Alisson et al. (2012), Brunner et al. (2011), Alisson (2010), and Mannan et al. (2020), reported low annoyance for strobic targets. Figure Jj shows that the SSVEP paradigm had the highest accuracy average, followed by the P300. It is possible that children can be more annoyed by the flashing than adults, but current research shows that children can perform well using SSVEP (cite Norton et al. 2018).

Goal or spatial targets were associated with MI and required focus to maintain a certain brain pattern. Long et al. (2012) associated MI to P300 and Allison et al. (2012) associated SSVEP and MI to control a 2D-cursor on the screen, and Mousavi et al. (2020) combined MI with ErrP to improve accuracy. We assume this type of target could be the most engagin for children as it is easier to add graphical elements with attractive and game-like designs. Long and colleagues acheived 92.84%, Allison et al. 60% and Mousavi and colleagues 75.33% in online trials.

Motor/tactile targets were associated with MI, motor execution, tactile ERP and tactile selective attention, ranging from 44.5 to 93.98% accuracy, with average of 78.81%. This target type, as well as all the other off-screen targets, can be especially beneficial for vision-impaired children. Nonetheless, the studies involving tactile stimuli had accuracies below 90% and only one study reported online results. Yao et al. (2014), Ahn et al. (2014) and Yao et al. (2017) combined MI with tactile selective attention. Yao et al. (2017) reported 86.1% accuracy offline, Yao et al.(2014) acheived 83.1%, and Ahn et al. (2014) attempted a simultaneous and sequential role of operation and acheived 60% and 71%, respectively. Breitwieser et al. (2016) included tactile ERP with steady-state somatosensory evoked potential for an online experiment, but the accuracy was of 44.5%.

Targets involving Mental tasks could be used in schools and academic settings. Mental targets require applied effort to visualize or do mathematical operations during trials. This type of target is usually associated with other inputs, such as NIRS (Shin et al. 2018) and fTCD (Khalaf et al. 2020). The weaknesses of mental task experiments is that they have a low ITR (4.7 bits/min for Shin and colleagues and 4.46 for Khalaf et al.) and that it could be difficult for children to be interested on those kinds of experiments as it involves not only focus, but also a task that might not be playful. Moreover, all the studies did not have online trials, preventing the assessment of the performance in real-world scenarios.

Finally, auditory cues were reported to be more difficult. Although it could also be a good alternative for children with impaired vision, An et al. (2014) reported a higher workload for off-screen stimuli when comparing visual and auditory P300. Auditory P300 had a lower accuracy (66.2% at its highest, compared to 85.4 for the visual P300) and it is known to be harder to learning (cite Nijboer et al., 2008). Kaongoen and Jo (2017) and Glowinsky et al. (2018) also included auditory cues but were both offline studies. Kaongoen and Jo was a preliminary study combining auditory steady-state response and auditory P300 with 85.33% accuracy and 9.11 bits/min. Glowinsky and colleagues tested auditory P300 with NIRS and EEG inputs. The final accuracy was 77.43%.

**Stimulus modalities** - We hypothesize that there would be different levels of complexity associated with each stimulus modality, especially for children. We suggest that modalities for kids, from the easiest to the hardest, based on the accuracies per modality seen in Figure K and Jj, would be visual, operant, auditory and tactile. Visual modalities and paradigms have been more widely used and tend to have better performance than other modalities. Operant is the second most popular modality and consistently performs with more than 70% accuracy. Not quite as high as visual, which usually attains more than 80% accuracy. Auditory and tactile modalities do not have enough studies to statistically confirm their lower performance compared to visual and operant. However, the studies included in our scoping review presented accuracies and ITR below the average of visual and operant modalities. We can also deduct that both auditory and tactile modalities would require a a higher level of auditory perception and body awareness, which might not be well developed in most children.

**Built-in correction/confirmation capabilities** - It was also made evident that having mechanisms to amend or confirm selections can increase performance. Mousavi et al. (2020) utilized ErrP to correct MI misclassification and the system had an improvement in accuracy. Similarly, Soekadar et al. (2015) implemented a task correction with EOG which resulted in a more intentional operation of the system. Fan et al. (2015) implemented confirmation mechanisms before the final selection utilizing SSVEP, resulting in one of the highest evaluated accuracies (99.07%). Such mechanisms can be beneficial for children, especially when starting to use hBCI.

# Conclusion

This scoping review analysed 42 papers that present 47 different systems. Papers were focused on clinically viable hBCI that were EEG-based and that had a diversity of inputs or brain signals attemtping to improve performance. Using taxonomy and other interface traits, we made some considerations as to how systems could be made more or less complex for users in general, but especially for children. Such considerations were based on accuracy and ITR results, and also qualitative responses presented by studies.

We conclude that hBCI systems that have a single brain signature, using a sequential role of operation with a singular visual stimulus modality and external input should have a lower complexity than other combinations. Additionally, interfaces using from 2 to 5 (or less than 37 for spellers) strobic targets, with single or double actions before selection can also attain good performances while keeping the system simple. The deductions made throughout this paper could serve as a guideline for future researchers that are developing hBCI for children.

# Limitations

The authors of this paper only considered EEG-based systems as they are the more commonly used, but hybrids that do not use EEG also exist. We also acknowledge that the sample size for statistical assumptions is small and some of our conclusions could be skewed due to the uneven distribution of studies across features. We also did not use criteria based on feedback for what might be important to consider for children using hBCI as we could not find studies with the peadiatric population. Future research might prove wrong some of the deductions made in this paper.

# References

(Scoping Studies: Towards a Methodological Framework, Hilary Arksey & Lisa O’Malley (2002))