**[ChNum]**

**[Designing Human-Computer Interaction with Neuroadaptive Technology]**

**[Stephen H. Fairclough]**

**[School of Psychology, Liverpool John Moores University, UK]**

**[NON PRINT ITEMS]**

**Abstract:**

Neuroadaptive systems are designed to make adaptations to the user interface based on implicit monitoring and analyses of neurophysiological data. This closed-loop approach to human-computer interaction can be used to enhance attentional regulation, sustain motivation and enable implicit selection from a range of options. While current research has focused on the development of sensors, refinement of signal processing and the accuracy of machine learning models, the design of the human-computer interaction (HCI) has been relatively neglected. The current chapter considers how neuroadaptive systems can enhance the quality of HCI, particularly with respect to adaptive information presentation, dynamic task adjustment and implicit detection of preference. To achieve these three functions, neuroadaptive technology must function as a self-regulating mechanism, capable of autonomous control in accordance with higher-level goals. The chapter explores this self-regulatory concept with respect to the design of higher-level goals and how the neuroadaptive interface may strike a balance between user agency and the autonomy of the system. The chapter concludes that the neuroadaptive interface should support the goals of the user in a way that is logical and transparent. Due to potential problems of robustness when neuroadaptive technology is used in the field, it is suggested that a provision for manual intervention is designed into the interface in order to preserve the agency of the user.

**Key Words: Neuroadaptive Technology; Human-Computer Interaction; Interface Design; Autonomous Technology; Human Factors**

# Introduction

Neuroadaptive technology was defined in 2003 by Hettinger and colleagues [1] as “an ensemble of computer-based displays and controls whose functional characteristics change in response to meaningful variations in the user’s cognitive and/or emotional states” (p. 220). This blueprint was inspired by earlier work on biocybernetic control [2-4], which provided fundamental elements for the development of physiological computing [5-7] and passive BCI [8-9], where implicit changes in brain and body data serve as inputs to a technological system.

Despite differences in terminology, these technological systems are united by a closed-loop design that subverts contemporary paradigms of human-computer interaction. Neuroadaptive technology permits implicit modes of communication that are ‘hands-free’ and do not require any physical input. These covert channels may be unconsciously activated from the perspective of the user [10-11]. Contemporary HCI is characterised as a dyadic relationship between people and machines where technology remains oblivious to the intentions, activities and psychological status of the human user. Hettinger et al (2003) described this type of HCI as asymmetrical in the sense that the operational status of a computer can be interrogated by a user but not vice versa. This absence can be rectified by symmetrical forms of neuroadaptive interaction [1]. When HCI is symmetrical, the computer can respond proactively to the intentions, emotions and cognitive status of its user. Hence, interaction with a neuroadaptive system can deliver implicit and reciprocal types of communication with the user, which both expand bandwidth of information exchange, i.e., by enabling tacit communication from person to computer via continuous changes in physiology, and enhance the autonomy of the machine [12], i.e., by allowing the computer to respond dynamically and proactively without manual input from the user.

Neuroadaptive interaction is presented as a solution to physical and cognition limitations on the capacity of users to communicate with technology, but the question of whether this innovation will ultimately improve or degrade the quality of HCI remains undetermined. Neuroadaptive forms of symmetrical HCI will undoubtedly increase the fidelity of information exchange between user and computer, but can this innovative and symmetrical mode of HCI create tangible benefits for the user population at large?

To answer that question, we must evaluate the concept of neuroadaptive technology in terms of perceived usefulness [13] rather than technical innovation; we must also assess the value of this emergent technology for the user population from both utilitarian and hedonic perspectives [14]. Any evaluation of value is relativistic and governed by the precise interaction between the individual and the technology [15]. For example, the value of brain-computer interfaces (BCI) for those with a limited capacity for communication are obvious, whereas the value for healthy users are perhaps more limited and ambiguous in comparison [16]. There are also aspects of neuroadaptive technology that may not appeal to healthy users, at least during the early phase of system adoption. For example, these systems will necessitate a sensor of some kind, which may be worn or implanted; in addition, and the user must cede a degree of autonomy to technology (and technological processes) to order to use the system. If neuroadaptive systems are to be adopted at scale within the population, the utility and hedonic value delivered by this technology to the user must be both significant and indisputable.

All varieties of neuroadaptive technology converge around a closed-loop feedback model that was originally derived from cybernetics [17]. Neuroadaptive technology is perhaps best analysed by dissecting this closed-loop into its constituent components. The beginning of the process is characterised by incoming data streamed from a sensor and processed and quantified to create an Input Function. The precise form of this function can vary considerably, from event-related responses to discrete stimuli to spontaneous fluctuations in dynamic signals. The psychological meaning of those signals is inferred via a Classification/Inference Engine and machine learning techniques are the dominant methodology for executing this analytic process, which often yields a categorical output, e.g. high vs. low mental workload (for system designed to assess mental workload level), anger vs. neutral vs. happy (for system designed to classify emotional states) [18], [19].

Current research in this field has focused heavily on sensor design, signal processing and machine learning, which represent the Input Function and Classification Engine. However, the closed-loop within neuroadaptive technology must also incorporate a Goal or Standard, the purpose of this goal is to describe the high-level goal of the technology. For example, the original biocybernetic loop described by Pope et al [20] was designed to utilise real-time changes in EEG to sustain a high level of task engagement; therefore, sustaining high task engagement was the Goal of the system. This specific system achieved its goal by activating and deactivating system automaton, effectively making the task more or less demanding. Adaptive automation at the interface represented the Output Function of the loop, i.e., the means by which the system achieved its higher-level goal.

