Philosophy & Technology (2023) 36:30
https://doi.org/10.1007/s13347-023-00636-5

RESEARCH ARTICLE

Neuroadaptive Technology and the Self: a Postphenomenological Perspective

Stephen Fairclough[1]

Received: 18 November 2022 / Accepted: 17 April 2023 / Published online: 26 April 2023

Abstract
Neuroadaptive technology (NAT) is a closed-loop neurotechnology designed to enhance human–computer interaction. NAT works by collecting neurophysiological data, which are analysed via autonomous algorithms to create actions and adaptations at the user interface. This paper concerns how interaction with NAT can mediate self-related processing (SRP), such as self-awareness, self-knowledge, and agency. We begin with a postphenomenological analysis of the NAT closed loop to highlight the built-in selectivities of machine hermeneutics, i.e., autonomous chains of algorithms that convert data into an assessment of psychological states/intentions. We argue that these algorithms produce an assessment of lived experience that is quantitative, reductive, and highly simplistic. This reductive assessment of lived experience is presented to the user via feedback at the NAT interface and subsequently mediates SRP. It is argued that congruence between system feedback and SRP determines the precise character of the alterity relation between human user and system. If feedback confirms SRP, the technology is regarded as a quasi-self. If there is a disagreement between SRP and feedback from the system, NAT is perceived to be a quasi-other. We argue that the design of the user interface shapes the precise ways in which NAT can mediate SRP.

Keywords Neuroadaptive technology · Agency · Self-referential processing · Algorithms · Postphenomenology · Alterity

Abbreviations

BCI	brain–computer interface
EEG	electroenceplogram
HCI	human–computer interaction
NAT	neuroadaptive technology
pBCI	passive brain–computer interface
SRP	self-referential processing

Stephen Fairclough
s.fairclough@ljmu.ac.uk

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

1 Introduction

Neuroadaptive technology (NAT) is an emergent form of human-computer interaction (HCI) where implicit signals from the brain and body are utilised to trigger commands or adapt software in real-time (Fairclough & Zander, 2022). NAT represents an extension of existing concepts such as passive brain–computer interfaces (pBCI) (Zander & Kothe, 2011) and physiological computing (Allanson & Fairclough, 2004; Fairclough, 2009; Fairclough & Gilleade, 2014). The NAT concept describes how the state and/or intention of a healthy user is quantified via neurophysiology to enable an implicit form of HCI via closed-loop control (Fairclough, 2017; Krol & Zander, 2022; Pope et al., 1995).

NAT represents an example of context-aware computing (Dey, 2018), designed to enhance HCI by permitting implicit control and promoting psychological states associated with effective performance, enjoyment, and health (Fairclough, 2022). The neuroadaptive model of cursor control described by Zander et al., (2016) represents an example of the former. In this case, users direct the cursor towards a target location based on implicit responses in the electroencephalogram (EEG) and, more importantly, without any requirement to exercise volitional control; see Kangassalo et., (2020) for another application of the same approach.

Like earlier incarnations of closed-loop control that have integrated real-time physiological monitoring (Fairclough 2009; Scerbo et al., 2003), NAT can also adapt the user interface in ways designed to shape the psychological state of the user. For instance, excessive levels of mental workload can be detected via neurophysiology (Gateau et al., 2018) and used to trigger technological interventions that aid the overloaded operator, such as intelligent aiding and adaptive automation (Brand & Schulte, 2021). NAT can also promote specific psychological states associated with good performance (e.g., task engagement, active learning, enjoyment, relaxation) or to mitigate states known to degrade performance (e.g., high workload, fatigue, anxiety, frustration). Recent applications of NAT for the regulation of psychological states include computer games (Fairclough et al., 2021), training in virtual reality (Dey et al., 2019), educational software (Walter et al., 2017), neurorehabilitation (Leamy et al., 2014), and automated systems (Di Flumeri et al., 2019).

Regardless of whether NAT is designed to enable implicit control or shape the psychological status of the user, the exchange of information between human and technology within the HCI can be described as symmetrical (Hettinger et al., 2003). In other words, users' existing ability to interrogate the operational status of a machine, e.g., available RAM and network activity, is mirrored by a technological capacity to monitor and assess intentions or the cognitive/emotional status of the user. The resulting exchange of information creates an implicit channel of communication between person and machine (Krol & Zander, 2022).

This symmetrical exchange creates a hybrid entity that is constructed in real-time wherein human data is amalgamated with algorithmic analysis (Dorrestijn, 2017; Froese, 2014; Verbeek, 2008), the outputs of which are conveyed as feedback at the user interface. This feedback allows the user to understand how his current psychological state, intention, or preference is currently appraised by the system. For example, if the NAT interface is designed to automatically offer help information when the person is frustrated, the user may reasonably interpret the appearance of help information as an indication that frustration was detected by the technology. Similarly, if the system automatically moves the cursor in a specific direction, this feedback conveys a technological interpretation of the user's current intention.

Real-time feedback at the NAT interface can influence self-referential processing (SRP), such as the user's sense of agency (Moore, 2016; Steinert et al., 2019) and self-awareness/self-knowledge (Li et al., 2010; Rapp, 2021; Rapp & Tirassa, 2017; Sharon, 2017). This technological mediation of SRP raises a question of whether feedback from the NAT interface supplements or subverts our human capacity for SRP. At one extreme, the freedom of the person to act with full autonomy is usurped by a NAT that automatically and pre-emptively translates implicit intentions into control actions. It could be argued that this kind of proactive technological intervention threatens human autonomy by "co-opting the flexibility that is human birth right" (Weiner 1950. p.105) and risks reducing the user to the status of a "human keypad" (Floridi, 2019) (p.382). Similarly, feedback from a NAT interface can distort self-awareness and self-regulation by biasing perceptions of cognitive capacity and emotional states. NAT has the potential to augment human cognition by creating efficient communication and promoting positive psychological states, but these benefits are only made possible by humans ceding part of their capacity for independent action and independent thought to a technological device (Fairclough, 2015).

