Technische Universität München
Fakultät für Elektrotechnik und Informationstechnik
Lehrstuhl für Kognitive Systeme

Error-related potentials in passive brain-computer interfaces for quantitative assessment and adaptation of robotic systems during human-robot interaction

Dipl.-Ing. (DH) Stefan Karl-Heinz Ehrlich, M.Sc.

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender:

Prof. Dr.-Ing. Werner Hemmert

Prüfer der Dissertation:

1.   Prof. Gordon Cheng, Ph.D.

2.   Prof. Cuntai Guan, Ph.D.

Die Dissertation wurde am 02.10.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 10.04.2020 angenommen.

Abstract

Successful collaboration require interaction partners to monitor and predict each other's behavior for establishing shared representations of beliefs, intentions, and goals. The ability to share mutual beliefs and intentions does not only enable better and more efficient communication, but also serves long term satisfaction of interactors, bonding, trust and relationship building. A transfer of this ability to artificial systems, such as autonomous robotic systems has not been achieved so far and progress along these lines has been widely acknowledged as key to more widespread acceptance of robotic systems in society. The development of robotic systems for interaction with human partners has so far been driven by two approaches:

- *Assessment:* in the domain of cognitive neuroscience, researchers systematically assess and validate robotic systems by studying social cognitive mechanisms in humans during human-robot interaction (HRI). These studies help engineers to design robotic systems to align better with human social traits and conventions. While this approach is scientifically well grounded, most studies employ subjective and behavioral outcome measures that are often not straightforwardly applicable in the design of new systems.

- *Adaptation:* In the domain of robotic engineering, researchers equip robots with user-adaptive mechanisms, including human intention detection, user-modelling, and approaches to handle co-adaptation in collaborative HRI. While this approach has developed fast, it is challenged by the limited perceptual capabilities of robotic systems which hinders the investigation and development of more complex HRI.

As will be demonstrated in this thesis, *passive brain-computer interfaces (BCI)* constitute a promising technology for addressing both real-time assessment and adaptation of autonomous systems, such as robots during HRI. Passive BCIs propose the decoding of user mental states from electro- or psychophysiological signals for augmenting human-machine interaction with information beyond standard sensory input deployable by machines to adapt to their human partner. Following this approach, we demonstrate that different mental states can be reliably decoded non-invasively from *electroencephalography (EEG)* recordings of human participants during interaction with real robotic systems. We focus here specifically on exploring the usefulness of *error-related potentials (ErrPs)*, a neural response measurable in the EEG in response to the human observation of self-committed errors or unexpected external events, and demonstrate that these ErrPs can be used to assess and adapt robot's behavior *to converge to the expectation of a human partner, this way facilitating and enhancing interaction performance*.

The outcomes of this thesis result from three experimental HRI studies and one simulation-based study involving different robotic systems and HRI scenarios and resolve in four main contributions: Regarding *assessment*, (1) the first experimental study demonstrates the feasibility to decode from the EEG of the human interaction partner, the intention to engage or to get engaged with a humanoid robot via gaze contact with an average offline accuracy of

80.4%. The results further demonstrate how differential human beliefs about the robot and the social role participants believed to have in the interaction are reflected in significant EEG modulations in response to human-robot gaze-contact, modulations that qualitatively resembled those of an ErrP. (2) In the second experimental study we showed that ErrPs can also be generated and decoded upon human observation of erroneous actions performed by a real humanoid robot. Cross-comparison with a screen-based cursor feedback indicated qualitatively similar brain activity modulations which suggested a common underlying neuronal origin. Quantitative comparisons showed however the ErrP effect strength to be dependent on the type of stimulus eliciting them, with significantly lower decodeability in response to actions performed by a real robotic system (average offline accuracy of 69%) as compared to screen-based stimuli (average offline accuracy of 91%). Regarding *adaptation*, (3) this thesis demonstrate that such decoded ErrPs can be used to online adapt robot behavior towards human expectations, which themselves could be volatile, such as the human adapting also to the robot (human-robot co-adaptation). This is addressed in a closed-loop experimental setup in which participants were asked to guess from the gazing behavior of a humanoid robot its underlying goal/intention while the robot's gaze behavior was adapted using online decoded ErrPs (average online accuracy of 81.8%). This closed-loop setup resulted in convergence of robot's gaze behavior towards patterns facilitating the interaction (significantly improved guessing performance) and thus an alignment of the robot's behavior with the human interaction partner's expectations. (4) This thesis further contributes with a computational model of human decision-making behavior and learning to explore, within the context of the previous experimental study, the influence of varying human and experimental factors on human-robot co-adaptation. The results of model-based simulations attributed a more important role to technical factors, specifically ErrP-decoders and methods for ErrP-based robot adaptation, and a minor role to human factors. These findings indicate that the ErrP-based human-robot co-adaptation approach is well generalizable across participants. Simulations also demonstrated that state-space extensions are feasible and mainly accompanied by longer co-adaptation convergence times, hence confirmed that the approach is well scalable to more complex HRI.

Zusammenfassung

Eine erfolgreiche Zusammenarbeit erfordert, dass die Interaktionspartner das Verhalten des anderen beobachten und vorhersagen, um gemeinsame Darstellungen von Erwartungen, Absichten und Zielen zu etablieren. Die Fähigkeit, gegenseitige Erwartungen und Absichten zu teilen, ermöglicht nicht nur eine bessere und effizientere Kommunikation, sondern dient auch der langfristigen Zufriedenheit von Interaktionspartnern, Bindung, Vertrauen und Beziehungsaufbau. Eine Übertragung dieser Fähigkeit auf künstliche Systeme, wie beispielsweise autonome Robotersysteme, ist bisher nicht erreicht worden, und Fortschritte in diese Richtung werden allgemein als Schlüssel zu einer breiteren Akzeptanz von Robotersystemen in der Gesellschaft anerkannt. Die Entwicklung von Robotersystemen für die Interaktion mit menschlichen Partnern wurde bisher traditionell durch zwei Ansätze vorangetrieben:

- *Assessment:* im Bereich der kognitiven Neurowissenschaften werden Robotersysteme systematisch bewertet und validiert, indem soziale kognitive Mechanismen beim Menschen während der Mensch-Roboter-Interaktion (HRI) untersuchen werden. Diese Studien helfen Ingenieuren, Robotersysteme zu entwerfen, die besser auf menschliche soziale Eigenschaften und Konventionen abgestimmt sind. Obwohl dieser Ansatz wissenschaftlich fundiert ist, verwenden die meisten Studien subjektive und verhaltensbedingte Messungen, die bei der Entwicklung neuer Systeme oft nicht einfach anwendbar sind.

- *Adaptation:* Im Bereich der Robotik werden Roboter mit benutzeradaptiven Mechanismen ausgestattet, einschließlich der Erkennung menschlicher Absichten, der Benutzermodellierung und Ansätze zur Handhabung von Koadaptation in der kollaborativen HRI. Obwohl sich dieser Ansatz schnell entwickelt hat, ist er durch die begrenzten Wahrnehmungsfähigkeiten von Robotersystemen eingeschränkt, was die Untersuchung und Entwicklung komplexerer HRI erschwert.

Wie in dieser Arbeit gezeigt wird, stellen *passive Brain-Computer Interfaces (BCI)* eine vielversprechende Technologie dar, um sowohl die Echtzeitbewertung als auch die Anpassung autonomer Systeme, wie beispielsweise Roboter während der HRI, anzugehen. Passive BCIs ermöglichen die Dekodierung von mentalen Zuständen der Benutzer, und bieten damit eine Erweiterung der üblichen sensorischen Informationen, die Maschinen üblicherweise zur Anpassung an ihren menschlichen Partner zur Verfügung stehen. Nach diesem Ansatz zeigen wir, dass verschiedene mentale Zustände aus den Aufzeichnungen der *Elektroenzephalographie (EEG)* von menschlichen Probanden während der Interaktion mit verschiedenen Robotersystemen zuverlässig und nicht-invasiv dekodiert werden können. Wir konzentrieren uns hier insbesondere auf die Untersuchung der Nützlichkeit von *Error-Related Potentials (ErrPs)*, einer neuronalen Reaktion, die im EEG als Reaktion auf die menschliche Beobachtung von selbstverschuldeten Fehlern oder unerwarteten externen Ereignissen messbar ist. Wir zeigen, dass diese ErrPs verwendet werden können, um *das Verhalten des Roboters zu validieren und*

in Echtzeit hinsichtlich der Erwartungen des menschlichen Interaktionspartners anzupassen, um somit die Interaktion zu erleichtern und zu verbessern.

Die Ergebnisse dieser Arbeit resultieren aus drei experimentellen HRI Studien am Menschen und einer simulationsbasierten Studie mit verschiedenen Robotersystemen und HRI-Szenarien und lösen sich in vier Hauptbeiträge auf: Bezüglich *Assessment*, (1) zeigt die erste experimentelle Studie die Machbarkeit, aus dem EEG des menschlichen Interaktionspartners, die Absicht einen Blickkontakt mit einem humanoiden Roboter herzustellen mit einer Offlinegenauigkeit von durchschnittlich 80.4% zu dekodieren. Die Ergebnisse zeigen weiter, wie sich unterschiedliche menschliche Erwartungen in Bezug auf den Roboter und die soziale Rolle, die die Testperson in der Interaktion zu haben glaubt, in signifikanten EEG-Modulationen als Reaktion auf den Blickkontakt zwischen Mensch und Roboter widerspiegeln, Modulationen, die qualitativ denen eines ErrPs ähneln. (2) In der zweiten experimentellen Studie zeigen wir, dass ErrPs auch nach menschlicher Beobachtung von Fehlverhalten eines echten humanoiden Roboters erzeugt und dekodiert werden können. Ein Quervergleich mit einem bildschirmbasierten Cursor-Feedback zeigte qualitativ ähnliche Modulationen der Gehirnaktivität, die auf einen gemeinsamen zugrunde liegenden neuronalen Ursprung hindeuteten. Quantitative Vergleiche zeigten jedoch, dass die Stärke des ErrP-Effekts von der Art des Stimulus abhängt und mit einer signifikant geringeren Dekodierbarkeit bei der menschlichen Beobachtung von realen Roboteraktionen (durchschnittliche Offlinegenauigkeit von 69%) im Vergleich zu bildschirmbasierten Stimuli (durchschnittliche Offlinegenauigkeit von 91%) einhergeht. Hinsichtlich *Adaptation*, (3) zeigt diese Arbeit, dass solche dekodierten ErrPs verwendet werden können, um das Roboterverhalten während der Interaktion an menschliche Erwartungen anzupassen. Erwartungen, die ihrerseits variable sein könnten, z. B. durch die menschliche Anpassung an den Roboter (Mensch-Roboter-Koadaption). Dies wird in einem closed-loop Versuchsaufbau untersucht, in dem die Probanden gebeten wurden, aus dem Blickverhalten eines humanoiden Roboters seine zugrunde liegende Intention abzuleiten, während das Blickverhalten des Roboters mit Hilfe von online dekodierten ErrPs (durchschnittliche Onlinegenauigkeit von 81.8%) angepasst wurde. Diese closed-loop Konfiguration führt zu einer Konvergenz des Blickverhaltens des Roboters welches die Interaktion erleichtert und damit eine effiziente Anpassung des Roboters an die Erwartungen des Probanden erreicht wird. (4) Diese Arbeit trägt weiterhin mit einem Computermodell des menschlichen Entscheidungsverhaltens und Lernens bei, um im Rahmen der vorherigen experimentellen Studie den Einfluss verschiedener menschlicher und experimenteller Faktoren auf die Mensch-Roboter-Koadaptation zu untersuchen. Die Ergebnisse der modellbasierten Simulationen ergaben, dass technische Faktoren, insbesondere hinsichtlich der ErrP-Dekoder und Verfahren zur fehlerbasierten Roboteranpassung, eine wichtigere Rolle als menschlichen Faktoren spielen. Diese Ergebnisse deuten darauf hin, dass der ErrP-basierte Mensch-Roboter-Koadaptionsansatz für alle Probanden gut generalisierbar ist. Simulationen zeigten auch, dass Zustands-Raum-Erweiterungen vor allem mit längeren Koadaptions-Konvergenzzeiten einhergehen, was darauf hindeutet, dass der Ansatz für komplexere HRI Szenarien gut skalierbar ist.