Designing a valid and robust signal processing pipeline for neuroadaptive technology is very important with respect to having an Input Function and Classification Engine that effectively measures changes in the psychological state of the user, which ensure that changes at the interface are triggered in a timely fashion. But it is equally important to design Output Functions at the interface that offer utility and hedonic value to the user. On a similar point, there should be a complementary relationship between the Output Function and the Goal of the system, because the former represents the means by the system achieves the latter. The closed-loop framework provides an integrated and holistic perspective on the design of neuroadaptive technology; the Input Function feeds into the Classification Engine, which triggers an Output Function at the human-computer interface that is consistent with the higher-level Goal of the system. Understanding of the process of how concepts are inferred from measures and translating action at the interface is essential for designing, prototyping and troubleshooting this type of neuroadaptive system.

The current chapter will focus on the topic of designing higher-level goals and output functions for neuroadaptive systems. The context for this discussion is provided by an analysis of the precise ways in which neuroadaptive technology can enhance the quality of human-computer interaction that we currently enjoy from existing technology. The following section will delineate the thorny issue of designing high-level goals for neuroadaptive technology, which begs a number of questions about the ultimate purpose of this technology. The final part of the chapter will consider how designers may strike a balance between the agency of the user and the autonomous, proactive qualities of neuroadaptive interaction.

# Can Neuroadaptive Technology Improve the Quality of Human-Computer Interaction?

The provision of value to the user is the ultimate goal of any technology. The exact manifestation of such value can take several forms, the technology could endow the user with a unique class of functionality (e.g., communicate using implicit signals from the brain) or yield a cumulative improvement over those benefits already received from existing technology. The same logic applies in the case of neuroadaptive systems, acceptability and adoption within the general population will be determined by the demonstrable value and utility of the technology.

The use of closed-loop control enables a neuroadaptive system to effectively regulate and respond to the changes psychological state of the user via a repertoire of adaptive responses at the interface. This regulatory capacity permits neuroadaptive technology to play an active, reciprocal role during an interaction with the user, triggering adaptive changes at the interface in a timely fashion to enhance communication and shape underlying psychological processes.

The categories of functionality enabled by neuroadaptive technology can be broadly divided into three groups: (i) adaptive information presentation, e.g., filtering, scheduling, format, (ii) dynamic adjustment of task demand, e.g., direction (increase/decrease), magnitude, timing, and (iii) implicit prediction of user preference. With respect to (i), the regulation of attention [21-23] determines which environmental stimuli are prioritised, how attention is divided between concurrent stimuli and whether attention is sustained over time. A neuroadaptive technology that accurately monitors the process of attentional regulation can shape the direction and intensity of human information processing via targeted adaptations at the interface, e.g. enhanced stimulus salience, suppression of distracting stimuli, introduction of additional information – see [24] as example. The same logic can be applied to neurodaptive systems that incorporate dynamic difficulty adjustment (DDA) [25] in order to optimise and sustain the engagement and motivation of the user (ii). In this case, task demand is adjusted dynamically upwards and downwards to stimulate and sustain investment of mental effort. The same self-regulatory function can be applied to the relationship between task demand and other psychological concepts, such as emotions and mental workload [26-27]. The third function is to widen communication bandwidth between person and computer by enabling implicit prediction and selection of user preferences (iii). The neuroadaptive mode of input control devised by Zander and colleagues [9] demonstrates how cursor movement towards a target on the screen can be directed by covert, evoked responses in EEG without any intentionality on the part of the user. This method conceptually overlaps with existing BCI methods where EEG responses are time-locked to the presentation of specific stimuli or ‘probes’, e.g. P300 speller [28], Steady-State Evoked Potentials (SSVEP) [29] – see [30] for another example of this approach. In this case, event-related neurophysiological responses are used to implicitly assess user preferences, which may reflect task-related goals or personal choice. This approach requires neither active thought nor any overt behavioural response, hence it is an excellent candidate to be used conventional modes of input control, such as a keyboard or joystick.

To assess how these three neuroadaptive functions can enhance HCI, we must define criteria to represent the quality of our interactions with technology. Table 1 presents a selected list derived from existing work [31-32]. Consideration of these criteria provides a structure within which to assess how HCI quality is potentially enhanced or jeopardised by the introduction of neuroadaptive functions.

|  |  |
| --- | --- |
| Criteria | Definition |
| Attention | Ability to keep attention of the user focused on his/her primary task |
| Calmness | Prevent the user feeling overwhelmed by incoming information |
| Context-Awareness | System can detect contextual information and proactively adapt |
| Transparency | User perception that system interaction occurs outside awareness |
| Robustness | System functions reliably across full range of operational conditions |
| Effectiveness | Performance (accuracy, speed) is enhanced by the system |
| Familiarity | Beneficial effects of system use are sustained over repeated interactions |
| Trust | User perception that the system will support goals |
| Satisfaction | User perception that system use is enjoyable and desirable |
| Privacy | Individual has the right to control handling, processing, storage and usage of personal information |
| Security | Personal data is protected from unauthorised access, malicious attacks and exploitation |

**Table 1. Selected criteria to describe the quality of human-computer interaction**

Neuroadaptive technology can enhance Attention by triggering changes at the user interface and this application is particularly important for HCI during safety-critical tasks, such as aviation [33], driving [34], medical operations [35] and defence/security applications [36] as well as the self-management of attention in augmented reality [37]. Attention can be enhanced by adapting the interface display or adjusting task demand or enabling the user to communicate preferences without intentionality. The same argument can be forwarded with respect to the Calmness criterion. Like affective computing systems, neuroadaptive technology can be configured to mitigate negative states, such as anxiety, boredom, frustration and disengagement, in ways that are dynamic and personalised to the individual. If the neuroadaptive system can effectively ‘manage’ the psychological state of the user attention, i.e., enhancing attention while sustaining calmness, then the Effectiveness of user performance will improve, e.g., fewer errors, less time required to complete task. It is also anticipated that neuroadaptive technology can improve the quality and reliability of performance in the presence of known task stressors like increased demand, sustained time-on-task, noise, fatigue etc.