This paper will analyse how interaction with NAT can mediate SRP, specifically perceptions of self and agency. We argue that NAT is based on a quantification of self that differs from lived experience in several significant ways. Therefore, feedback from the system will inevitably fall on a continuum between agreement and disagreement with SRP. Using a postphenomenological framework, we analyse the technological intentionality of NAT with a particular emphasis on the instrumental selectivities of autonomous algorithms and the dyadic relationship between user and interface. We argue that the degree of discrepancy between system feedback and SRP determines the precise type of alterity relation between person and machine, specifically whether users perceive NAT as a quasi-self or quasi-other. This dichotomy has a profound influence on how SRP are mediated during any interaction with NAT. It is concluded that the potential of NAT to supplement or subvert SRP is dependent on the design of the user interface, which in turn, determines the precise character of the Alterity relation.

The remainder of the paper is structured in three sections. The first will analyse processes within the NAT closed loop from a postphenomenological perspective (Ihde, 1990; Rosenberger & Verbeek, 2015). The second section presents a comparison between lived experience and the quantification of lived experience by technology. The final section will consider the alterity relation and how SRP are mediated by interaction with NAT.

2 A Postphenomelogical Analysis of Neuroadaptive Technology

This postphenomenological analysis will consider the different dimensions of technological intentionality exhibited by NAT. Technological intentionality describes how technological artifacts actively mediate the relationship between human consciousness and the world (Mykhailov, 2020; Ihde, 1990; Rosenberger & Verbeek, 2015); it has been argued that interaction with technology shapes the ways in which reality becomes meaningful for a person (Verbeek, 2005).

Analysing technological intentionality is complicated by the fact that NAT is not a single technological artifact but manifests as series of technological elements that are interconnected in a closed-loop configuration. Specifically, neurophysiological data are captured from wearable sensors (Shi et al., 2020), and the resulting raw data is processed by signal processing algorithms to increase the signal-to-noise ratio, e.g., Mumtaz, Rasheed, and Irfan (2021). Inferences from these "cleaned" data are derived from machine learning algorithms (Appriou et al., 2020), which trigger action or adaptation at the user interface. The functions of each technological element (sensors, algorithms, adaptive software) are stabilised by the closed-loop configuration of the system (Verbeek, 2005). Therefore, we should regard NAT as an ecosystem where technological elements are directed towards other technological elements in a sequential and recursive formation.

The various technological elements within NAT encompass a series of processes, e.g., data collection, signal processing, and machine learning, which are concealed from the user. These unseen processes constitute what Husserl would have termed the inner horizon of the NAT system (Mykhailov & Liberati, 2022). This inner horizon can only be experienced by the person indirectly when they interact with the user interface; therefore, feedback from the NAT interface is the primary conduit for any technological mediation of SRP. Furthermore, feedback from the NAT interface is mutable and dynamic, fluctuating in real-time with each iteration of the closed loop.

It should also be noted that NAT functions as a machine with an agenda (Fairclough, 2015; Fairclough, 2021). In other words, these systems are designed to serve a prescribed goal, e.g., reduce frustration, identify preference, and sustain task engagement, and each system incorporates a repertoire of adaptive responses that are dedicated to this system goal. The co-existence of human and machine goals creates a type of hybrid intentionality (Verbeek, 2008), where interaction between person and NAT yields a new entity with a hybrid intentionality. In his 2008 analysis, Verbeek described various forms of cyborg intentionality where technological artifacts are implanted into the human body. While this is a design option for NAT, most of the systems currently under discussion are designed to work with wearable/removable sensors. Therefore, any merging between human and technological intentionality is flexible rather than permanent (Rapp, 2021), and this hybrid intentionality manifests as an exchange of data in the realm of the infosphere (Floridi, 2007, 2014).

Interaction with NAT is most accurately represented by a specific category of hybrid intentionality called immersion (Rosenberger & Verbeek, 2015). Their usage of immersion describes a category of intentionality that is bi-directional and reflexive, which is completely unrelated to immersion as studied in HCI research (Cummings & Bailenson, 2016; Jennett et al., 2008). The bidirectionality of NAT is captured by the fact that humans are working with a technology that is simultaneously directed towards themselves, e.g., detection of psychological states and classification of preferences. Also, the presentation of feedback at the NAT interface is reflexive (Serbedzija & Fairclough, 2012), i.e., humans perceive their lived experience from the perspective of those processes of data collection, analysis, and inference at work within the inner horizon of the technology.

Therefore, the type of technological intentionality exhibited by the NAT ecosystem actively mediates SRP via feedback that reflects bidirectional communication and is self-referential. The precise nature of this mediation depends on the degree of association between lived experience (particularly the experience of self via introspective and interoception) and the quantification of that lived experience produced by the system. NAT creates a real-time representation of the self that is quantified via a "haze of technological activity" (p. 617) (Mykhailov, 2020), which incorporates various forms of technological intentionality encompassed within a connected ecosystem of technological elements.