Acknowledgements

This thesis would not have been possible without the help and support of many people during the last years. I would like to express my deep gratitude to all of them.

First and foremost, I would like to thank my advisor Prof. Gordon Cheng, who gave me the opportunity to carry out my research as a doctoral student. I am especially grateful that he always encouraged me to pursue unconventional ideas; his support, patience, motivation, and guidance were invaluable for the success of this thesis. I also want to thank my former advisor Prof. Cuntai Guan (NTU, Singapore) who gave me the opportunity to carry out my master thesis at the Agency for Science, Technology and Research in Singapore and introduced me to research on brain-computer interfaces. His guidance during my master thesis was key decisive for the research carried out during my doctoral thesis.

I would like to also express my deep gratitude to my mentor Dr. Ana Alves-Pinto (TUM Klinikum rechts der Isar, Munich) who not only greatly supported me on the professional, but also on the personal level. Also, I would like to thank my former mentor Prof. Agnieszka Wykowska (IIT, Genoa) who introduced me to the world of cognitive neuroscience and greatly inspired the research carried out during my thesis.

Furthermore, I would like to thank all former and present colleagues at the Chair for Cognitive Systems for inspiring discussions and support: Florian Bergner, Dr. Sae Franklin, Dr. Mohsen Kaboli, Jasmin Kajopoulos, Dr. Barbara Kühnlenz, Dr. Pablo Lanillos, Quentin Leboutet, Dr. Emmanuel Dean Leon, Alireza Malekmohammadi, Dr. John Nassour, Constantin Uhde, Erhard Wieser. Also, I would like thank all external collaborators with whom I had the honor to work together: Prof. Renée Lampe and Dr. Ana Alves-Pinto (TUM Klinikum rechts der Isar, Munich), Prof. Cuntai Guan and Prof. Kat R. Agres (NTU, NUS, A*STAR, Singapore), Prof. Josef P. Rauschecker and Dr. Jessica Philipps-Silver (Georgetown University, Washington DC).

None of this would have been possible without the love and patience of my family, my life companion, and friends. I would like to express my heartfelt gratitude to my mother, father, and brother. A very special thank you to my loving companion, Caro Kania, who kept my life in balance and supported me with her unconditional love throughout the last years. I am in deep gratitude that you shared this journey with me. And finally, also to my dear friends Peter Abeling, Philipp Schütze, Byron-Lim Steffan, Frederik Landwehr, and Christian Müller who had to suffer through many of my monologues about new experiments, data analyses and philosophical implications, but nonetheless always encouraged me to continue.

At last, I would like to thank the funding sources that made this thesis possible, in particular, the Deutsche Forschungsgemeinschaft (DFG) through the International Graduate School of

Science and Engineering (IGSSE), the Elite Network Bavaria (ENB) through the MSNE master program in NeuroEngineering, and the TUM Faculty Graduate Center for electrical and computer engineering (FGZ-EI).

List of Publications

Journal articles

J5: Schönleitner, F. M., Otter, L., **Ehrlich, S. K.**, & Cheng, G. (2020). Calibration-Free Error-Related Potential Decoding with Adaptive Subject-Independent Models: A Comparative Study. IEEE Transactions on Medical Robotics and Bionics. DOI: `https://doi.org/10.1109/TMRB.2020.3012436`

J4: **Ehrlich, S. K.**, Agres, K. R., Guan, C., & Cheng, G. (2019). A closed-loop, music-based brain-computer interface for emotion mediation. PloS one, 14(3), e0213516. DOI: `https://doi.org/10.1371/journal.pone.0213516`

J3: **Ehrlich, S. K.**, & Cheng, G. (2019). A Feasibility Study for Validating Robot Actions Using EEG-Based Error-Related Potentials. International Journal of Social Robotics, 11(2), 271-283. DOI: `https://doi.org/10.1007/s12369-018-0501-8`

J2: **Ehrlich, S. K.**, & Cheng, G. (2018). Human-agent co-adaptation using error-related potentials. Journal of Neural Engineering, 15(6), 066014. DOI: `http://dx.doi.org/10.1088/1741-2552/aae069`

J1: Alves-Pinto, A., **Ehrlich, S.**, Cheng, G., Turova, V., Blumenstein, T., & Lampe, R. (2017). Effects of short-term piano training on measures of finger tapping, somatosensory perception and motor-related brain activity in patients with cerebral palsy. Neuropsychiatric disease and treatment, 13, 2705. DOI: `https://doi.org/10.2147/ndt.s145104`

Peer-reviewed conference papers

C8: **Ehrlich, S. K.**, & Cheng, G. (2019). A computational model of human decision making and learning for assessment of co-adaptation in neuro-adaptive human-robot interaction. In 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC) (pp. 264-271). IEEE. DOI: `https://doi.org/10.1109/SMC.2019.8913872`

C7: Schönleitner, F. M., Otter, L., **Ehrlich, S. K.**, & Cheng, G. (2019). A comparative study on adaptive subject-independent classification models for zero-calibration error-potential decoding. In 2019 IEEE International Conference on Cyborg and Bionic Systems (CBS) (pp. 85-90). IEEE. (Finalist best paper award). DOI: `https://doi.org/10.1109/CBS46900.2019.9114494`

C6: Braun, J. F., Díez-Valencia, G., **Ehrlich, S. K.**, Lanillos, P., & Cheng, G. (2019). A prototype of a P300 based brain-robot interface to enable multi-modal interaction for patients with limited mobility. In 2019 IEEE International Conference on Cyborg and Bionic Systems (CBS) (pp. 78-84). IEEE. (Finalist best paper award). DOI: `https://doi.org/10.1109/CBS46900.2019.9114484`

C5: **Ehrlich, S.**, Guan, C., & Cheng, G. (2017). A closed-loop Brain-Computer Music Interface for continuous affective interaction. In Orange Technologies (ICOT), 2017 International Conference on (pp. 176-179). IEEE. DOI: `https://doi.org/10.1109/icot.2017.8336116`

C4: **Ehrlich, S.**, Alves-Pinto, A., Lampe, R., & Cheng, G. (2017). A simple and practical sensorimotor EEG device for recording in patients with special needs. In Neurotechnix2017, CogNeuroEng 2017, Symposium on Cognitive Neural Engineering. DOI: `https://doi.org/10.5220/0006559100730079`

C3: **Ehrlich, S.**, & Cheng, G. (2016). A neuro-based method for detecting context-dependent erroneous robot action. In Humanoid Robots (Humanoids), 2016 IEEE-RAS 16th International Conference on (pp. 477-482). IEEE. DOI: `https://doi.org/10.1109/humanoids.2016.7803318`

C2: Hagerer, G. J., Lux, M., **Ehrlich, S.**, & Cheng, G. (2015). Augmenting affect from speech with generative music. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (pp. 977-982). ACM. DOI: `https://doi.org/10.1145/2702613.2732792`

C1: **Ehrlich, S.**, Wykowska, A., Ramirez-Amaro, K., & Cheng, G. (2014). When to engage in interaction—And how? EEG-based enhancement of robot's ability to sense social signals in HRI. In Humanoid Robots (Humanoids), 2014 14th IEEE-RAS International Conference on (pp. 1104-1109). IEEE. DOI: `https://doi.org/10.1109/humanoids.2014.7041506`

Workshop/conference abstracts and poster presentations

P14: **Ehrlich, S. K.**, & Cheng, G.: Passive BCI for assessment and adaptation of robotic systems during human-robot interaction. Abstract and presentation at BCI: Science & Practice, Samara 2020.

P13: Dimova, V., **Ehrlich, S. K.**, & Cheng, G.: ErrP Components of Self- and Agent-Related Errors in Human-Agent Collaboration. Abstract and presentation at BCI: Science & Practice, Samara 2020.

P12: Hoxha, I., Del Duca, F., **Ehrlich, S. K.**, Bergner, F., Berberich, N., & Cheng, G.: Improving user comfort in auditory-steady-state-response brain-computer interface by using a co-adaptive stimulus. Abstract and poster presentation at FENS2020.

P11: Popova, M., Tschiersch, M., Berberich, N., **Ehrlich, S. K.**, Franklin, D., & Cheng, G.: A bump-attractor spiking neural network for motor adaptation and washout based on Norepinephrine release in primary motor cortex. Abstract and poster presentation at CBS2020.

P10: Tschiersch, M., Popova, M., Berberich, N., **Ehrlich, S. K.**, Franklin, D., & Cheng, G.: A bump-attractor spiking neural network for motor learning based on Norepinephrine release. Abstract and poster presentation at FENS2020.

P9: Malekmohammadi, A., Jacobs, J. T., Uhde, C., **Ehrlich, S. K.**, Philipps-Silver, J., Toivi-ainen, P., Rauschecker, J. P., & Cheng, G.: Brain-To-Sound Computer Interfaces: Neurofeed-back of Music for Entrainment, Interaction and Neurorehabilitation. Poster at IGSSE Forum 2019, Raitenhaslach, Germany. (Finalist best poster award)

P8: Weiß, L., Zillekens, M., **Ehrlich, S. K.**, Rinklin, P., Cheng, G., & Wolfrum, B.: Bridging the gap: From abstract oddball experiments towards the classification of sparse events using a hybrid EEG-fNIRS approach. Abstract at ARTscientific Artinis NIRS symposium 2019, The Netherlands.

P7: Ehrlich, S. & Cheng, G.: Detecting erroneous robot behavior from the electroencephalo-gram (EEG) during human robot interaction. Abstract and poster presentation at Bernstein Conference 2016, Berlin, Germany.

P6: Ehrlich, S., Alves-Pinto, A., Lampe, R., & Cheng, G.: A simple and practical prototype of a sensorimotor EEG device. Abstract and poster presentation at Bernstein Conference 2016, Berlin, Germany.

P5: Ehrlich, S., Wykowska, A., Ramirez-Amaro, K., & Cheng, G. (2015). Predictive modeling of higher-cognitive processes involved in social engagement based on electroencephalog-raphy (EEG). Abstract and poster presentation at Bernstein Conference 2015, Heidelberg, Germany.

P4: Ehrlich, S., Wykowska, A., Ramirez-Amaro, K., & Cheng, G. (2015). Predictive modeling of higher-cognitive processes involved in social engagement based on electroencephalogra-phy (EEG). Poster presentation at Young Researcher's Retreat 2015, Nara, Japan.

P3: **Ehrlich, S.**, Wykowska, A., Ramirez-Amaro, K., & Cheng, G. (2015). Predictive modeling of higher-cognitive processes involved in social engagement based on electroencephalography (EEG). Abstract and poster presentation at CuttingEEG Workshop 2015, Berlin, Germany.