User interaction with a neuroadaptive system can occur without any awareness on the part of the individual, hence the Transparency of the interaction is enhanced. A high speed of data exchange coupled with minimal awareness of this symmetrical HCI creates a design bias towards unconscious modes of interaction [38-39]. In addition, the absence of any tangible interface reinforces the transparency of the technology. Due to this inconspicuous quality of the interaction, it is important that users Trust the process of neuroadaptive control to support their activities. Adaptive changes at the neuroadaptive interface may be obvious or obscure from the perspective of the user [6]. If the neuroadaptive interface is perceived to confer conspicuous benefit, the trust of the user is enhanced [40-41].

The modus operandi of neuroadaptive functionality is to create technology that demonstrates Context-Awareness. The system creates a dynamic model of the user that informs an intelligent process of software adaptation [12], [42]. Repeated interactions with neuroadaptive technology over time should improve Familiarity in at least two senses. Users develop realistic expectations about the quality of system adaptation, which facilitate effective ‘partnership’ with the system. Secondly, repeated usage yields increased data about the individual and different task contexts, which can enhance the accuracy of the classification process; i.e., the system is able to develop a subject-dependent model and achieve greater specificity of personalisation, for further elaboration, see [12], [43].

It should be acknowledged that a technological system based on the collection of neurophysiological data collected outside of the laboratory has significant potential for system error. The closed-loop structure ensures that erroneous data captured as an input is propagated into the classification engine, leading to an incorrect output function at the interface. Even if perfect data collection and excellent classification accuracy are assumed, the neuroadaptive system is still capable of misfiring at the interface, i.e., adaptive responses can have unanticipated consequences or are perceived as ineffective or confusing or mistimed by the user. The criterion of Robustness represents the greatest single challenge to be overcome if neuroadaptive technology is to improve the quality of HCI in everyday use. This criterion encompasses both the propagation of error through imperfect data and the designing an adaptive repertoire of the system, i.e., do adaptations work as planned? Do they consistently support the goals of the user? The potential for high system error also represents the most significant threat to user Satisfaction, if the neuroadaptive system cannot perform to a reliable standard, it will not be adopted by the user population at any kind of scale.

Neuroadaptive technology is reliant on multidimensional data input from the brain and body that is continuous, quantitative and representative of unconscious as well as conscious processes [44]. A requirement to implicitly monitor the user presents multiple threats for the Privacy of the individual. The most obvious being that adaptive changes at the interface occur in a public space, where they can be observed and interpreted by co-workers and superiors. The raw data from any neuroadaptive interaction could be processed and discarded, but information about the user can still be inferred from any record of change at the adaptive interface. As we saw when considering the criterion of familiarity, there are advantages in retaining a data archive from each individual in order to improve user modelling, which opens a number of concerns around data governance and anonymity [45]. For those working in large organisations, there is also the question of who has the legitimate right to access, utilise and share those datasets accumulated from repeated interaction with neuroadaptive technology. The criterion of Security concerns the integrity and degree of data protection that is necessary to allay privacy concerns from a potential user of a neuroadaptive technology. Raw neurophysiological data will be transferred from the sensor to another entity for signal processing, such as cloud-based computing, e.g., [46], and subsequently communicated to the interface. There is the challenge of securing this protocol for communication and data transfer from unauthorised access [47-48]. If neuroadaptive data requires storage, the resulting databases must have sufficient protection to prevent malicious attacks and attempts by third parties to exploit the data.

Neuroadaptive technology has the potential to both enhance and degrade the quality of contemporary HCI. When considering the potential effects of neuroadaptive functions, we must apply a utilitarian logic that weighs potential advantages against possible disadvantages for the user. Such considerations open a debate about whether the tangible benefits of neuroadaptive technologies can outweigh the potential drawbacks of their introduction into real-world applications and everyday use. The next section will continue in this speculative vein as we consider how neuroadaptive systems encapsulate higher-level goals that shape the process of interaction with the user.

# Design of Goals for Neuroadaptive Technology

Closed-loop control is designed to achieve a specific Goal. In the case of neuroadaptive technology, the incorporation of goals into closed-loop control yields

a machine with an agenda [6], [10], [12] designed to fulfil a function (e.g., identify user preference) and proactively influence the psychological state of the user. It is important to define the goals of neuroadaptive adaptation with great care, particularly as the technology can achieve its goals with or without awareness on the part of the individual. It is equally important to ensure continuity between system goals and adaptation of the user interface.

This section is concerned with the design of system goals for neuroadaptive interfaces. It is both natural and logical that users expect the goals of the system to coincide with their own wishes, aspirations, and objectives. In addition, the high rate of data exchange between human and machine is likely to necessitate a degree of autonomous function on behalf of the system (see next section). Increasing system autonomy inevitably weakens human control, which represents an additional impetus to ensure that neuroadaptive interfaces operate in users’ best interests.

It is important that any system goal incorporated into the closed-loop of neuroadaptive technology is inherently desirable from a human perspective. Consideration of this proposition begs a fundamental question about the teleology of neuroadaptive technology [49]. In general terms, technologies can be conceived as tools to promote productivity, efficiency and effectiveness [50]. Other authors make the supplemental point that technological tools should enhance human performance in a way designed to “enfranchise, not enslave” (p.60) [49]; see also [51]. However, the definition of inherently desirable human goals within a technological framework represents a huge challenge for system design. Michel Foucault [52] made a basic distinction between technologies of production and technologies of the self. The former are tools permitting the production, manipulation and transformation of information or physical objects. Technologies of self, on the other hand, are a collection of techniques for effecting self-regulation of thoughts, actions and behaviour, the purpose of which is to achieve happiness or wisdom. According to Foucault, the ultimate purpose of any technology of the self is rooted in Greco-Roman maxims to both “know yourself” and “take care of yourself.”

Taking a cue from Foucault, neuroadaptive systems constitute a literal technology of self. Adaptive changes at the interface can be configured to promote productivity (good performance) or advance self-knowledge and self-care, or strike a humanitarian balance between both directives, which are not irreconcilable, i.e., a system designed primarily to enhance performance (reduce error) can also advance self-care as a collateral effect (reduced anxiety). It is also possible for directives towards productivity and self-care to come into conflict, e.g., for safety-critical tasks, the goal of sustaining effective performance in the face of high task demand would be prioritised over a goal to sustain a state of calmness.