These technological elements actively mediate SRP via different dynamics. A significant proportion of the NAT ecosystem is devoted to sensor technologies and associated algorithms for signal processing and data analysis. This sensing apparatus is configured to process specific categories of data in particular ways (Ihde, 1990; Wiltse, 2014). There are a number of instrumental intentionalities and built-in selectivities associated with sensors and algorithms, which can reveal or conceal different aspects of SRP in the existential dimension and magnify or reduce facets of SRP in the epistemological dimension (Kiran, 2015). For example, wearable sensors are manufactured to capture a specific type of neurophysiological activity (e.g., electrocortical, neurovascular) and designed to collect data features from defined cortical locations via the montage/configuration of sensors (e.g., prefrontal, parietal, temporal). This selectivity is necessary to target a particular aspect of lived experience, e.g., mental workload, frustration, engagement, and fatigue, which determines the types of feedback received at the user interface. This selectivity of design determines which aspects of lived experience are revealed/magnified at the expense of others. With respect to an existential dimension, feedback from the user interface focuses on one aspect and conceals other aspects of lived experience; the availability of feedback on one aspect of lived experience can also occlude and contaminate any attempt to perceive that dimension of lived experience in a non-technologically mediated way. From an epistemological perspective, system feedback that fluctuates in real-time magnifies an awareness of a specific aspect of lived experience by increasing the fidelity and objectivity of self-referential information available to the person, but only with reference to that specific aspect.

Both existential and epistemological forms of technological mediation are predetermined by the specifics of those algorithms incorporated into the NAT ecosystem. These algorithms are concerned with the transformation, quantification, and interpretation of human experience, which can be characterised as a hermeneutic relation (Ihde, 1990). All hermeneutic relations within NAT are governed by autonomous algorithms, each of which express a particular type of instrumental intentionality by actively interpreting data within a prescriptive way (Wellner, 2020; Wiltse, 2014). Traditionally, hermeneutic relations describe a translational process between person and technology, but Luan (2020) recently proposed the category of machine hermeneutics wherein mediation occurs directly between technological elements. These technology-technology relations describe how outputs from one algorithm are relayed to a second algorithm for subsequent interpretation (Luan, 2020) and are particularly relevant to the NAT closed loop. For example, if EEG data from a wearable sensor is filtered and subjected to artifact correction, this process represents one form of machine hermeneutic, i.e., a technological element (signal analysis) operates upon data from another technological element (sensor) that is directed towards the user. These technology–technology relations are a common feature within the NAT closed loop as data processing runs as an autonomous pipeline, and therefore, built-in selectivity is an inevitable feature of technology. The feedback received at the user interface reveals and magnifies a selective quantification of lived experience assembled from a chain of machine hermeneutics (Just & Latzer, 2016; Luan, 2020).

A combination of bidirectional communication, autonomous algorithmic processing, and reflexive feedback allows the NAT user interface to function as a self-sustaining technological actor. The user communicates implicitly with this technological actor via implicit data monitoring, which provokes an adaptive response from the NAT interface. This symmetrical mode of human-computer interaction (Hettinger et al., 2003) creates a dyadic dialogue that mimics key aspects of human-human communication, such as assessment, reciprocity, and action, creating an Alterity relation (Ihde, 1975, 1990). The fact that NAT can perceive and respond autonomously to user states/intentions in ways that are both timely and appropriate (Fairclough, 2021) creates a dialogue that allows the user interface to function as a technological actor.

This Alterity relation mediates human perception of SRP along practical and ethical dimensions (Kiran, 2015). With respect to the former, the NAT interface enables the person to communicate their state and intentions to technology without any overt behaviour or volitional intention. By interpreting these data, the NAT system can adapt software or execute a control action at the user interface. As stated in Section 1, this type of autonomous adaptation can support self-regulation (e.g., enhancing engagement, promoting enjoyment, mitigating frustration); it can also be utilised to trigger control actions based on user preferences. If we view this functionality from the perspective of an enabling/constraining dynamic (Kiran, 2015), it can be argued that NAT enables the user to delegate the "normal" process of self-regulation (of internal states) to technology, and this transference inevitably constrains and suppresses our human capacity for self-regulation. Similarly, with respect to implicit modes of control, NAT enables an implicit process of decision/selection/action without any need for the user to explicitly formulate or communicate their wish or intention, e.g., Zander et al., (2016); however, this implicit mode of control both constrains and impinges on our human capacity for choose and act in an independent fashion.

From an ethical perspective, technological mediation can involve the person in decision-making as an ethical agent or alienate individuals by limiting their capacity for self-determination, e.g., by formulating choices in specific ways or restricting choices (Verbeek, 2011). As stated earlier, the NAT interface is designed to promote desirable psychological states or mitigate undesirable ones; see Fairclough, (2021) for a fuller discussion. However, the concept of software acting on behalf of the user or promoting the state of the user in a "desirable" way is paternalistic at best and dictatorial at worse. Also, from a practical perspective, the practice of translating human goals and values into the process of software design is inherently challenging (Friedman et al., 2006), and if there is a mismatch between what the user wants and what the system does, the former will be overruled by the latter, usurping the autonomy of the person.

Feedback from the NAT interface can enable the person to engage with SRP in a deeper fashion, to actively introspect and deepen self-awareness, but this type of mediation of SRP is double-edged. Also, the underlying "dataveillance" model (Zimmer, 2008) of NAT creates an uneven communication dynamic where technology proactively intervenes at the interface, thus forcing the user into a reactive position. When feedback is presented, the user must ask "was that me?" (i.e., was that feedback or function reflective of my lived experience or what I wanted to do?) or "was that you?" (i.e., was that feedback or function an artifact of the way in which algorithms have interpreted my lived experience?). Because feedback at the NAT interface is derived from an algorithmic interpretation of lived experience, the nature of the Alterity relation falls onto a continuum between two extreme positions. Users can relate to the Alterity of the NAT interface as either: (i) a quasi-self with an expectation that feedback from the interface will correspond with lived experience or current intentions, or (ii) a quasi-other that is characterised by discrepancy between feedback and lived experience and "an objectness to which humans relate" (Ihde, 1975) (p. 153). In the case of (ii), the degree of objectivity offered by system feedback may be valued (Ihde, 1975) as these qualities counteract the inherent subjectivity of human experience. On the other hand, the existence of discrepancy can alienate and confuse the user by creating the sense of a divided- or split-self.