P2: **Ehrlich, S.**, Guan, C., & Cheng, G.: Closed-loop Interaction with Affective Brain-Music Interface. Abstract and poster presentation at Bernstein Conference 2013, Tübingen, Germany.

P1: **Ehrlich, S.**, Guan, C., & Cheng, G.: Monitoring emotional states in the loop and its potential applications. Poster presentation and talk at Biomedical Signal Analysis Conference 2013, Rio de Janeiro, Brazil.

Contents

List of Figures

List of Tables

List of Algorithms

List of Acronyms

U

V

1. Introduction

1.1. Motivation

Successful social interaction and collaboration require interaction partners to monitor and predict each other's behavior [1] for establishing shared representations of beliefs, intentions, and goals [2]. The ability to share mutual beliefs and intentions does not only enable better and more efficient communication between peers, but also serves long term satisfaction of interactors, bonding, trust and relationship building [3]. Humans and animals with social capabilities are excellent examples of organisms that have developed perceptual and cognitive mechanisms capable of realizing this process. A transfer of these mechanisms to autonomous artificial systems, such as autonomous robotic systems has not been achieved so far. This is due to the limited perceptual and cognitive capabilities of robotic systems [4] and to a limited understanding of human social cognition [5]. Progress along this line of research is widely acknowledged as the main path to more widespread acceptance and application of autonomous robotic systems in society [6], since robotic systems, that were until now largely perceived as sophisticated tools, may soon become entities which actively contribute to our society. Examples for such developments can be found in the domains of autonomous driving cars, dialogue systems and chatbots, as well as collaborative robotic systems. These systems have in common a close interaction with humans and with this they exert a significant influence on their human interaction partners. And while appropriate safety mechanisms constitute the most critical factor in such systems, aspects considering the human acceptance of the technical system as an interaction partner play as well an integral role. Questions that are frequently raised in the context of the design of such systems are (see [4, 7]): How to foster the human's willingness to engage and persistently interact with the system? How to realize synergistic and efficient interaction by exploiting both the human and machine's individual capabilities? How to establish long term trust and relationship building between humans and machines?

One increasingly important way to approach these questions in a quantitative and systematic fashion is by exploiting direct measures of brain activity from the human interaction partner that are modulated by that interaction [6]. As will be demonstrated in this thesis, closed-loop brain-machine interfaces constitute a very promising technology for real-time assessment and adaptation of autonomous machines to human's expectations and preferences. Indeed, all above mentioned research questions can to some extent be subsumed under the consideration of *how to align the machine's decisions and behavior with human's expectations and subjective preferences.* Because, if both expectations and preferences of the human interaction partner are met, comfort, trust, role definition and relationship building may automatically evolve towards human conventions. By addressing this central research question, this thesis

contributes to the development of a technology that can ultimately bring humans and machines closer together.

1.2. Problem definition

This thesis focuses on a specific subfield of HMI (human-machine interaction), namely collaborative HRI. Within this domain, further focus is put on HRI involving autonomous robotic systems with some form of agency. Meant here are HRI scenarios in which both human and robot are equally responsible and independent interaction partners, each with potentially different goals and / or attitudes towards solving a task or shaping the interaction. Examples are collaborative assembly tasks in factory setups or social human-robot communication in elderly care or service robotics.

Two main approaches towards improving robotic systems and interfaces in collaborative HRI have been pursued so far: (1) offline assessment and validation of robotic systems and HRI, and (2) online user-adaptive robotic systems and interfaces. Both approaches are subsequently introduced and current challenges and limitations are discussed.

1.2.1. Assessment: Offline validation of robotic systems and HRI

In the domain of cognitive neuroscience and experimental psychology, a typical approach to improve HRI is to offline assess robotic systems in experimental studies involving human participants. In experimental tasks different interaction settings, such as different robot appearance and behavior, are typically presented to human participants. Settings and/or interactions are rated/evaluated using different types of measures [8]: questionnaires, video recordings, interviews; performance measures such as reaction time, error rates or time to complete the task; behavioral measures such as eye tracking or motion tracking; electrophysiology and neuroimaging measures (for an overview the reader is referred to [6]). These measures inform, for example, about which scenarios or experimental conditions are preferred by human participants, and provide therefore guidelines to robot engineers. Examples are the work by Weiss and colleagues, e.g. [9] in 2010 in which they employed an autonomous mobile robot exploring the city, specifically public places in Munich. Among other aspects, they studied the human's (random passers-by) willingness to help the robot by pointing out directions to navigate to an intended goal. The study was based on the observation through experimenters and a customized questionnaire administered to passers-by after they engaged with the robot. Another example is the work by Wykowska and colleagues in 2015 [10] in which they demonstrated the sensitivity of human perception to behavioral characteristics related to perceived intentionality of robots. Specifically, they investigated if participants were able to tell apart if particular robot actions were pre-programmed, human-controlled, or modelled after human behavior (see Figure 1(a)). Here, performance measures were employed, in particular participants' judgment and related response reaction times. While such approaches are based on clear, scientifically grounded and testable hypotheses, the majority of research works still use exclusively performance measures and/or self-assessment using standardized or customized

questionnaires [11] since they are inexpensive to apply and easy to implement [6] (see for example the "Godspeed Questionnaire Series" - Goodspeed [12], the "Robotic Social Attributes Scale" - RoSAS [13], the "Usability, Social acceptance, User experience, and Societal impact questionnaire" - USUS [14], or the "NASA Task Load Index" - NASA-TLX [15]). However, subjective ratings usually disrupt natural interaction and are therefore typically administered at the end of the experiment (see above mentioned example of Weiss and colleagues [9]). As such, obtained results rarely inform about participants' assessment immediate to a (particular) robot action or behavior. This also accounts for performance measures, which are often reported as average across multiple trials to yield conclusive results about the contrasted conditions (repetitive execution of the same experimental condition to overcome trial-to-trial variations; see above mentioned example of Wykowska and colleagues [10]). While performance measures are objective, questionnaire-based self-assessment ratings are subjective and thus can be biased by participants' individual interpretation. The need for more objective and immediate measures for the assessment of robot behavior has been widely acknowledged, e.g. in [16]. Along this line, Wiese and colleagues recently argued that the best way to "make robots more social [would be a] systematic experimental approach based on behavioral and physiological neuroscience methods such as motion/eye-tracking, electroencephalography, or functional near-infrared spectroscopy embedded in interactive human-robot paradigms" [6] (see Figure 1(b)). Approaches employing such objective neuro-cognitive measures have been presented, although less frequently, for instance using EEG [17], or fMRI (functional magnetic resonance imaging) [18, 19] (see Figure 1(c)). In summary, while offline assessment of HRI is scientifically strongly grounded it is limited by being largely based on non-immediate, non-objective measures. This hinders the smooth and swift application of that acquired knowledge into the design of robots and is not suitable to application in more dynamic HRI.

1.2.2. Adaptation: Online user-adaptive robotic systems and interfaces

In the domain of robotics and engineering, a more common approach is to employ robotic systems with mechanisms to automatically adapt (during interaction) to their human interaction partners. These "user-adaptive systems consist of autonomous agents that are able to use some manner of information on their user in order to adapt to them" [21]. It has long been established that user-adaptive human-computer interfaces lead to significantly improved acceptance as compared to non-adaptive ones [22]. In the domain of social robotics and HRI, previous works have shown that robot adaptation significantly improves human's perceived robot trustworthiness [23]. One can distinguish here between three strands of research: (1) adaptation of robots to users (e.g. [24, 25], for an overview the reader is referred to [26]), (2) modeling of human adaptation to robots (e.g. [27] or [28]), and (3) human-robot co-adaptation (or mutual adaptation, e.g. [29] and [30]). The first two strands consider only one-way adaptation of either the robot to the human or vice-versa. An example is the work by Kühnlenz and colleagues in 2013 who employed an expressive robot head (see Figure 2(a), left) which adapted facial expressions towards user's emotional states (predicted from user's facial expressions, see Figure 2(a), right). This automatic adaptation towards congruency of human-robot emotional states increased the willingness of users towards helping the robot [31, 32].
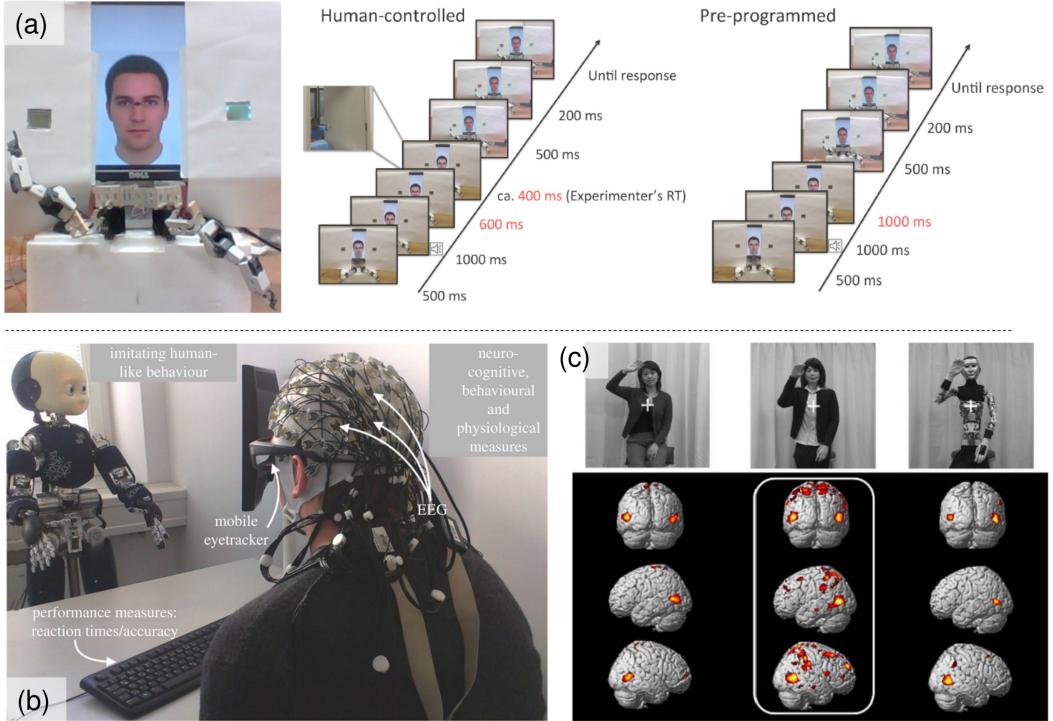
Figure 1 Assessment and validation of HRI: (a) Experimental setup and protocol testing the sensitivity of human perception to robot actions which was pre-programmed, human-controlled, or modelled after human behavior. To contrast conditions, performance measures were employed, e.g. the participants' judgment and related response reaction times. [Reprinted from Wykowska et al., 2015 [10]] (b) Example experimental set-up in which a human interacts with a humanoid robot iCub, while behavioral, neural and physiological measures are taken to examine the human social cognition. [Reprinted from Wykowska et al. 2016 [5]]. (c) Experimental study exploring and demonstrating differential neural responses to same actions performed by entities with varying degree of humanness: a human (top left), an android (top middle), and a humanoid robot (top right) [Reprinted from Ishiguro, 2015 [20], adapted from Saygin et al., 2011 [18]].