A teleological analysis of neuroadaptive technology poses the question of whether “desirable” psychological states (and system goals) can be meaningfully disentangled from “undesirable” ones in this context. Periods of intense mental effort may be necessary when task demand is high, but sustained effort investment is often associated with feelings of anxiety [53]. Designing technology to exclusively promote “desirable” psychological states may be inadvisable as well as impossible. In their analysis of affective computing systems, Picard and colleagues [54] take inspiration from Huxley’s ‘Brave New World’ [55] and characterise a technology designed solely to promote positive emotional states as ‘computational soma’ (p.16). The point of their analysis being that unconditional positive affect can have a stupefying effect on the human being, and in particular situations, such as learning a new skill, ‘negative’ experiences of frustration serve as the impetus to increase motivation and achieve mastery [56]. Learning, effective performance and even self-care are not necessarily synonymous with positive psychological states and this ambiguity complicates the process of designing a technology with goals that are universally benign.

Moving from a conceptual space to practical considerations introduces a second layer of complexity, specifically how can system goals be translated into neuroadaptive functions at the user interface? The interface serves as a proxy for a high-level directive of the system and the internal coherence of the system design is dependent on consistency between those high-level goals and neuroadaptative functions at the user interface. While frameworks for the incorporation of values into the design process do exist [57], the goal of closed-loop control must be defined unambiguously in order to create a repertoire of adaptive changes at the interface that are united by a common purpose.

The hierarchical model of behavioural self-regulation described by William Powers [58] captures this process of transforming goals into actions that are differentiated by levels of abstraction. Figure 1 represents an example of this scheme applied to the design of a neuroadaptive cockpit interface, configured to enhance safety by improving attentional self-regulation. The summit of the hierarchy is a system concept formulated as a directive, e.g., to ‘do’ something, such as preserve safety or enhance self-care. This concept informs a subordinate Principle that represents a global strategy for achieving this directive. In the case of the example presented in Figure 1, safety is enhanced by a principle of sustaining attention. These Principles define discrete groupings of adaptive strategies available at the Program level. The Sequence level represents concrete instances of those strategies as specific instances of interface adaptation as experienced by the user. The number and variety of adaptive responses that can be achieved at the Sequence level is synonymous with the repertoire of adaptive responses [10], i.e., the scope and precision of the adaptive interface.

Within this hierarchical approach [59-60], the specification of each subordinate levels is informed by its superordinate. Hence, Principles are defined by the System Concept, and those Programs of adaptive strategies are aligned with these Principles. A descent through the hierarchy ensures consistency as each level both constrains and defines the range of options available to the one below. In this way, we move from abstract directives to the specification of an adaptive interface in a fashion designed to preserve continuity and consistency from top to bottom.



Figure 1: Hierarchical Design of Goals and Adaptive Responses for Hypothetical Neuroadaptive Cockpit Interface

Credit: Author

While the hierarchical approach illustrated in Figure 1 ensures a degree of internal consistency between goals and interface design, ample room remains for disagreement between users and designers on the semantics operating within this framework. A degree of dissension is a perennial issue when designing human-machine systems, but this issue is particularly problematic in the current case, given the one-to-many relationship between upper and lower levels of the hierarchy. A user-centred strategy represents one approach to ensure continuity between goals and sequences, so adaptations of the interface are formulated in a way that are intuitive from the perspective of the end-user.

The real challenge encapsulated in Figure 1 is the creation of an adaptive repertoire at the Sequence level that effectively and transparently supports the goals of the individual. Direct experience of the adaptive interface exerts an enormous influence on user perceptions of satisfaction, trust and familiarity (see Table 1). While the hierarchical relationship between goal and output is fixed within the hierarchy, there is scope for personalisation at the Program and Sequence level. With respect to the latter, the system can attempt to match adaptive responses to the preferences of the individual via a second-order process of monitoring and adaptation [12], [42], [43], [61].

In the current example of hierarchical self-regulation (Figure 1), the process of matching goals to adaptation is formulated as a top-down process. There is another option encapsulated by the probe-based approach [11] wherein the system monitors the neurophysiological response from the user to adaptive changes at the interface, which are subsequently utilised to construct a second-order model of user preferences. In principle, this bottom-up process can differentiate successful interface adaptations from unsuccessful ones by capturing event-based responses from the user. For example, Error-Related Negativity (ERN) and Error Positivity (Pe) are two evoked cortical potentials sensitive to a mismatch between expected and actual outcomes, which can be triggered by observing the response from a technological system [62]. In the context of the current discussion, adaptive responses that consistently trigger an implicit “error/mismatch” response from the user would be removed from the functional repertoire over a period of usage. By pruning the repertoire of Sequences associated with a particular Program (Figure 1), the system can personalise the interface to the individual by increasing the probability of a preferred/expected adaptive responses. The neuroadaptive model of cursor control represents a working example of this process, firstly by differentiating preferred from non-preferred directions of cursor movement via ERP analysis and subsequently increasing the probability of preferred cursor movements via reinforcement learning [9], see [30] for a second example of this approach.

Translation of a goal or directive into a specific function at the neuroadaptive interface is a complex process for the designer and riddled with the potential for mismatch. One major problem concerns the semantics of the hierarchical process shown in Figure 1, particularly preserving continuity between the system concept and those sequences experienced at the interface. At a socio-technical level, we have the additional complication of creating hierarchical systems of goal regulation that genuinely support the needs of the user, whether these are defined with respect to productivity or self-care and with respect to who defines the primary directive of the system, e.g., employers or employees.