This postphenomenological perspective on NAT has revealed how the technological intentionality of the system emerges from a confluence of different technological elements within the closed loop. From the perspective of the user, this technology actively mediates the relationship between the self and SRP by creating a hybrid form of intentionality that is bidirectional and reflexive. This hybrid intentionality is inherently selective, reflecting the selection of sensors and the design of algorithms that reveal and magnify some aspects of lived experience at the expense of others. NAT is configured to enable implicit modes of communication and control, which creates an Alterity relation that can engage the person as an extension of self or alienate the person as a manifestation of a technological "other". The next section will expand on the theme of those built-in selectivities that are an inherent part of the system by considering how lived experience is quantified by algorithms within NAT and re-represented to the human user.

3 Lived Experience vs. Quantification of Lived Experience

The quantification of psychological states (cognitions/feelings/intentions) via neurophysiological data is a fundamental element of NAT. This technology is designed to translate intentions and psychological states from the realm of first-person, lived experience into a quantitative format that is compatible with computational processing. Translation from the realm of human to the machine represents a border crossing (Hayles, 1999) and the resulting data has a paradoxical quality, concisely captured by the phrase "dematerialised materialism" (Seltzer, 1992) (p.14). Hayles (1999) makes a related point when she compares "the plenitude of embodiment" from lived experience with "the (relative) sparseness of abstraction" (p.99) that characterises a quantification of that lived experience. We take this observation as a starting point to compare the qualities of first-person lived experience with a quantification of that experience in Table 1.

Table 1 contrasts a reductive quantification of self with the rich and nuanced first-person experiences of cognition and interoceptive awareness (Khalsa et al., 2009). The transition from the left to the right-hand column of Table 1 is characterised by a process of enormous simplification and selectivity, which is suboptimal with respect to representative fidelity, but essential for the lived experience of the person to be converted into machine-compatible formats. It is also significant that any quantification of lived experience are delivered from a third-person perspective. This characteristic is rooted in positivism and its dominant influence on the empirical tradition in psychology (Tolman, 1992). For the user of NAT, feedback from the user interface functions as a reflexive form of self-representation (Von Foerster, 1984), where intentions and thoughts are simultaneously experienced in first-person alongside a real-time quantification of those lived experiences, as measured from the third-person vantage of an impersonal "other".

As a secondary issue for the purpose of the current paper, we should also note that quantification of lived experience creates a permanent data record that nullifies the transience and privacy of first-person experience and introduces the possibility of incursion, such as the possibility of a BCI being "hacked" (Ienca et al., 2018; Ienca & Haselager, 2016) or information being stolen from wearable sensors (Schukat et al., 2016).

Table 1 Characteristics of lived experience vs. quantification of experience

Lived experience	Quantification of experience
Rich	Sparse
Embodied	Immaterial
Qualitative	Quantitative
Temporary	Permanent
Private	Public
Secure	Pregnable
First-person perspective	Third-person perspective

The primary purpose of the analysis shown in Table 1 is to demonstrate how the quantification of intentions, cognitions, and emotions differs in fundamental ways from lived experience. The process of quantification creates a selective, impoverished, and highly simplified representation of lived experience, bereft of qualitative nuance, and observed from a third-person perspective. The same highly-processed quantification of experience is the basis of the real-time feedback delivered to the user of a NAT. Given the stark contrast shown in Table 1, one questions whether a quantification of lived experience that is so simplified and algorithmically derived can achieve even the minimal degree of congruence with those thoughts, intentions, and feelings that are experienced by the person. But, perhaps the more pressing question is whether NAT users can integrate feedback from such an impoverished representation into their SRP, and how does that process of integration transform SRP as a direct result? The fundamental differences between both perspectives can disorientate the user who perceives intentions, thoughts, and feelings from a first-person perspective while simultaneously receiving feedback on those dimensions of lived experience via the "technological gaze" of the system (Lewis, 2020).

Despite this potential for disruption, a reductive quantification of lived experience is the price of using NAT and similar technologies to monitor and quantify psychological concepts in real-time. Technology can only represent lived experience by analogy and at the cost of enormous simplification (Hayles, 1999). The key issue for the user is how to reconcile lived experience with this quantified representation, which is the topic of the next section.

4 Alterity and the Design of the User Interface

The Alterity relation describes how the person relates to feedback from NAT as a technological actor (Ihde, 1975, 1990). From the perspective of the user, this Alterity relation is characterised by four attributes: (1) selectivity, i.e., one aspect or dimension of lived experience is represented by feedback, (2) reflexivity, i.e., feedback from the system is self-referential, (3) reactivity, i.e., feedback changes in response to changes in user state or intention in real-time, and (4) vulnerability, i.e., neurophysiological sensors can access dimensions of human experience that are concealed from other people (Liberati & Nagataki, 2019). As stated in the previous section, the presentation of feedback at the user interface must be reconciled or consolidated with SRP, a process that bears similarities to the act of double-embodiment described by Buongiorno (2019), where technology mediates the first-person experience of internal states like cognition and emotion (Liberati, 2019).

The precise way in which SRP are mediated by NAT depends on the design characteristics of feedback presented via the user interface. According to the framework outlined by Tromp et al., (2011), the design of a user interface is characterised by its visibility and force. Visibility in this case being quite literally whether feedback at the interface is designed to capture the attention of the user. Alternatively, the designer can opt for a low-visibility option, where feedback must be actively solicited by the user, e.g., changes to available menu options that are only observable by selecting that menu. NAT interfaces that are designed to be highly visible will strongly mediate SRP by confronting the person with a quantification of lived experience that is difficult to ignore. This "glass-box" approach is aligned to the perspective of explainable artificial intelligence (Barredo Arrieta et al., 2020; Gunning et al., 2019), because high-profile/high-resolution feedback allows the user to develop an understanding of those autonomous algorithms working within the inner horizon of NAT. This transparency engenders trust by allowing the user to develop causal models about the contingencies between lived experience and feedback from the interface (Shin, 2021). In general, mediation of SRP via NAT is expected to occur as a collateral effect of working with this technology. However, the impact of mediation on SRP is most pronounced when feedback is presented overtly, clearly, and with high resolution.