In many set-ups such one-way adaptation can be sufficient, and may, as demonstrated in the former example, implicitly trigger beneficial user adaptations. However, explicit considerations of both way adaptations are presumably particularly important to progress in the domain of collaborative HRI where robot acceptance and trust play a critical role in overall task success. For instance, a robot which solely adapts to human preferences may result in sub-optimal task performance because the human partner may not be aware of the optimal policy. In contrast, a robot which insists on executing an optimal policy against the human's preferences may erode the human's trust in the robot and as such lead to decreasing team performance [33]. Even though research on human-robot co-adaptation being a reasonable approach in collaborative HRI, there are only a few recent works by Nicolaidis and colleagues [29, 30, 34]. They proposed a probabilistic "bounded memory model" to describe and predict the human's adaptability in a simulated table-carrying task (see Figure 2(b)). This model is integrated with the robot's decision policy and allows the robot to adapt to the human partner in case the human insists on his/her policy. Otherwise the robot would guide the human partner towards the optimal policy. This led, in addition to improved task performance, also to higher satisfaction levels in the tested participants. While these works clearly show the importance of robot adaptation, progress along this line of research has proven challenging [4, 7]. A critical prerequisite for a system to be user-adaptive is to have sufficient and reliable information about the user [22]. While in HCI (human-computer interaction) with constrained input modalities

(computer mouse, keyboard) information about the user can relatively easily be inferred, the latter is less straightforward in natural, unstructured and dynamically changing environments [4]. State-of-the-art perceptual and reasoning capabilities of robotic systems are still limited and often augmented with additional information/feedback channels (e.g. buttons for providing feedback to the robot, as discussed in [26]) or specifically tailored to controlled interaction scenarios. This can as well be observed in the experimental setup of the table-carrying task of Nicolaidis and colleagues [29], in which human decisions were restricted to two options carried out by click-actions in the simulated environment (rotate table clock- or anticlockwise, see Figure 2(b)). In summary, while robot adaptation has been proven highly effective up to methods handling mutual adaptation in HRI, extensions to more realistic environments are thwarted by limitations of robot perceptual and reasoning capabilities.
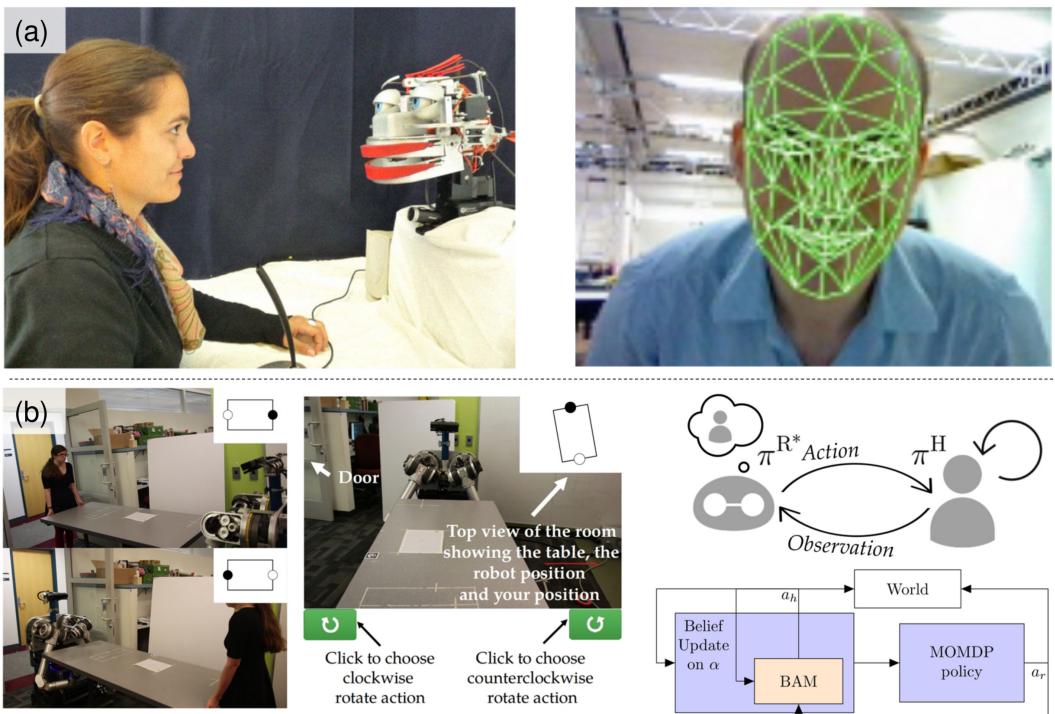
Figure 2 User-adaptive robotic systems and interfaces: (a) Experimental setup of a social interactive HRI task in which the effect of the robot head EDDIE mirroring / adapting to human facial expressions was investigated. [Reprinted from Gonsior et al. 2011 [31]] (b) Experimental setup of the simulated human-robot table-carrying task. Human participants could take decisions out of two options (rotate table left or right). To realize co-adaptation, the robot's actions were set to follow a decision policy linked to the "bounded memory model" which in turn predicts the participant's level of adaptability [Reprinted from Nikolaidis et al. 2016 [29] and 2017 [34]].

1.3. Novel approaches: Passive BCI-based assessment and adaptation of robotic systems and HRI

The approach of utilizing implicit measures from electro- or psychophysiological signals for augmenting HRI (or HCI) is a relatively young sub-field of BCI research and commonly referred as *passive BCI*. Passive BCIs aim for continuously providing information about the user's state (emotional responses, intentions, motivations) to the technical system - most of the time without the user being aware of it. The basic concept is to employ BCIs as a modality

for both understanding and augmenting a person's experience in interaction with machines [35, 36, 37, 38]. A series of neuronal correlates, potentially useful for such passive BCIs, have been explored and reported so far (see further information in Chapter 3, Section 3.1.1). This thesis focuses on neuronal correlates arising from *error-/performance-monitoring* processes in the brain, implicated in decision making, learning and adaptation, conflict and error-handling, both in collaborative as well as social interaction [39] (for further information, see Chapter 2, Section 2.3).

Error-/performance monitoring processes are also implicated in interaction between humans [40], and contribute to successful monitoring and prediction of other's behavior for establishing shared representations of goals and intentions to maintain or improve interaction performance [3]. There is increasing evidence that humans behave in a similar manner when interacting or collaborating with autonomous machines or robots [5]. Thus, in order to maintain or improve interaction performance in HRI, humans are likely not just to monitor and judge their own actions and decisions, but also those of the robots in order to predict future outcomes and be able to adapt swiftly to and interact efficiently with the robot.

Error-/performance monitoring processes have been shown to be manifested as characteristic signal deflections in the human EEG. These so-called ErrPs have been observed to be elicited by self-inflicted errors, processing of conflicting stimuli, as well as negative feedback and unexpected events (for further information, see Chapter 2, Section 2.4). While the study of ErrPs has a longstanding history in cognitive neuroscience, ErrPs have only recently been proposed as a valuable source of information about the human's expectations and subjective preferences towards robot behavior [41, 42, 43]. ErrPs are decodable from real time EEG signals and have features that are similar across tasks (for further information, see Chapter 4, Section 4.1.1). This makes them a suitable and task-independent complement to existing immediate and objective methods for validating and adapting HRI (or HMI/HCI in general). The approach is illustrated in Figure 3: EEG signals are recorded from the human partner during HRI; ErrPs are decoded upon occurrence of unexpected/erroneous robot actions observed by the human. Information derived from decoded ErrPs are subsequently either used to label robot actions for post-hoc validation, or utilized as feedback signal for online adaptation of robot behavior.

Therefore, the approach allows for both offline validation (binary assessment of individual events or robot actions) and online adaptation of robot behavior (feedback signal), and may as such serve as valuable complement to both the cognitive science as well as the engineering approach while simultaneously compensating for some of their drawbacks:

- *Assessment:* The approach goes in line with the proposition of Wiese and colleagues in 2017 according to which systematic validation of HRI must include objective neuro-cognitive measures [6]. Instead of using a variety of measures, the here proposed passive BCI focuses on one particular neuro-cognitive measure (e.g. ErrPs) which informs about
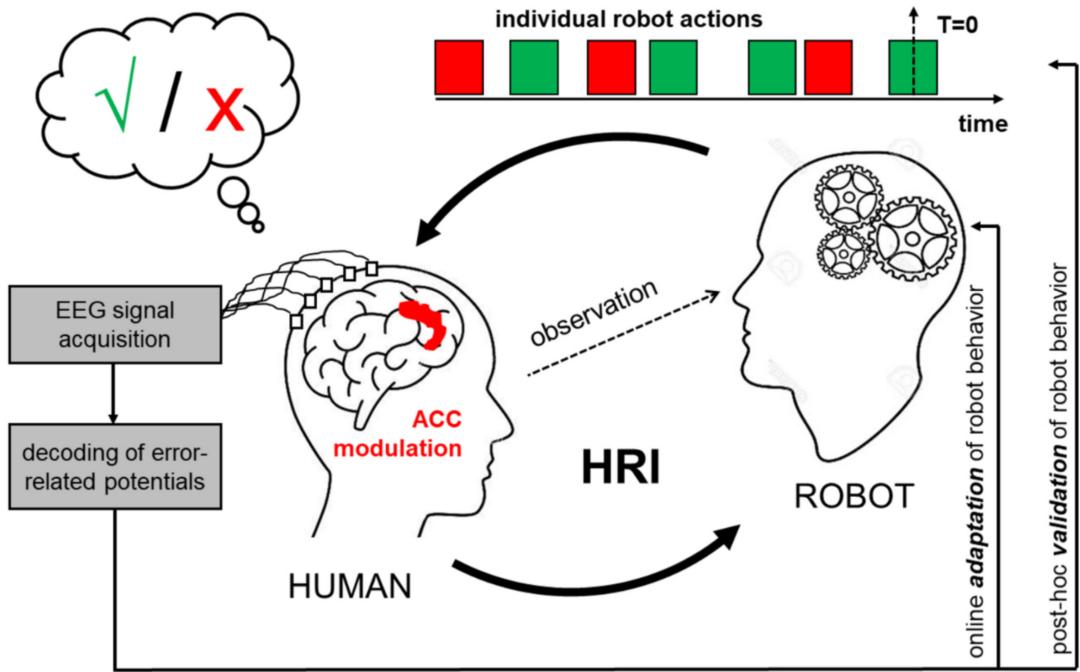
Figure 3 Passive BCI-based assessment and adaptation of HRI: In interaction with a robot, ErrPs decoded from the human interaction partner's EEG signals are used to assess/validate individual robot actions and/or to online adapt the robot's behavior during HRI.

human's expectations and preferences. This facilitates data analysis and interpretation, while preserving scientific grounding. Decoding ErrPs do not require participant's attention or awareness. Hence, there is no need to interrupt the interaction. Furthermore, ErrPs can be real-time decoded in single-trial upon occurrence of events observed by the participant, allowing as such for immediate binary assessment of individual robot actions.

- *Adaptation:* The feasibility to decode ErrPs in real-time, makes this approach useful also for online adaptation of robot behavior (feedback signal). Information about the human's assessment of individual robot actions can be used as feedback for correction of erroneous robot behavior or adaptations in line with human expectation. Online decoded ErrPs can furthermore serve as a complementary information channel in user-adaptive HRI to support the inference of human intentions and goals.

1.4. Contributions

The outcomes of this thesis result from three experimental HRI studies and one simulation-based study involving different robotic systems and HRI scenarios and resolve in four main contributions, with two addressing the research field of *assessment* and two addressing the research field of *adaptation*.