A bottom-up approach presents the possibility of creating what is probably the ultimate form of user-centred design, a system that can adapt to the needs of a particular individual operating in a specific context. However, there are reasons for healthy cynicism about the robustness of this approach outside of the laboratory. For the final section of this chapter, we will consider the interaction between user and system in greater detail at the Sequence level, specifically with respect to autonomous functions and the ability of the user to preserve agency at the neuroadaptive interface.

# Neuroadaptive Autonomy: Striking a Balance

Interaction with neuroadaptive technology is achieved via a continuous stream of neurophysiological data collected without the conscious awareness of the user. This implicit channel of communication tends to favour autonomous interaction, where software adapts without any input from the user. A design bias towards increased system autonomy may reflect the influence of active BCI, where output is a substitution for an effector like a hand or a foot and must be activated automatically for a seamless experience of input control. However, neuroadaptive technology encompasses greater scope for intelligent software adaptation [10], both with respect to range of available function and mechanism of delivery at the interface.

As described in Section 2, neuroadaptive systems enable three broad classes of function - adaptive information presentation, adaptive demand adjustment and implicit prediction of user preference. The predicament for the interface designer across all three functions is how to strike a balance between high-speed, symmetrical HCI while preserving the autonomy of the user. This dilemma goes deeper than simply providing a mechanism for the user to retain control. A system that combines high rates of information exchange (between human and system) with autonomous function effectively blurs the distinction between user and machine. Actions at the interface are no longer triggered by discrete volitional actions directly attributed to a person, neuroadaptive ‘actions’ are derived from an amalgamation of neurophysiological monitoring, statistical modelling and machine logic. The actor at the heart of a neuroadaptive interaction is a quantified symbiosis of information technology and ‘live’ neurophysiological data as opposed to an individual human agent. Within this hybrid entity [63], goals, performance and liabilities are merged between person and machine, to the point where they are impossible to disentangle in practical terms, see [12] for further discussion.

There is a long tradition in human factors research where implicit signals from the brain and body are utilised in order to enable adaptive modes of automation [44]. Furthermore, human factors research has produced a graded analysis of system automation that is pertinent to a consideration of how neuroadaptive interfaces should be implemented. The Levels of Automation (LOA) analysis [64] presents an analysis of system automation wherein the degree of automation varies along a continuum from manual to fully automated functions. LOA provides with a framework within which to consider the full range of design options available for any neuroadaptive interface. Lower levels of system automation (LO3-4) are concerned with filtering information to support attentional regulation. This category of automation pertains to the neuroadaptive function (i) described in Section 1. Adapting the provision of information at the interface is a reversible category of interface adaptation, i.e., the user can revert from the adapted to the original modality or format of information presentation. The two other neuroadaptive functions (ii) and (iii) from Section 1 are both action-based, they represent an adjustment of task demand or the selection of a preferred item. This distinction is important because the potential for negative consequences increases for irreversible adaptive functions as opposed to revokable adjustments of information provision at the interface [65].

Neuroadaptive functions can be executed with (LOA5) or without (LOA7) approval from the user. In the case of the former, the user requires feedback of a pending action and a means to grant approval. LOA6 represents a midpoint between both options where the function is executed, albeit with a limited time window to enable a manual veto. In all examples discussed so far, the system provides the user with explicit feedback that a neuroadaptive function has been activated. This feedback may be tacit, e.g., changes to the display format, or explicit involving an additional display element, e.g., an icon indicating the current format of the display. At LOA8, functions are executed autonomously and feedback must be requested by the user, otherwise it is not provided. At LOA9, the provision of feedback remains at the discretion of the system and LOA10 represents full automation without any possibility of feedback to the user; see Table I (p. 287) of [64] for full definitions.

There are use cases for neuroadaptive technology where optimising HCI may only be achieved by enabling higher levels of automation at LOA8 and upwards. If the system goal is to promote states of challenging, skilful cognitive performance via adjustment of task demand, then any requirement for the user to provide manual approval would be counterproductive. Similarly, if we design a system to expand communication bandwidth via an implicit neurophysiological channel, the latter must function autonomously to be effective. There are also cases where it is appropriate to explicitly request confirmatory input from the user (LOA5). Given significant concerns related to the robustness of neuroadaptive systems ‘in the wild’ (see Table 1 and Section 2), it is advisable to incorporate user confirmation into neuroadaptive interfaces when the task is safety critical. Therefore, a pilot or driver can revert to conventional modes of information provision or even disable neuroadaptive functionality if data quality is compromised and the interface cannot function as it was designed. For those neuroadaptive systems designed to identify preferred items from an array, seeking manual confirmation prior to selection would allow the user to identify and prevent erroneous decisions made by the system. The resulting data provides a useful metric of system effectiveness and could be incorporated into training data for a classification algorithm, see [66-67] as examples of this approach. It is important that neuroadaptive interfaces incorporate a level of automation that is suitable for: (1) application domain (e.g., safety-critical or not), (2) type of functionality (information provision vs. activation of function), and (3) the predicted robustness of closed-loop in a particular setting.

Given that neuroadaptive systems generally incorporate classification algorithms based on machine learning methods, it is logical that classification accuracy may be low or unstable if only a limited set of training data is available; this is particularly an issue for systems that utilise subject-dependent learning, i.e., when data are specific to a single user. In these cases, when an algorithm is still in the process of acquiring the necessary corpus of training data, automation could be capped to LOA5 and all adaptive functions require user approval until the algorithm is optimised. With repeated interaction, functions are executed but with an associated time delay to enable a manual veto from the user (LOA6), provided the delay does not impede effective interaction. After a long and sustained period of usage, when classification accuracy is high and stable, and the user is sufficiently familiar with the adaptive interface, the system is permitted to move to LOA8 and above. A cumulative increase of LOA with experience provides an opportunity for user trust to accurately reflect the robustness of the system (Table 1, see Section 2), both of which should improve with repeated use.