This dimension of visibility is supplemented by the force of the interface wherein "strong" interfaces activate functions autonomously without input from the user, whereas "weaker" interfaces are characterised by implicative interventions, such as confirmation dialogues and selective highlighting that suggests a likely course of action (Tromp et al., 2011). This dimension is identical to the more detailed continuum of manual to autonomous control systems described by Parasuraman et al., (2000). Autonomous activation of functions at the user interface can also functions as a form of user feedback, e.g., if the difficulty of a neuroadaptive game increased, this change in game demand provides tacit that the system has assessed the player to be bored. Similarly, a less forceful interface design that offers a confirmatory dialogue represents a form of feedback to be consolidated with SRP.

Let us consider a hypothetical scenario to understand the interaction between feedback at the NAT user interface and SRP in greater detail. Imagine a computerised learning program designed to detect mental overload in order to optimise the learning experience. If the student is assessed to be mentally overloaded, the system automatically offers additional teaching support. In the first instance, consider a case where the user experiences mental overload, which automatically and promptly initiates a highly visible offer of teaching support from the system. In this example, there is a correspondence between lived experience and feedback from the interface. This feedback confirms the integrity of SRP - and the NAT interface is perceived as a "quasi-self", i.e., an algorithmic extension of SRP. This is positive from the user's perspective, i.e., the system "works", but there is a long-term risk that perceiving the system as a quasi-self may prompt users to favour feedback from the NAT interface over SRP, due to the scientific and medical authority imbued by sensor technology and machine intelligence (Schukat et al., 2016). This possibility is confirmed by research where subjective self-assessment of psychological states was observed to converge with a technological assessment of psychological states through repeated interactions with NAT (Fairclough et al., 2015).

Now let us consider two instances where feedback at the NAT interface deviates from SRP (Fairclough, 2009):

(1) The user feels that he is mentally overloaded but there is no visible response at the NAT interface. In this case, the user may question whether he is sufficiently overloaded to trigger additional teaching support from the system. He may also ponder whether his self-appraisal of mental workload is accurate; alternatively,

he could question whether the algorithmic assessment of mental overload performed by the system is prone to the production of false negatives.

(2) The machine offers additional teaching support, but the user does not perceive herself to be mentally overloaded in any subjective sense. In this case, the user may question the sensitivity of her SRP with respect to self-assessed mental workload, or she may assume that the discrepancy lies with an algorithmic assessment of mental overload that is too finely tuned and disposed towards the production of false positives.

The presence of discrepancy forces a binary choice onto the user, either feedback is correct and SRP are erroneous, or vice versa. In both cases, the Alterity relation is characterised by NAT interface functioning as a quasi-other in a Sartrean sense, subverting the integrity of SRP and forcing the person to either transcend or appropriate this algorithmic assessment (Sartre, 1994). But even if feedback from the NAT interface is judged to be erroneous and is subsequently transcended, the presence of visible feedback will mediate SRP indirectly by forcing the user into an introspective reflection on current lived experience.

The same logic extends to those NAT systems designed to extend human agency by triggering commands at the interface. Users receive feedback on this process by observing forceful NAT interfaces acting autonomously on their behalf, which serves as a counterpoint to their human capacity to choose with full cognizance. If there is a correspondence between intention and action at the interface, then a feeling of agency (Moore, 2016) is preserved, and the interface is regarded as a quasi-self. If the action at the interface deviates from intentionality, the user experiences a lack of agency and disrupted autonomy—and a strong sense that he is neither responsible nor accountable for that particular action (Kellmeyer et al., 2016).

In both types of interaction, users assess congruence between SRP and algorithmic assessment via feedback, but the decision to appropriate or transcend is shaped by the visibility and force of the interface design. Tromp et al., (2011) described four categories of design based on the levels of visibility and force of the user interface, each of which may mediate SRP in different ways. A persuasive interface combines high visibility with low force. This type of interface design might incorporate a "live" gauge indicating the current level of mental workload, when high levels are registered, the system prompts the user with a dialogue box, e.g., "is there anything I can do to help?" The availability of highly visible, high-resolution feedback prompts repeated comparisons between users' lived experience and the algorithmic assessment provided by the system even in the absence of the dialogue box. With respect to technological mediation in the practical domain, a persuasive interface enables the user to integrate system feedback into SRP with high resolution and to implicitly communicate their mental workload status to the technology. The presence of a confirmatory dialogue also involves the user in the decision as to whether their experience of mental overload could be resolved by the presentation of teaching support (Kiran, 2015). On the other hand, a coercive interface would present the same highly visible feedback to the user but would automatically present teaching support. This interface constrains the user's possibilities for self-regulation and runs the risk of alienating the person through a paternalistic mode of communication (Kiran, 2015).

If visibility is low and the force of the interface is weak, we have what Tromp et al., (2011) called a seductive interface. In this example of interface design, there is no continuous display of current state, the user only receives irregular feedback from the interface when the target state of mental overload has been detected. This low-visibility option does not create as many opportunities for the user to compare lived experience with the assessment of the system compared to the previous two examples, but like the persuasive design, it involves the user in the decision of whether to provide teaching support or not. The fourth type of interface design is called decisive implicative and simply communicates by action, in this case, by delivering teaching support. In this case, the design of the interface constrains the user's options and mediates SRP by delivering a solution to mitigate an undesirable user state in much the same way as a coercive interface design.