Assessment:

Contribution 1[1]: We conducted experimental study 1 to validate the feasibility of using direct brain measures from human interaction partners for the assessment of HRI. In an initial pilot study, we focused on the question as to what extent direct brain measures can be informative with regard to the nature of gaze contact events with a real humanoid robot. In particular, we were interested whether the intention to engage in interaction with a humanoid robot and the type of social role the human partner took in the engagement process can be decoded from modulations of oscillatory EEG activity. In a follow-up study with more participants, we performed a complementary data analysis on effects of brain activity modulations under different conditions of human-robot gaze-contact events. Resulting modulations were in line with the characteristics of ErrPs and could be related to (mis-)matching beliefs of the human partner about underlying goals of the robot. This gave rise to focusing on the effect of ErrPs in further works. The outcomes of these studies were:

- (1a) Confirmation of feasibility to decode from modulations of oscillatory EEG activity (i) the intention to engage in gaze-contact with a humanoid robot and (ii) the type of social role the human partner assumed in the engagement process.

- (1b) Explorative analysis of brain activity modulations in response to human-robot gaze-contact events suggesting involvement of performance monitoring processes in the engagement process of the proposed HRI scenario.

Contribution 2[2]: As a follow up to the initial studies, we conducted experimental study 2 addressing the feasibility of observing and decoding ErrPs in response to real robotic systems. Most previous studies on ErrPs focused on simple symbolic screen-based stimuli; a systematic investigation of whether and to what extent ErrPs arise from human observation of incorrect actions performed by real robotic systems was lacking. To address this, we investigated as to which extent the observability and decodability of ErrPs vary in response to human observation of incorrect robot actions compared to those arising in response to a simplified stimulus in procedurally identical experimental protocols. The outcomes of the study were:

- (2a) Confirmation of feasibility to observe and decode ErrPs in response to incorrect actions performed by a real humanoid robot.

---

[1] This contribution was in part published as a conference paper at the IEEE International Conference on Humanoid Robots in 2014 [44].
[2] This contribution was in part published as a conference paper at the IEEE International Conference on Humanoid Robots in 2016 [42] and as a journal article in the International Journal of Social Robotics in 2019 [45]. The EEG dataset collected during this study is available for download via: `https://github.com/stefan-ehrlich/dataset-ErrP-HRI`.

- (2b) Systematic cross-comparison of different stimuli (cursor, robot) revealed qualitatively similar brain activity modulations, suggesting a common underlying process related to error-/performance monitoring. Comparisons also revealed quantitative differences which suggest ErrPs to be dependent on the type of stimulus eliciting them.

Adaptation:

Contribution 3[3]: Experimental study 3 investigated as to which extent ErrP-based adaptations of robot behavior are possible when human interaction partners also adapt or dynamically change their expectations towards the robotic system. This is addressed in a closed-loop experimental setup in which participants were asked to guess from the gazing behavior of a real humanoid robot its underlying goal/intention while the robot's gaze behavior was adapted using online decoded ErrPs. This relaxation of interaction constraints - permitting mutual adaptation - is particularly important in scenarios in which adaptations of the human to the robot are expected or even necessary for successful interaction, such as in collaborative or social interactive HRI. The outcomes of this study were:

- (3a) Demonstration of feasibility to successfully mediate co-adaptation via online decoded ErrPs during HRI with a real humanoid robotic system.

- (3b) Demonstration of feasibility to utilize ErrPs as delayed reward signal for updating sequences of robot actions instead of single occurrences. This is an important prerequisite for transferring ErrP-based robot adaptation to more complex HRI where robot actions are likely to occur in rapid succession or being embedded in continuous behavior.

Contribution 4[4]: Our previous case study on human-agent co-adaptation left open a number of fundamental questions about the functionality of the approach in the context of varying human and technical factors. To address these questions, we propose a computational model describing participants' decision-making and learning behavior in the previously introduced HRI task and perform a series of model-based simulations to assess factors influencing co-adaptation. The outcomes of this work were:

- (4a) A computational model to describe and predict human decision making and learning within the context of experimental study 3. The proposed computational model allows the simulation and thus planning of future empirical studies in a resource-saving manner and thereby providing a means for accelerating progress along this line of research.

- (4b) Identification of the main technical and human factors influencing co-adaptation in ErrP-based HRI by model-based simulations of ErrP-based human-robot interaction with varying experimental conditions. The results provide guidelines for future technical im-

---

[3] This contribution was published as a journal article in Journal of Neural Engineering in 2018 [46]. The EEG dataset collected during this study is available for download via: `https://github.com/stefan-ehrlich/dataset-ErrP-coadaptation`.
[4] This contribution is in part covered in a conference paper, accepted for publication in the IEEE International Conference on Systems, Man and Cybernetics in 2019 [47].

provements of ErrP-decoders and alternative methods for ErrP-based robot adaptation and furthermore confirm that extensions of the approach to more complex HRI are feasible.

Datasets: Besides research contributions, this work resulted in several EEG datasets which were made publicly available via the following link: `https://github.com/stefan-ehrlich`. Over the last years, these datasets were used for teaching purposes and provided the basis for several student works in the study research questions complementary to those addressed in this thesis. In a similar manner, the data may further serve the BCI community in the future.

1.5. Thesis outline

This dissertation consists of seven chapters.

Chapter 2 - Background and foundations. This chapter introduces research areas/topics as well as important terminology relevant to this thesis. The review starts by providing a background on the history and recent developments in the field of BCI and introduces EEG as the neuro-recording method used in this work. Furthermore, the concept of error-/performance monitoring and its manifest in the human EEG, ErrPs, are introduced.

Chapter 3 - Initial studies: EEG-based assessment of intention to engage in HRI. This chapter presents the design and empirical results of experimental study 1 assessing the feasibility of using human EEG activity to derive human intention to establish gaze contact with a humanoid robot. The results showed decodable modulations of EEG oscillatory activity that could be related to the different beliefs of the participant about the underlying goals of the robot. A post-hoc analysis furthermore revealed characteristic modulations which resembled those of ErrPs. Consequently, the following works focused on ErrPs as a unitary electrophysiological effect.

Chapter 4 - A feasibility study for validating robot actions using EEG-based error-related potentials. This chapter presents the design and empirical results of experimental study 2 assessing the feasibility of using ErrPs to validate robot actions during HRI. The results confirmed feasibility to observe and decode ErrPs arising upon observation of erroneous / unexpected robot behavior, but also pointed towards challenges and limitations of the approach. This work served as a precursor to the subsequent study on ErrP-based mediation of human-agent co-adaptation.

Chapter 5 - Human-agent co-adaptation using error-related potentials. This chapter presents the development of a closed-loop BCI, and the design and empirical results of experimental study 3 exploring the feasibility to mediate co-adaptation in human-agent interaction

using ErrPs online decoded from a human interaction partner. The results confirmed feasibility and demonstrated successful ErrP-based human-robot co-adaptation.

Chapter 6 - Computational modelling of ErrP-based human-agent co-adaptation. This chapter presents the development of a computational model of human decision-making and learning to explore, within the context of the experimental study 3, the influence of varying human and experimental factors on human-robot co-adaptation. The results of model-based simulations attributed a more important role to technical than to human factors, and furthermore confirmed scalability of the ErrP-based human-robot co-adaptation approach to more complex HRI.

Chapter 7 - Conclusions and Outlook. This chapter provides a final overview of the contributions of this thesis, summarized in concluding remarks. Furthermore, an outlook of the future steps driving this line of research is provided.

2. Background and foundations

This chapter summarizes background, foundations and terminology relevant to this thesis. The survey starts with a general introduction to BCI in Section 2.1 and continues with an introduction of EEG in relation to alternative neuro-recording methods in Section 2.2. This is followed by a description of the neural basis of human performance monitoring in Section 2.3. The main effect relevant to performance monitoring, observable in the human EEG, ErrPs, are introduced in Section 2.4. The chapter ends with a summary in Section 2.5. The relevant literature is presented in chronological order to provide the reader a notion of the historic development of the covered fields up to the concepts and terminologies relevant to this thesis. More specific literature is presented as related work in Chapters 3-6 and cross-referenced in the respective sections of this chapter.

2.1. Brain-Computer Interfaces

A BCI, or brain-machine interface (BMI)[1], is a system that establishes a direct communication pathway between a brain and an external device. Traditionally a BCI is defined as a device that *"enables communication without movement, and must (1) rely on direct measures of brain activity, (2) provide feedback to the user, (3) operate online, and (4) rely on intentional control".* according to the definition of Wolpaw and colleagues (2002, [48]). A more recent definition according to Brunner and colleagues in 2015 defines a BCI as *"[...] a system that measures CNS [(central nervous system)] activity and converts it into artificial output that replaces, restores, enhances, supplements or improves natural CNS output and thereby changes the ongoing interactions between the CNS and its external or internal environment"* [49]. This is an extension to the traditional definition as it includes systems that do not rely on intentional control. Furthermore, it extends the definition to any device that exploits signals reflecting CNS activity, and not focusing on brain activity exclusively. A BCI consists of three parts (see Figure 4): (i) a measurement device that acquires signals containing information about the user's CNS activity, (ii) a decoder that translates in real-time these acquired signals into computer commands, and (iii) a device that executes these commands and provides feedback to the user. This way, BCIs bypass the normal pathways of communication with the external world, and as such provide users an additional and/or supplementary communication channel. One distinguishes here between *invasive* (based on surgically implanted electrodes or microarray recordings) and *non-invasive* (based on recording techniques that do not require implantation such as EEG) BCIs. While the history of BCI dates back to the invention of the EEG by Hans Berger in 1924 [50], the term "brain-computer interface" was only coined about 50 years later with the development of the first EEG-based BCI by Jacques Vidal in 1973 which employed VEP (visual evoked potential)s to estimate user's gaze direction [51]. Ever since, research on BCI has constantly gained momentum (for an overview the reader is referred to the following

---

[1] Although BCIs and BMIs follow the same principle, historically it has been established to use the term BCI for systems employing non-invasive and BMI for systems employing invasive neuro-recording methods.

historical and more recent survey articles [48, 52, 53, 54, 55, 49, 56]); but research in BCI is still relatively young with most works having been published throughout the last 15-20 years [49]. Early works were almost exclusively focused on *BCIs for communication and control* [48]. These are BCIs with the purpose of establishing an alternative communication channel for patients with severe motor deficits, as in ALS (amyotrophic lateral sclerosis), CLIS (complete locked-in-syndrome), or paralysis due to brainstem injury. EEG features were translated into commands, for example to perform binary selections of simple "yes"/"no" responses, multi-target selection for character spelling, over 2D cursor control, up to controlling wheelchairs, robotic prostheses, and exoskeletons. Prominent examples are the works of Nicolelis and colleagues [57, 58], as well as the work of Donoghue, Schwartz and colleagues [59, 60]. The second field of application emerged from the "closed-loop" nature of BCI systems: note that in BCI applications decoded mental commands are not only executed but also fed back to the user visually or via other perceptual modalities (see Figure 4). Closing the loop allows BCI users to monitor their own brain processes, which can trigger awareness, learning, recovery or even enhancement of brain functionality (neuroplasticity). *BCIs for neurorehabilitation* [53] have brought about novel treatment protocols for different neurological disorders, among others, but most notably for post-stroke motor rehabilitation (for a recent overview, see [61]). Prominent examples are works by Guan and colleagues [62] as well as the work by Daly and colleagues [63]. As such, early works on BCI were mostly focused on clinical applications for different patient populations. For healthy users on the other hand, contemporary neural recording techniques were too imprecise to be useful as an alternative input modality for HMI (e.g. non-invasive techniques such as EEG) or harmful and/or too expensive and therefore not justifiable in terms of cost-benefit trade-off (e.g. surgically implanted recording techniques). Along this line, a new class of non-invasive BCIs were proposed by Zander and colleagues in 2011, so called *passive BCIs* [36] with the specific purpose to serve the healthy population. These are BCIs which, in contrast to *active BCIs*, do not require voluntary or intentional control by the user, but rather decode mental states without the user necessarily being aware of it. The decoded information can be used to augment and as such enrich natural HMI with an additional information channel. It can be used for example for attention lapse detection or workload assessment during demanding or dangerous tasks, such as driving a car or operating heavy machinery [64]. Other examples are in the domain of gaming and entertainment [65, 66], and in the context of social HRI, such as in [67]. A detailed overview of passive BCIs for HRI is provided in Chapter 3, Section 3.1.1. For an overview of BCI applications and common BCI paradigms the reader is referred to Table 1 and Table 2, respectively, and furthermore pointed to a recent comprehensive survey by Abiri and colleagues [56].