This graded and longitudinal approach to autonomous function allows us to design neuroadaptive interfaces that are sensitive to the context of the usage scenario. By permitting inputs from the user, it is possible to improve the robustness of the technology and actively cultivate trust in the adaptive process. There is a danger of ‘confirmation fatigue’ given the high rate of information exchange between human and machine, but this shortcoming can be mitigated by careful design of feedback at the interface.

# Summary

Neuroadaptive interfaces have been described with respect to three core functions: adaptive information presentation, adaptive difficulty adjustment, and identification of a preferred category of response. The value and utility of these functions is based on the potential of closed-loop control to imbue technology with a capacity for context-awareness. By using neurophysiological signals as a continuous and implicit channel of communication, neuroadaptive systems can support performance and self-care by dynamically adapting the interface to select specific options autonomously or cultivate some psychological states while mitigating others. The utility of neuroadaptive technology is jeopardised by a significant potential for error when utilised in the real-world, from data collection to the process of inference that informs the adaptive interface. The dubious robustness of these systems in the field and concerns around data privacy and security represent significant risks for large-scale adoption of this emergent technology.

The closed-loop model for neuroadaptive technology is a blueprint for a machine with an agenda. In other words, neuroadaptive systems are designed to achieve a high-level goal. Defining and designing goals that can function within a quantitative model of closed-loop control presents a number of challenges, especially with respect to consistency, coherence and effectiveness. However, the formalisation of system goals is important because the designer must give serious thought to the raison d'être of the neuroadaptive system. Research on designing technology to comply with high-level directives based on human goals is a developing field of investigation, e.g., [68]. In practical terms, neuroadaptive interfaces that fail to support the goals and objectives of the user in a transparent fashion will not add value or utility from a human perspective.

There is a school of thought that high-speed communication and enhanced human-computer bandwidth are the primary benefits to the user offered by neuroadaptive technology. This emphasis on speed of communication represents an argument for neuroadaptive interfaces that function autonomously without imposing upon human attention. However, there are occasions when human intervention in neuroadaptive functions is both desirable and beneficial for the quality of the human-computer interaction and to preserve human agency. By designing neuroadaptive functions with graded levels of automation that permit manual inputs, system robustness can be improved.

The next phase of research on neuroadaptive technology must embrace the design of the entire system, across the closed-loop, from sensors and signal processing to the experience of the user at the interface. The latter has been under-researched in comparison to the former, but the neuroadaptive interface is the locus of this technology where any utility and value will be experienced by the user.

ACKNOWLEDGEMENT

The author would like to acknowledge the comments and feedback received from Alexander Karran on an earlier version of this draft.

**[Chapter References]**

[1] L. J. Hettinger, P. Branco, L. M. Encarnaco, and P. Bonato, “Neuroadaptive technologies: applying neuroergonomics to the design of advanced interfaces,” *Theor. Issues Ergon. Sci.*, vol. 4, no. 1–2, pp. 220–237, 2003.

[2] A. T. Pope, E. H. Bogart, and D. S. Bartolome, “Biocybernetic system evaluates indices of operator engagement in automated task,” *Biol. Psychol.*, vol. 40, pp. 187–195, 1995.

[3] L. J. Prinzel, M. W. Scerbo, F. G. Freeman, and P. J. Mikulka, “A bio-cybernetic system for adaptive automation,” in *Human Factors and Ergonomics Society 39th Annual Meeting*, 1995, pp. 1365–1369.

[4] M. W. Scerbo, F. G. Freeman, and P. J. Mikulka, “A brain-based system for adaptive automation,” *Theor. Issues Ergon. Sci.*, vol. 4, no. 1–2, pp. 200–219, 2003.

[5] J. Allanson and S. H. Fairclough, “A research agenda for physiological computing,” *Interact. Comput.*, vol. 16, no. 5, 2004, doi: 10.1016/j.intcom.2004.08.001.

[6] S. H. Fairclough, “Fundamentals of physiological computing,” *Interact. Comput.*, vol. 21, no. 1–2, 2009, doi: 10.1016/j.intcom.2008.10.011.

[7] S. H. Fairclough and K. Gilleade, *Advances in Physiological Computing*. Springer, 2014.

[8] T. O. Zander and C. Kothe, “Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general,” *J. Neural Eng.*, vol. 8, pp. 1–5, 2011.

[9] T. O. Zander, L. R. Krol, N. P. Birbaumer, and K. Gramann, “Neuroadaptive technology enables implicit cursor control based on medial prefrontal cortex activity,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 52, pp. 14898–14903, 2016, doi: 10.1073/pnas.1605155114.

[10] S. H. Fairclough, “Physiological computing and intelligent adaptation.,” in *Emotions and affect in human factors and human-computer interaction.*, San Diego, CA, US: Elsevier Academic Press, 2017, pp. 539–556.

[11] L. R. Krol, P. Haselager, and T. O. Zander, “Cognitive and affective probing: a tutorial and review of active learning for neuroadaptive technology,” *J. Neural Eng.*, vol. 17, no. 1, p. 12001, Jan. 2020, doi: 10.1088/1741-2552/ab5bb5.

[12] S. Fairclough, *A Closed-Loop Perspective on Symbiotic Human-Computer Interaction*, vol. 9359. 2015.

[13] V. Venkatesh and F. D. Davis, “A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies,” *Manage. Sci.*, vol. 46, no. 2, pp. 186–204, 2000, doi: 10.1287/mnsc.46.2.186.11926.

[14] B. J. Babin, W. R. Darden, and M. Griffin, “Work and/or Fun: Measuring Hedonic and Utilitarian Shopping Value,” *J. Consum. Res.*, vol. 20, no. 4, pp. 644–656, 1994.

[15] R. Sánchez-Fernández and M. Á. Iniesta-Bonillo, “The concept of perceived value: A systematic review of the research,” *Mark. Theory*, vol. 7, no. 4, pp. 427–451, 2007, doi: 10.1177/1470593107083165.

[16] B. Allison, B. Graimann, and A. Graser, “Why use a BCI if you are healthy?,” in *ACE Workshop - Brainplay’07: Brain-Computer Interfaces and Games*, 2007.