To summarise, the user interface responds to real-time changes in neurophysiology and functions as a technological actor, creating an Alterity relation that mediates SRP. The way in which feedback is presented at the user interface has a profound effect on how SRP are mediated by the technology. It is possible for a NAT that is perceived as a quasi-self that subverts or even occludes SRP, but that type of relation would require consistently high levels of congruence between lived experience and its algorithmic quantification. Discrepancies between SRP and feedback at the interface must be resolved by either favouring technological assessment (appropriation) over SRP or vice versa (transcendence). In both cases, the specific design of the user interface with respect to visibility and force determines the types of technological mediation that occur in the practical and ethical dimensions.

5 Summary and Conclusions

This paper was concerned with the ways in which interaction with NAT can mediate SRP, such as self-awareness, self-knowledge, and a sense of agency. We have described NAT as a closed-loop neurotechnology designed to enhance HCI in several ways. A postphenomenological analysis of this closed loop identified the prominent role played by machine hermeneutics and the alterity relation with respect to the technological intentionality of the system and how it can mediate SRP.

We have argued that machine hermeneutics produce an assessment of psychological states and intentions that is quantitative, reductive, and simplistic. This process of simplification is also completely necessary to support a closed-loop technology that operates in real-time. This simplified quantification of lived experience is conveyed to the user via the design of the interface, which must be consolidated with SRP and the lived experience of the user.

This process of consolidation is crucial in determining how SRP are mediated by NAT. Features of interface design, such as visibility and force, determine the efficacy of this process. Visible feedback is easier to assimilate and compare with SRP than opaque formats. Regardless of specific formats, feedback from the NAT interface may be congruent or divergent from SRP. We have argued that the degree of congruence between feedback and SRP determines the precise nature of the Alterity relation between user and technology. When system feedback confirms SRP, the

NAT is viewed as a quasi-self, i.e., an extension of SRP. If there is a discrepancy between feedback and SRP, NAT is viewed as a quasi-other, i.e., a computational assessment of lived experience that is distinct from actual lived experience. We argue that the potential for NAT to disrupt or subvert of SRP is more likely when the technology is perceived to function as a quasi-self. If NAT is viewed as a quasi-other, the user can transcend feedback from the system in one of two ways, by either disregarding it completely in favour of SRP or by using feedback from the system to supplement SRP. We conclude that the design of the user interface determines the precise nature of this alterity relation, shaping how SRP are mediated by interaction with NAT.

Acknowledgements The author would like to acknowledge comments received on an earlier version of this paper from Alexander Karran, Christopher Baker, and Kellyann Stamp, as well as the constructive feedback from two anonymous reviewers.

Author Contribution SF is the sole author of the work.

Data Availability Not applicable.

Declarations

Ethics Approval and Consent to Participate Not applicable

Consent for Publication Not applicable

Competing Interests The author declares no competing interests.

References

Allanson, J., & Fairclough, S. H. (2004). A research agenda for physiological computing. *Interacting with Computers, 16*(5). https://doi.org/10.1016/j.intcom.2004.08.001

Appriou, A., Cichocki, A., & Lotte, F. (2020). Modern machine-learning algorithms: For classifying cognitive and affective states from electroencephalography signals. *IEEE Systems, Man, and Cybernetics Magazine, 6*(3), 29–38. https://doi.org/10.1109/MSMC.2020.2968638

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Brand, Y., & Schulte, A. (2021). Workload-adaptive and task-specific support for cockpit crews: design and evaluation of an adaptive associate system. *Human-Intelligent Systems Integration, 3*(2), 187–199. https://doi.org/10.1007/s42454-020-00018-8

Buongiorno, F. (2019). Embodiment, Disembodiment and Re-embodiment in the Construction of the Digital Self. *Humana Mente, 12*(36).
Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology, 19*(2), 272–309. https://doi.org/10.1080/15213269.2015.1015740
Dey, A., Chatburn, A., & Billinghurst, M. (2019). Exploration of an EEG-based cognitively adaptive training system in virtual reality. *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 220–226. https://doi.org/10.1109/VR.2019.8797840
Dey, A. K. (2018). Context-aware computing. In *Ubiquitous computing fundamentals* (pp. 335–366). Chapman and Hall/CRC.
Di Flumeri, G., De Crescenzio, F., Berberian, B., Ohneiser, O., Kramer, J., Aricò, P., Borghini, G., Babiloni, F., Bagassi, S., & Piastra, S. (2019). Brain–computer interface-based adaptive automation to prevent out-of-the-loop phenomenon in air traffic controllers dealing with highly automated systems. *Frontiers in Human Neuroscience, 13*, 296. https://doi.org/10.3389/fnhum.2019.00296
Dorrestijn, S. (2017). The care of our hybrid selves: Ethics in times of technical mediation. *Foundations of Science, 22*(2), 311–321. https://doi.org/10.1007/s10699-015-9440-0
Fairclough, S. (2015). A closed-loop perspective on symbiotic human-computer interaction. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 9359). https://doi.org/10.1007/978-3-319-24917-9_6
Fairclough, S. H. (2009). Fundamentals of physiological computing. *Interacting with Computers, 21*(1–2). https://doi.org/10.1016/j.intcom.2008.10.011
Fairclough, S. H. (2017). Physiological computing and intelligent adaptation. In *Emotions and affect in human factors and human-computer interaction* (pp. 539–556). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-801851-4.00020-3
Fairclough, S. H. (2021). Designing human-computer interaction with neuroadaptive technology. In S. H. Fairclough & T. O. Zander (Eds.), *Current Research In Neuroadaptive Technology*. Elsevier.
Fairclough, S. H., & Gilleade, K. (2014). *Advances in physiological computing*. Springer.
Fairclough, S.H., Karran, A. J., & Gilleade, K. (2015). Classification accuracy from the perspective of the user: Real-time interaction with physiological computing. *Conference on Human Factors in Computing Systems - Proceedings*, *2015-April*. https://doi.org/10.1145/2702123.2702454
Fairclough, Stephen H. (2022). Designing human-computer interaction with neuroadaptive technology. In Stephen H Fairclough & T. O. B. T.-C. R. in N. T. Zander (Eds.), *Current Research In Neuroadaptive Technology* (pp. 1–15). Academic Press. https://doi.org/10.1016/B978-0-12-821413-8.00006-3
Fairclough, S. H., Dobbins, C., & Stamp, K. (2021). Classification of game demand and the presence of experimental pain using functional near-infrared spectroscopy. *Frontiers in Neuroergonomics, 2*. https://doi.org/10.3389/fnrgo.2021.695309
Fairclough, S. H., & Zander, T. O. (2022). *Current Research In Neuroadaptive Technology*. Academic Press.
Floridi, L. (2007). A look into the future impact of ICT on our lives. *The Information Society, 23*(1), 59–64. https://doi.org/10.1080/01972240601059094
Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press.
Floridi, L. (2019). Marketing as control of human interfaces and its political exploitation. *Philosophy & Technology, 32*(3), 379–388. https://doi.org/10.1007/s13347-019-00374-7
Friedman, B., Kahn Jr., P. H., & Borning, A. (2006). Value sensitive design and information systems. *Human-Computer Interaction and Management Information Systems: Foundations*, 1–27.
Froese, T. (2014). Bio-machine hybrid technology: A theoretical assessment and some suggestions for improved future design. *Philosophy & Technology, 27*(4), 539–560. https://doi.org/10.1007/s13347-013-0130-y
Gateau, T., Ayaz, H., & Dehais, F. (2018). In silico vs. over the clouds: On-the-fly mental state estimation of aircraft pilots, using a functional near infrared spectroscopy based passive-BCI. *Frontiers in Human Neuroscience, 12*. https://doi.org/10.3389/fnhum.2018.00187
Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G.-Z. (2019). XAI: Explainable artificial intelligence. *Science*. *Robotics, 4*(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120
Hayles, K. N. (1999). *How we became posthuman*. University of Chicago Press.