2.2. Electroencephalography (EEG)

EEG was invented by Hans Berger in 1924 [50] and measures electrical fields generated by populations of synchronously firing neurons. It has been established ever since as one of the most widely used non-invasive technologies for monitoring brain activity and is currently used for diagnosis (e.g. epilepsy) and long-term monitoring (e.g. sleep monitoring) in
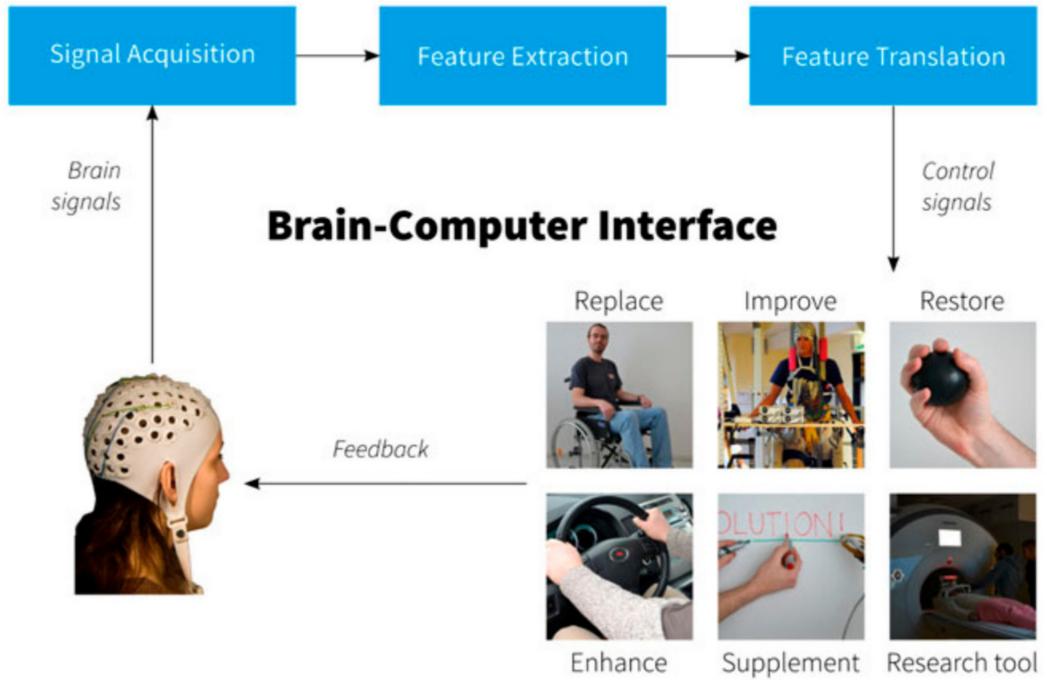
Figure 4 Brain-computer interface architecture: Principle of a brain-computer interface including application scenarios [Reprinted from Brunner et al., 2015 [49]].

Table 1 Overview of BCI applications

Purpose	Target group	Example applications
Communication and control	paralyzed patients and patients suffering from motor disorders	character spelling, cursor movement, wheelchair control
Neuro-rehabilitation	treatment of attention deficit hyperactivity disorder and motor disorders in stroke	attention training embedded in video games, robot guided movement training
Passive augmentation of HCI/HMI	healthy users	gaming, e-learning, HRI

clinical contexts; for research in human cognitive neuroscience; for BCI applications allowing paralyzed individuals rudimentary communication with their environment; and finally for BCI-based neuro-rehabilitation of sensorimotor disorders (e.g. in stroke). While EEG provides exceptional temporal resolution, it is limited in spatial resolution. EEG signals have furthermore a very low SNR (signal to noise ratio) and are often contaminated by non-brain activity artifacts such as due to environmental noise, electrode movement, and muscular activity. Some of the more recently developed neuro-imaging/recording methods can compensate for these drawbacks, such as fMRI, positron emission tomography (PET), fNIRS (functional near infrared spectroscopy) or invasive technologies, such as intracortical EEG (ECoG), and single- and multi-electrode recordings. From a BCI engineering point of view however, above-mentioned alternatives pose other serious limitations: they are very expensive or difficult to access, immobile / non-portable, and in part invasive which is accompanied by serious health risks. EEG technology on the other hand is comparably inexpensive and comfortable, non-

invasive, portable, and allows mobility of participants during experiments, therefore being still the technology employed in most BCI research.

Measurement principle: When neurons fire, electrical pulses are sent along the axon connecting several neurons (Figure 5(a)). These electrical pulses result in extracellular electrical fields that can - if strong enough - penetrate the layers of the cortex, cerebrospinal fluid, skull and skin and be picked up with surface electrodes positioned outside the skull (Figure 5(b)). Usually, the electrical field emerging of a single firing neuron is too weak to be measurable outside the skull, but if many neurons fire synchronously the summation of individual electric fields can generate electric dipoles strong enough to be picked up with surface electrodes. For this effect to work, axons need to be spatially aligned in parallel and neurons need to fire in synchrony, otherwise individual electrical fields may cancel out. This property of spatial parallel alignment of axons is given in most parts of the cortex [68]. The electric dipole generated has a scalp topographical distribution which can be derived from multichannel EEG recordings (Figure 5(c)). EEG allows to measure a series of electrophysiological effects which are described in detail below. These effects provide insight about different neuronal processes, such as implicated in stimuli perception and evaluation, sensorimotor planning, attention processes and decision making.
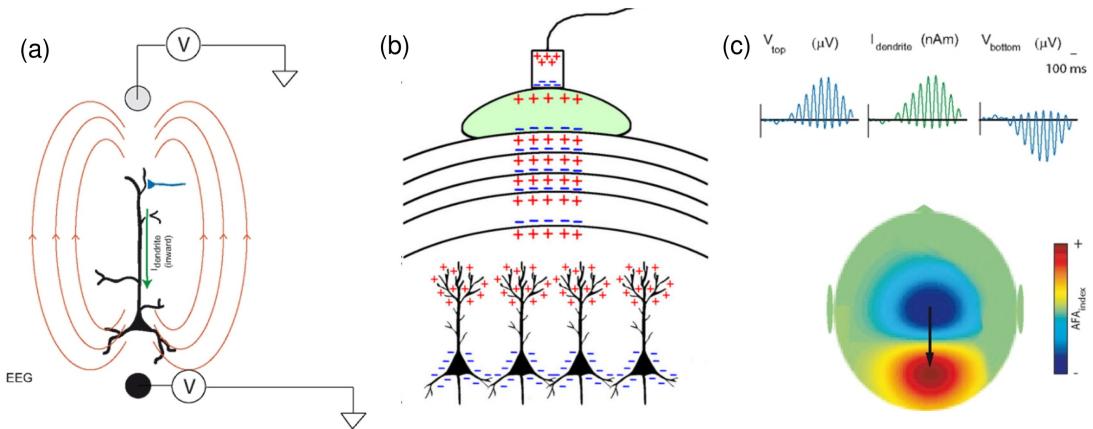
Figure 5 Measurement principle of EEG: (a) When neurons fire, electrical pulses are sent along the axon which creates an electrical field. (b) For this field to be strong enough to be picked up with surface electrodes, axons need to be spatially aligned in parallel and neurons need to fire in synchrony. (c) This generates an electrical dipole whose scalp topographical distribution can be derived from multichannel EEG recordings [(a,c) Adapted from Mazaheri & Jensen, 2010 [69]; (b) Adapted from Jackson & Bolger, 2014 [68]].

Slow cortical potentials: SCP (slow-cortical potential)s are low frequency negative or positive potential shifts arising in response to motor preparation, planning and execution. Among the most prominent is the *Bereitschaftspotential (or readiness potential)*, which is a negative signal deflection beginning 500-100 ms before self-initiated movement [70]. It has been shown that users can learn to voluntarily regulate these potentials after training using immediate feedback [71] which renders this effect suitable for BCIs for communication and control.

Oscillatory modulations: The frequency spectrum of EEG signals has been traditionally subdivided into several bands, *delta* (1-4 Hz), *theta* (4-7 Hz), *alpha* (8-13 Hz), *beta* (14-30 Hz), and

gamma ($>$ 30 Hz), and each of these bands have been associated with different physiological and/or mental states [72]. Among the most prominent are SMR (sensorimotor rhythm)s, which are oscillatory modulations in alpha- (also called mu-band) and beta-band over sensorimotor areas [73]. SMRs have been associated with voluntary motor planning, motor imagery and motor execution and manifest as modulations of the EEG frequency spectrum, so called ERS/ERD (event-related synchronization/desynchronization) [74]. As for SCP, users can learn to voluntarily modulate these SMRs which is why this effect has been employed in BCIs for communication and control [75]. Modulations of EEG oscillatory activity have also been associated with different levels of attention (drowsiness-alertness), mental workload, and task engagement [76] which is particularly relevant in the context of passive BCIs (see Chapter 3, Section 3.1.1). Furthermore, modulations of EEG oscillatory activity have been observed in response to observation of high-frequent repetitive stimuli, such as visual flickering stimuli, but also auditory and tactile stimuli [77]. These so-called steady state evoked potentials (SSEPs) manifest as sharp (low bandwidth) modulations of the EEG spectrum at frequencies corresponding to those of the stimuli.

Event-related potentials: ERPs are voltage deflections in the EEG time-locked to events or stimuli presented to participants and are assumed to arise from bursts (synchronization and desynchronization) of neuronal activity in specific frequency bands [78]. ERPs are usually reported as average waves across individual trials (single occurrence of stimulus presented to the participant) to overcome trial-by-trial variations and the generally low SNR of EEG signals. Averaging across individual trials of ERPs highlights specifically modulations that are phase-locked with respect to onset of the event; non-phase locked components are as such cancelled out. ERPs are composed of well-defined components (see Figure 6) which are positive and negative deflections of the ERP at specific latencies. Usually, earlier components ($<$ 100 ms) are associated with lower-level perceptual processes, whereas later components (up to 700 ms) are associated with higher-level cognitive, evaluative processes [78]. ERP components are termed according to their latency with regard to stimulus onset and peaking (e.g. P1 corresponding to a positive deflection at a latency of 100 ms post stimulus).

Table 2 Overview of common EEG-based BCI paradigms: Control schemes are divided into paradigms that allow for asynchronous self-paced operation of the BCI (e.g. continuous cursor control) vs. paradigms that allow for synchronous initiation of discrete commands at specific moments (e.g. selection of a target among several options). Further distinctions are made with regard to active (voluntary control based on movement intentions), reactive (voluntary control based on differential attention to external stimuli) and passive (passive measurement without requiring voluntary user actions) paradigms.