[17] N. Wiener, *Cybernetics: Control and communication in the animal and the machine*, 2nd ed. Cambridge, Mass.: M.I.T. Press, 1948.

[18] D. Novak, M. Mihelj, and M. Munih, “A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing,” *Interact. Comput.*, vol. 25, pp. 154–172, 2012.

[19] P. J. Bota, C. Wang, A. L. N. Fred, and H. Plácido Da Silva, “A Review, Current Challenges, and Future Possibilities on Emotion Recognition Using Machine Learning and Physiological Signals,” *IEEE Access*, vol. 7, pp. 140990–141020, 2019.

[20] L. J. Prinzel, A. T. Pope, and F. G. Freeman, “Physiological self-regulation and adaptive automation,” *Int. J. Aviat. Psychol.*, vol. 12, no. 2, pp. 179–196, 2002.

[21] S. E. Petersen and M. I. Posner, “The attention system of the human brain: 20 years after,” *Annu. Rev. Neurosci.*, vol. 35, pp. 73–89, Jul. 2012, doi: 10.1146/annurev-neuro-062111-150525.

[22] M. Esterman and D. Rothlein, “Models of sustained attention,” *Curr. Opin. Psychol.*, vol. 29, pp. 174–180, 2019, doi: https://doi.org/10.1016/j.copsyc.2019.03.005.

[23] M. Corbetta, G. Patel, and G. L. Shulman, “The Reorienting System of the Human Brain: From Environment to Theory of Mind,” *Neuron*, vol. 58, no. 3, pp. 306–324, 2008, doi: 10.1016/j.neuron.2008.04.017.

[24] J. Müller and A. Schulte, “Concept of an Adaptive Cockpit to Maintain the Workflow of the Cockpit Crew,” in *Intelligent Human Systems Integration 2020*, 2020, pp. 952–958.

[25] R. Hunicke, “The case for dynamic difficulty adjustment in games,” *ACM Int. Conf. Proceeding Ser.*, vol. 265, pp. 429–433, 2005, doi: 10.1145/1178477.1178573.

[26] S. K. Ehrlich, K. R. Agres, C. Guan, and G. Cheng, “A closed-loop, music-based brain-computer interface for emotion mediation,” *PLoS One*, vol. 14, no. 3, pp. 1–24, 2019, doi: 10.1371/journal.pone.0213516.

[27] M. Chaouachi and C. Frasson, “Mental Workload, Engagement and Emotions: An Exploratory Study for Intelligent Tutoring Systems,” in *Intelligent Tutoring Systems*, 2012, pp. 65–71.

[28] E. W. Sellers, A. Kubler, and E. Donchin, “Brain-computer interface research at the university of south Florida cognitive psychophysiology laboratory: the P300 speller,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 221–224, Jun. 2006, doi: 10.1109/TNSRE.2006.875580.

[29] C. Guger *et al.*, “How Many People Could Use an SSVEP BCI?,” *Front. Neurosci.*, vol. 6, p. 169, 2012, doi: 10.3389/fnins.2012.00169.

[30] L. Kangassalo, M. Spapé, and T. Ruotsalo, “Neuroadaptive modelling for generating images matching perceptual categories,” *Sci. Rep.*, vol. 10, no. 1, p. 14719, 2020, doi: 10.1038/s41598-020-71287-1.

[31] R. M. Carvalho, R. M. de Castro Andrade, K. M. de Oliveira, I. de Sousa Santos, and C. I. M. Bezerra, “Quality characteristics and measures for human–computer interaction evaluation in ubiquitous systems,” *Softw. Qual. J.*, vol. 25, no. 3, pp. 743–795, 2017, doi: 10.1007/s11219-016-9320-z.

[32] ISO/IEC 25000., “Software Engineering—Software Product Quality Requirements and Evaluation (SQuaRE)—Guide to SQuaRE.,” 2014.

[33] F. Dehais, M. Causse, F. Vachon, N. Régis, E. Menant, and S. Tremblay, “Failure to Detect Critical Auditory Alerts in the Cockpit: Evidence for Inattentional Deafness,” *Hum. Factors*, vol. 56, no. 4, pp. 631–644, 2014, doi: 10.1177/0018720813510735.

[34] K. Kircher and C. Ahlstrom, “Minimum Required Attention: A Human-Centered Approach to Driver Inattention,” *Hum. Factors*, vol. 59, no. 3, pp. 471–484, 2017, doi: 10.1177/0018720816672756.

[35] H. N. Modi, H. Singh, A. Darzi, and D. R. Leff, “Multitasking and Time Pressure in the Operating Room: Impact on Surgeons’ Brain Function,” *Ann. Surg.*, vol. 272, no. 4, 2020.

[36] S. Chiappino, L. Marcenaro, and C. Regazzoni, “Selective attention automatic focus for cognitive crowd monitoring,” in *2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance*, 2013, pp. 13–18, doi: 10.1109/AVSS.2013.6636609.

[37] R. McKendrick, “Chapter 18 - Mobile Neuroergonomics: Action, Interfaces, Cognitive Load, and Selective Attention,” in *Neuroergonomics*, H. Ayaz and F. Dehais, Eds. Academic Press, 2019, pp. 111–116.

[38] K. Hoshi and J. Waterworth, “Unconscious Interaction and Design,” in *Primitive Interaction Design*, Cham: Springer International Publishing, 2020, pp. 75–86.

[39] P. Negri, L. Gamberini, and S. Cutini, “A Review of the Research on Subliminal Techniques for Implicit Interaction in Symbiotic Systems,” in *Symbiotic Interaction*, 2014, pp. 47–58.

[40] J. D. Lee and K. A. See, “Trust in automation: designing for appropriate reliance,” *Hum. Factors*, vol. 46, no. 1, pp. 50–80, 2004.