Hettinger, L. J., Branco, P., Encarnaco, L. M., & Bonato, P. (2003). Neuroadaptive technologies: Applying neuroergonomics to the design of advanced interfaces. *Theoretical Issues in Ergonomic Science, 4*(1–2), 220–237.

Ienca, M., & Haselager, P. (2016). Hacking the brain: Brain–computer interfacing technology and the ethics of neurosecurity. *Ethics and Information Technology, 18*(2), 117–129. https://doi.org/10.1007/s10676-016-9398-9

Ienca, M., Haselager, P., & Emanuel, E. J. (2018). Brain leaks and consumer neurotechnology. *Nature Biotechnology, 36*(9), 805–810. https://doi.org/10.1038/nbt.4240

Ihde, D. (1975). The experience of technology: Human-machine relations. *Cultural Hermeneutics, 2*(3), 267–279. https://doi.org/10.1177/019145377500200304

Ihde, D. (1990). *Technology and the Lifeworld*. Indiana University Press.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*(9), 641–661.

Just, N., & Latzer, M. (2016). Governance by algorithms: Reality construction by algorithmic selection on the Internet. *Media, Culture & Society, 39*(2), 238–258. https://doi.org/10.1177/0163443716643157

Kangassalo, L., Spapé, M., & Ruotsalo, T. (2020). Neuroadaptive modelling for generating images matching perceptual categories. *Scientific Reports, 10*(1), 14719. https://doi.org/10.1038/s41598-020-71287-1

Kellmeyer, P., Cochrane, T., Müller, O., Mitchell, C., Ball, T., Fins, J. J., & Biller-Andorno, N. (2016). The effects of closed-loop medical devices on the autonomy and accountability of persons and systems. *Cambridge Quarterly of Healthcare Ethics, 25*(4), 623–633. https://doi.org/10.1017/S0963180116000359

Khalsa, S. S., Rudrauf, D., Feinstein, J. S., & Tranel, D. (2009). The pathways of interoceptive awareness. *Nature Neuroscience, 12*(12), 1494–1496. https://doi.org/10.1038/nn.2411

Kiran, A. H. (2015). Four Dimensions of Technological Mediation. In R. Rosenberger & P.-P. Verbeek (Eds.), *Postphenomenological investigations: Essays on human-technology relations* (pp. 123–139). Lexington Books.

Krol, L. R., & Zander, T. O. (2022). Defining neuroadaptive technology: The trouble with implicit human-computer interaction. In Stephen H Fairclough & T. O. B. T.-C. R. in N. T. Zander (Eds.), *Current Research In Neuroadaptive Technology* (pp. 17–42). Academic Press. https://doi.org/10.1016/B978-0-12-821413-8.00007-5

Leamy, D. J., Kocijan, J., Domijan, K., Duffin, J., Roche, R. A. P., Commins, S., Collins, R., & Ward, T. E. (2014). An exploration of EEG features during recovery following stroke – Implications for BCI-mediated neurorehabilitation therapy. *Journal of NeuroEngineering and Rehabilitation, 11*(1), 9. https://doi.org/10.1186/1743-0003-11-9

Lewis, R. S. (2020). Technological Gaze: Understanding how technologies transform perception. In A. Daly, F. Cummins, J. Jardine, & D. Moran (Eds.), *Perception and the inhuman gaze: Perspectives from philosophy, phenomenology, and the sciences*. Routledge.

Li, I., Dey, A., & Forlizzi, J. (2010). A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 557–566). ACM. https://doi.org/10.1145/1753326.1753409

Liberati, N. (2019). Emotions and digital technologies. *HUMANA.MENTE. Journal of Philosophical Studies, 12*(36), 292–309.