Paradigm	Description	Type	neuronal effect
Event-related potentials (ERP)	selective for variations in presented stimuli	reactive+passive; dependent; event-based	evoked + induced activities
Steady-state evoked potentials (SSEP)	selective for attended oscillating targets	reactive, dependent, continuous	evoked activities
Slow cortical potentials (SCP)	low frequency signal fluctuations due to movement preparation	active, independent, continuous	induced activities
Sensorimotor rhythms (SMR)	selective for voluntary motor planning, imagery and execution	active, independent, continous	spontaneous oscillatory modulations
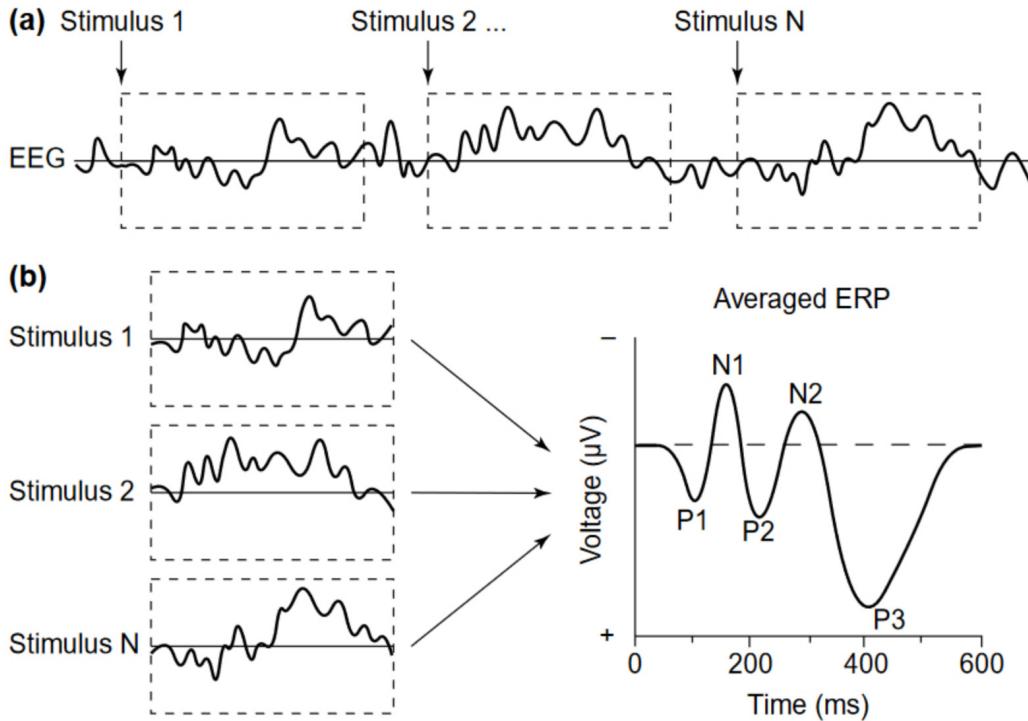
Figure 6 Extraction of ERPs from EEG signals: (a) Stimuli (1... N) are presented during EEG recording (b) ERPs are computed by segmenting EEG data following each stimulus (trial) and averaging across trials to obtain the averaged ERP waveform. ERP components are termed according to their latency with regard to stimulus onset and peaking (e.g. P1 - positive deflection at a latency of 100 ms post stimulus) [Reprinted from Luck et al., 2000 [78]].

2.3. Neural basis of error-/performance monitoring

Error-/performance monitoring[2] subsumes neural mechanisms involved in goal-directed behavior, decision making, planning and execution of tasks but also in flexible adaptive behavior and error handling [39]. Research on human and animal performance monitoring has a long tradition and dates back to the work of Rabbitt in the 1960s who investigated behavioral effects of error correction in choice-response tasks [79]. The first physiological evidence for performance monitoring in humans was reported by Bechtereva and Gretchin in the 1960s by means of measurements with implanted depth electrodes [80].

All studies, from fMRI to EEG and those employing invasive electrophysiology recordings, point towards several sub-regions in the medial frontal wall of the human brain as being involved in performance monitoring (see Figure 7). This is suggested to happen through a distributed network involving many other regions of the brain, but the electrophysiological responses to performance monitoring problems have been most consistently recorded in the anterior midcingulate cortex, which is part of the posterior MFC (medial frontal cortex) containing the pre-SMA (supplementary motor areas), also [39]. Several computational models for human performance monitoring in the medial prefrontal cortex have been proposed so far, as summarized in [81], and more recently in [82]. The majority of the proposed models employed

---

[2] Referred to as "performance monitoring" for the remainder of this thesis

a reinforcement learning-based computational framework, in particular temporal-difference learning in actor-critic architectures [83]. A comprehensive overview of computational models of error-/performance monitoring is provided in Chapter 6, Section 6.1.1.
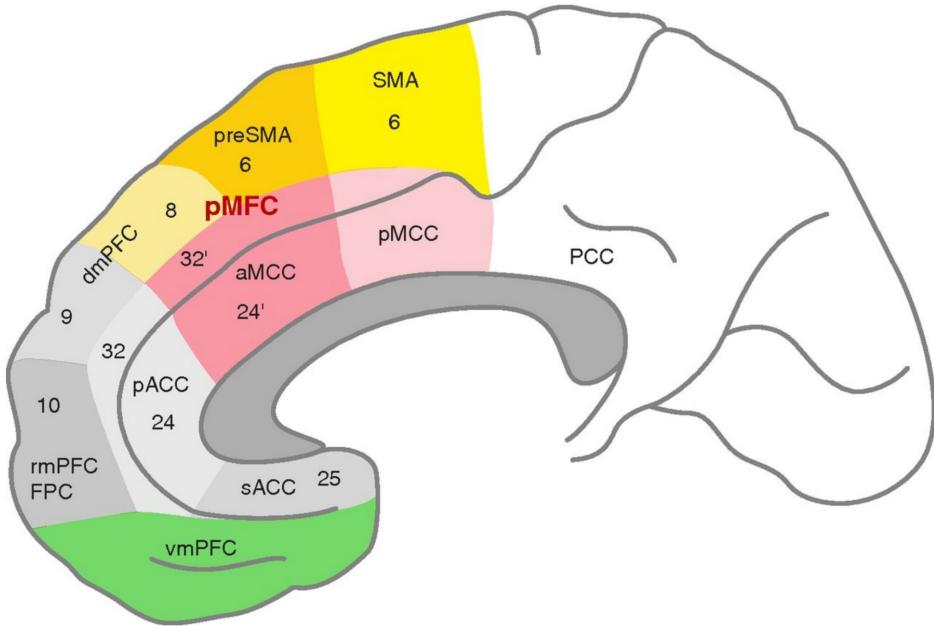
Figure 7 Brain regions implicated in human performance monitoring: Subregions of the medial frontal wall involved in performance monitoring; ACC (anterior cingulate cortex) and SMA [Reprinted from Ullsperger et al., 2014 [39]].

A multitude of experimental paradigms have been used to disentangle sub-functionalities of performance monitoring. Generally, participants are led to commit errors in a certain context, and can be given, or not, external feedback. One of the most prominent paradigms uses the (speeded) *CRT (choice reaction time task)* in which stimuli are presented in a continuous stream and participants are required to perform a binary decision on each stimulus (e.g. left vs. right button) as fast as possible [84]. A simplified version is the *Go/NoGo task*: one of the stimuli requires participants to make a motor response (go), whereas the other to withhold a response (no-go). In some cases, the experiment design allows the participant to correct post-error his/her initial response. Other paradigms, such as the *Eriksen flanker task* [85], use interfering stimuli to cause an error: the target is flanked by non-target stimuli which correspond either to the same directional response as the target (congruent flankers), to the opposite response (incongruent flankers) or to neither (neutral flankers). Such paradigms are used to assess the ability to suppress responses that are inappropriate in a particular context and focus on the investigation of selective attention. Another paradigm using interfering stimuli is the *Stroop task* [86]. Here, participants are presented color-printed color-words, with some words being printed in a color not denoted by the name, in which case there is a higher likelihood that participants commit an error when asked to name the color of the word. Another category of experimental paradigms subsumes gambling tasks in which the outcome of the choice is not estimable by the participant and after choice/response communicated temporally delayed by external feedback. An example is the *Iowa Gambling Task* developed by Bechara

and colleagues in 1994 [87] which is widely used in the study of cognition, emotions, and executive functions.

Performance monitoring processes are also implicated in social cognition and processing of affects [88]. Indeed, brain regions that have been associated with different aspects of social cognition highly overlap with those active in performance, error and outcome monitoring. A comprehensive meta-analysis by Amodio and Frith [40], addressing the perspective of understanding the role of the MFC in social cognition, proposed the separation of the MFC into three functional divisions: (1) the control and monitoring of action (see Figure 8, posterior rostral MFC), (2) tasks involving self-knowledge, and -reflection, person perception, and mentalizing (see Figure 8, anterior rostral MFC), and (3) the monitoring of outcomes related to punishment and reward (see Figure 8, orbital MFC). Amodio and Frith further proposed posterior rostral MFC and orbital MFC functionalities to be complementary in the process of monitoring and guiding behavior. They suggested the posterior rostral MFC as being involved in guiding behavior by monitoring the *value of different possible actions* and the orbital MFC being involved in guiding behavior in terms of the *value of possible outcomes*. In other words, if action selection is irrelevant for the value of possible outcomes, posterior rostral MFC is not activated, but anticipation of outcome value will be performed by orbital MFC, e.g. the anticipation of regret associated with a decision, as demonstrated in [89]. The location of the anterior rostral region of the MFC between the two regions (posterior rostral MFC and orbital MFC), suggests that it has access to information about both actions and outcomes. Recent evidence suggests that the more anterior regions of the MFC are involved in more complex social cognitive processing, e.g. when performance monitoring involves more abstract representations of the value of actions and outcomes [39].