[41] E. J. de Visser, R. Pak, and T. H. Shaw, “From ‘automation’ to ‘autonomy’: the importance of trust repair in human–machine interaction,” *Ergonomics*, vol. 0139, pp. 1–19, 2018, doi: 10.1080/00140139.2018.1457725.

[42] N. Serbedzija and S. Fairclough, “Reflective pervasive systems,” *ACM Trans. Auton. Adapt. Syst.*, vol. 7, no. 1, 2012, doi: 10.1145/2168260.2168272.

[43] E. A. Kirchner, S. H. Fairclough, and F. Kirchner, “Embedded Multimodal Interfaces in Robotics: Applications, Future Trends, and Societal Implications,” in *The Handbook of Multimodal-Multisensor Interfaces: Language Processing, Software, Commercialization, and Emerging Directions - Volume 3*, Association for Computing Machinery and Morgan &amp; Claypool, 2019, pp. 523–576.

[44] E. Byrne and R. Parasuraman, “Psychophysiology and adaptive automation,” *Biol. Psychol.*, no. 42, pp. 249–268, 1996.

[45] S. Fairclough, “Physiological data must remain confidential,” *Nature*, vol. 505, no. 7483, 2014, doi: 10.1038/505263a.

[46] J. K. Zao *et al.*, “Pervasive brain monitoring and data sharing based on multi-tier distributed computing and linked data technology,” *Front. Hum. Neurosci.*, vol. 8, p. 370, 2014, doi: 10.3389/fnhum.2014.00370.

[47] M. Ienca and P. Haselager, “Hacking the brain: brain–computer interfacing technology and the ethics of neurosecurity,” *Ethics Inf. Technol.*, vol. 18, no. 2, pp. 117–129, 2016, doi: 10.1007/s10676-016-9398-9.

[48] P. Kellmeyer, “Big Brain Data: On the Responsible Use of Brain Data from Clinical and Consumer-Directed Neurotechnological Devices,” *Neuroethics*, 2018, doi: 10.1007/s12152-018-9371-x.

[49] P. A. Hancock, *Mind, Machine and Morality: Towards a Philosophy of Human-Technology Symbiosis*. Ashgate, 2009.

[50] W. B. Arthur, *The Nature of Technology: What It Is and How It Evolves*. Penguin, 2009.

[51] N. Wiener, *The Human Use Of Human Beings: Cybernetics & Society*. Boston: Da Capo Press, 1954.

[52] M. Foucault, *Technologies of the Self: A Seminar with Michel Foucault*. University of Massachusetts Press, 1988.

[53] G. R. J. Hockey, “Compensatory control in the regulation of human performance under stress and high workload: a cognitive-energetical framework.,” *Biol. Psychol.*, vol. 45, pp. 73–93, 1997.

[54] J. Klein, Y. Moon, and R. W. Picard, “This computer responds to user frustration: theory, design and results,” *Interact. Comput.*, vol. 14, pp. 119–140, 2002.

[55] A. Huxley, *Brave New World*. Harmondworth, England: Penguin, 1932.

[56] R. W. Picard *et al.*, “Affective learning - a manifesto,” *BT Technol. J.*, vol. 22, no. 4, pp. 253–269, 2004.

[57] B. Friedman, P. H. Kahn Jr., and A. Borning, “Value Sensitive Design and Information Systems,” *Human-Computer Interact. Manag. Inf. Syst. Found.*, pp. 1–27, 2006, doi: 10.1145/242485.242493.

[58] W. T. Powers, “A systems approach to consciousness,” in *Psychobiology of Consciousness*, J. M. Davidson and R. J. Davidson, Eds. New York: Plenum, 1980, pp. 217–242.

[59] W. T. Powers, “Feedback: Beyond behaviour,” *Science (80-. ).*, vol. 179, pp. 351–356, 1973.

[60] C. S. Carver and M. F. Scheier, *On the Self-Regulation of Behaviour*. Cambridge, UK: Cambridge University Press, 1998.

[61] N. B. Serbedzija and S. H. Fairclough, “Biocybernetic loop: From awareness to evolution,” in *2009 IEEE Congress on Evolutionary Computation, CEC 2009*, 2009, doi: 10.1109/CEC.2009.4983195.

[62] E. J. de Visser *et al.*, “Learning From the Slips of Others: Neural Correlates of Trust in Automated Agents,” *Front. Hum. Neurosci.*, vol. 12, p. 309, 2018, doi: 10.3389/fnhum.2018.00309.

[63] P. A. Hancock and M. H. Chignell, “Adaptive control in human-machine systems,” in *Human Factors Psychology*, P. A. Hancock, Ed. Amsterdam: North-Holland, 1987, pp. 305–345.

[64] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, “A model for types and levels of human interaction with automation,” *IEEE Trans. Syst. Man, Cybern. - Part A Syst. Humans*, vol. 30, no. 3, pp. 286–297, May 2000, doi: 10.1109/3468.844354.

[65] L. Onnasch, C. D. Wickens, H. Li, and D. Manzey, “Human performance consequences of stages and levels of automation: An integrated meta-analysis,” *Hum. Factors*, vol. 56, no. 3, pp. 476–488, 2014, doi: 10.1177/0018720813501549.

[66] N. Kosmyna, F. Tarpin-Bernard, and B. Rivet, “Adding human learning in Brain-Computer Interfaces (BCIs): Towards a practical control modality,” *ACM Trans. Comput. Interact.*, vol. 22, no. 3, pp. 1–37, 2015, doi: 10.1145/2723162.

[67] S. H. Fairclough, A. J. Karran, and K. Gilleade, “Classification accuracy from the perspective of the user: Real-time interaction with physiological computing,” in *Conference on Human Factors in Computing Systems - Proceedings*, 2015, vol. 2015-April, doi: 10.1145/2702123.2702454.

[68] M. O. Riedl, “Human-centered artificial intelligence and machine learning,” *Hum. Behav. Emerg. Technol.*, vol. 1, no. 1, pp. 33–36, 2019, doi: 10.1002/hbe2.117.