Liberati, N., & Nagataki, S. (2019). Vulnerability under the gaze of robots: Relations among humans and robots. *AI & Society, 34*(2), 333–342. https://doi.org/10.1007/s00146-018-0849-1

Luan, S. T. (2020). The hidden dimensions of human–technology relations. *Philosophy & Technology, 33*(1), 141–165. https://doi.org/10.1007/s13347-019-00349-8

Moore, J. W. (2016). What is the sense of agency and why does it matter? *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01272

Mumtaz, W., Rasheed, S., & Irfan, A. (2021). Review of challenges associated with the EEG artifact removal methods. *Biomedical Signal Processing and Control, 68*, 102741. https://doi.org/10.1016/j.bspc.2021.102741

Mykhailov, D. (2020). The phenomenological roots of technological intentionality: A postphenomenological perspective. *Frontiers of Philosophy In China, 15*(4), 612–635.

Mykhailov, D., & Liberati, N. (2022). A study of technological intentionality in C++ and generative adversarial model: Phenomenological and postphenomenological perspectives. *Foundations of Science*. https://doi.org/10.1007/s10699-022-09833-5

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 30*(3), 286–297. https://doi.org/10.1109/3468.844354
Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology, 40*, 187–195.
Rapp, A. (2021). Wearable technologies as extensions: A postphenomenological framework and its design implications. *Human–Computer Interaction*, 1–39. https://doi.org/10.1080/07370024.2021.1927039
Rapp, A., & Tirassa, M. (2017). Know thyself: A theory of the self for personal informatics. *Human–Computer Interaction, 32*(5–6), 335–380. https://doi.org/10.1080/07370024.2017.1285704
Rosenberger, R., & Verbeek, P.-P. (2015). A field guide to postphenomenology. In R. Rosenberger & P.-P. Verbeek (Eds.), *Postphenomenological investigations: Essays on human-technology relations* (pp. 9–41). Lexington Books.
Scerbo, M. W., Freeman, F. G., & Mikulka, P. J. (2003). A brain-based system for adaptive automation. *Theoretical Issues in Ergonomics Science, 4*(1–2), 200–219. https://doi.org/10.1080/1463922021000020891
Sartre, J.-P. (1994). *Being and Nothingness*. Routledge: An Essay in Phemonological Ontology.
Schukat, M., McCaldin, D., Wang, K., Schreier, G., Lovell, N. H., Marschollek, M., & Redmond, S. J. (2016). Unintended consequences of wearable sensor use in healthcare. *Yearbook of Medical Informatics, 25*(1), 73–86.
Seltzer, M. (1992). *Bodies and Machines*. Routledge.
Serbedzija, N., & Fairclough, S. (2012). Reflective pervasive systems. *ACM Transactions on Autonomous and Adaptive Systems, 7*(1). https://doi.org/10.1145/2168260.2168272
Sharon, T. (2017). Self-tracking for health and the quantified self: Re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. *Philosophy & Technology, 30*(1), 93–121. https://doi.org/10.1007/s13347-016-0215-5
Shi, Y., Liu, R., He, L., Feng, H., Li, Y., & Li, Z. (2020). Recent development of implantable and flexible nerve electrodes. *Smart Materials in Medicine, 1*, 131–147. https://doi.org/10.1016/j.smaim.2020.08.002
Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies, 146*, 102551. https://doi.org/10.1016/j.ijhcs.2020.102551
Steinert, S., Bublitz, C., Jox, R., & Friedrich, O. (2019). Doing things with thoughts: Brain-computer interfaces and disembodied agency. *Philosophy & Technology, 32*(3), 457–482. https://doi.org/10.1007/s13347-018-0308-4
Tolman, C. W. (1992). *Positivism In psychology: Historical and contemporary issues*. Springer.
Tromp, N., Hekkert, P., & Verbeek, P.-P. (2011). Design for socially responsible behavior: A classification of influence based on intended user experience. *Design Issues, 27*(3), 3–19. https://doi.org/10.1162/DESI_a_00087
Verbeek, P. (2005). *What things do: Philosophical reflections on technology, agency and design*. Penn State Press.
Verbeek, P. P. (2008). Cyborg intentionality: Rethinking the phenomenology of human-technology relations. *Phenomenology and the Cognitive Sciences, 7*(3), 387–395. https://doi.org/10.1007/s11097-008-9099-x
Verbeek, P.-P. (2011). *Moralizing technology: Understanding and designing the morality of things*. Chicago University Press.
Von Foerster, H. (1984). *Observing systems*. Intersystems Publications.
Walter, C., Rosenstiel, W., Bogdan, M., Gerjets, P., & Spüler, M. (2017). Online EEG-based workload adaptation of an arithmetic learning environment. *Frontiers in Human Neuroscience, 11*. https://doi.org/10.3389/fnhum.2017.00286
Weiner, N. (1950). *The human use of human beings: Cybernetics & society*. Houghton Mifflin.
Wellner, G. (2020). Material hermeneutic of digital technologies in the age of AI. *AI & Society*. https://doi.org/10.1007/s00146-020-00952-w
Wiltse, H. (2014). Unpacking digital material mediation. *Techné: Research in Philosophy and Technology, 18*(3), 154–182.
Zander, T. O., & Kothe, C. (2011). Towards passive brain-computer interfaces: Applying brain-computer interface technology to human-machine systems in general. *Journal of Neural Engineering, 8*, 1–5.
Zander, T. O., Krol, L. R., Birbaumer, N. P., & Gramann, K. (2016). Neuroadaptive technology enables implicit cursor control based on medial prefrontal cortex activity. *Proceedings of the National Academy of Sciences, 113*(52), 14898–14903. https://doi.org/10.1073/pnas.1605155114

Zimmer, M. (2008). The gaze of the perfect search engine: Google as an infrastructure of dataveillance. In A. Spink & M. Zimmer (Eds.), *Web Search* (pp. 77–99). Springer. https://doi.org/10.1007/978-3-540-75829-7_6

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.