2.4. Error-Related Potentials (ErrPs)

Research on performance monitoring was significantly boosted by the discovery of the ERP, measurable non-invasively via EEG [90, 91, 92, 93]. Ever since, findings from ERP analyses related to performance monitoring have been consistently supported by studies using other neuro-recording methods such as fMRI [94]. The effects are commonly subsumed under the term ErrP, but consist of several individual effects, also called ERP-components. The most prominent component, termed ERN (error-related negativity) (or Ne), is a frontocentral (typically observed at EEG recording channels FCz/Cz) negative potential deflection about [50...100] ms post error response, being elicited for example in speeded CRTs. It is preceded by a positive deflection, also frontocentrally, occurring about [-50...0] ms prior to the response and another positive deflection around [100...250] ms post response (see Figure 9(a), left). Another prominent component is the FRN (feedback-related negativity), characterized by a stronger negative deflection occurring about 250 ms after the presentation of external negative, compared to positive, feedback. Finally, the N2 (or N200) component is time-locked to the presentation of conflicting stimuli, for instance novel or mismatching stimuli [95], and precedes the participant's response. It is characterized by a stronger negative deflection around 300 ms
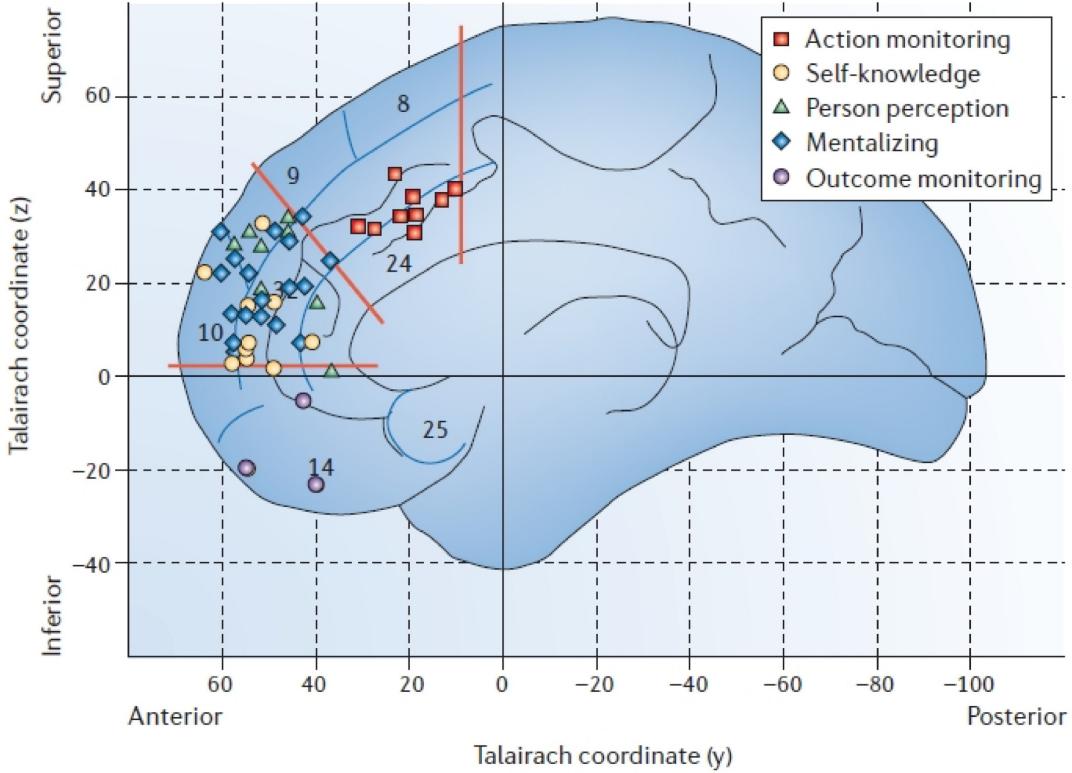
Figure 8 Human performance monitoring in social cognition: A meta-analysis of brain areas in MFC implicated in social cognition tasks involving self-knowledge, person perception, and mentalizing, activate areas in the anterior rostral MFC. By contrast, activations from action-monitoring tasks occur in the posterior rostral region of the MFC, and tasks involving the monitoring of outcomes occur in the orbital MFC [Reprinted from Amodio & Frith, 2006 [40]].

post incongruent, compared to congruent stimuli. It has been suggested to reflect the level of conflict and to be strongly interrelated with the post-error response ERN [96]. In all three cases (ERN, FRN, N2) modulations are spatially located over frontocentral areas (see Figure 9(b,c)). Moreover, all of these deflections manifest most prominently in theta band (5-7 Hz), but it is still under debate whether from theta phase-locking and/or power modulations [97]. Owing to the similar topographical distributions, the ERN and FRN are understood to arise from the same underlying processing system but under different circumstances [98]; functional equivalence of ERN and N2 have also been suggested [99]. Source localization and fMRI studies consistently pointed the origin to be in the medial PFC (prefrontal cortex), specifically, in the ACC and the SMA [39]. These areas in turn have been implicated in conflict and error coding, social cognition, and effortful control [82]. In summary, ErrPs are implicated in the monitoring of internal and external events and arise from and as such encode deviations of predictions about observable events, e.g. differences between actual and expected outcomes [100, 101]. Besides the above-mentioned early components at latencies around 200 ms post stimulus, ErrPs have also been associated with components at later latencies: a positive deflection around 300 ms (P3 or P300 component) and a negative deflection around 400 ms (N4 or N400 component) both over fronto-central and fronto-parietal recording sites, respectively [102, 103, 104]. The N200 in conjunction with the P300 component is sometimes referred as the *N2-P3-complex* [102, 103, 104, 105]. An additional late positive component (P600, around 600 ms post stimulus) has also been frequently reported in the context of error processing in

CRTs [106] and the experience of syntactic and semantic anomalies in language comprehension [107, 108]. Further information about ErrPs is provided in the respective chapters of this thesis: An overview of single-trial decoding of ErrPs from ongoing EEG signals is provided in Chapter 4, Section 4.1.1; an overview of research works using ErrPs for adaptation of HMI is provided in Chapter 5, Section 5.1.1.
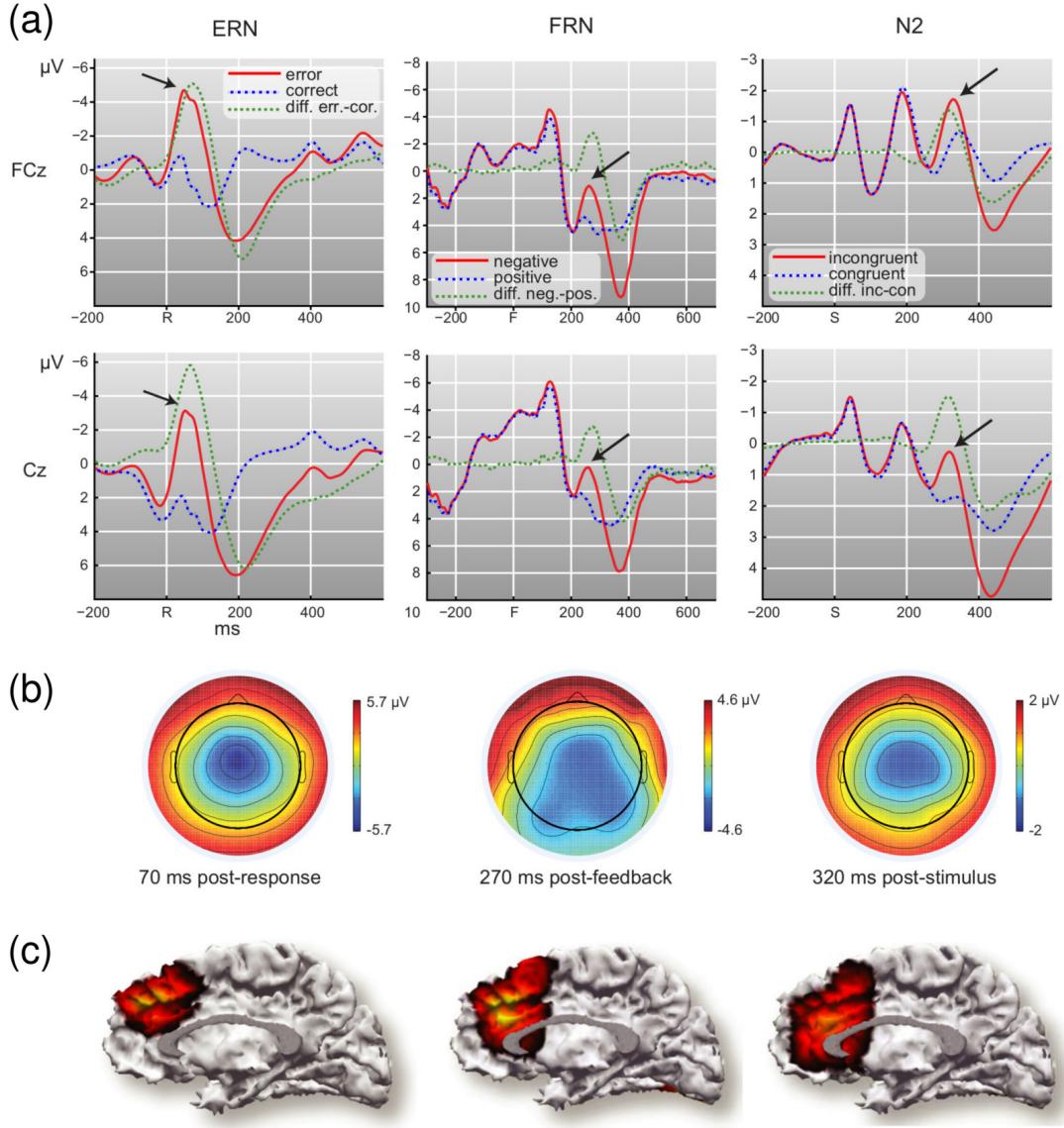
Figure 9 Manifestation of error-related potentials in EEG and fMRI data: (a) event-related potentials (ERP) in response to self-inflicted errors (ERN, response-locked), negative feedback (FRN, feedback-locked), or pre-response conflict (N2, stimulus-locked) at EEG electrodes FCz and Cz. (b) corresponding topographies of difference waves from (a). (c) source localizations of ERN, FRN, and N2 [Adapted from Ullsperger et al., 2014 [39]; original from Gruendler et al., 2011 [109]].

2.5. Summary

BCI research and related applications are still in its infancy. Despite the fascination this field has sparked both on scientific communities and the broad public, the main body of research has mainly been published during the last 20 years [49]. Early works were almost exclusively focused on patient populations suffering from severely limited mobility and communi-

cation abilities. More recent works investigated the usability for neuro-rehabilitation and also proposed passive BCI-based applications for the healthy population. Therefore, early works almost exclusively focused on motor related neuronal correlates, whereas recently, the BCI community has gained increasing interest in neuronal correlates related to error-/performance monitoring. The corresponding manifestation in the EEG, ErrPs, may constitute valuable information about the user's state during HMI and thus be a useful feedback signal to augment HCI and HRI. While there are examples of both invasive and non-invasive BCIs, the main recording technology is still EEG, which provides insight about neuronal processes in form of modulations of oscillatory activity and modulations of time-locked responses to external events (e.g. ERPs), the latter being particularly relevant in the context of error-/performance monitoring. EEG is comparably inexpensive, non-invasive and thus risk-free, as well as portable and thus more readily deployable in practical applications than alternative neuro-recording methods.

3. Initial studies: EEG-based assessment of human intention to engage in HRI

The current limited ability of humanoid robots to interpret social cues conveyed by humans restricts fluency and naturalness in HRI. We propose a method to derive directly from the brain activity of the human interaction partner two aspects of social engagement: (1) the *intention to initiate gaze-contact* and (2) the distinction between the observer being *initiator or responder* of an established gaze contact between human and robot. We propose these measures to constitute valuable input to humanoid robots for deciding *when (timing)* and *how (social role)* to engage in interaction with a human. We developed an interaction paradigm with the humanoid robot *iCub*, and recorded the associated brain activity via EEG, in a pilot study and a follow-up full study. The pilot study focused on the analysis and modeling of oscillatory EEG activity and revealed biologically plausible brain activity patterns for both processes of social engagement. By using SVM (support vector machine) classifiers with RBF (radial basis function) kernel we showed that these patterns can be modeled with high within-participant average accuracies of $Acc(accuracy) = 80.4\%$ for (1) and $Acc = 77.0\%$ for (2). The follow-up study focused on the analysis of ERP time locked to the moment of gaze-contact between human and robot. Results showed the ERP N200 and P350 components to have different characteristics depending on which partner, the human (YI) or the robot (RI), initiated the interaction. These findings demonstrate that processes of mismatch detection as well as the evaluation of social context depend on the participant's belief about its social role during the gaze-contact event. The observed spatio-temporal activation patterns resembled those typical for ErrPs which suggests the main effect to be related to performance monitoring processes.

3.1. Introduction

In everyday lives, humans are embedded in rich social environments. It is typical of humans to seek social contact, which is intrinsically very rewarding [110, 111]. In this respect the willingness or intention to be engaged is a crucial aspect of social interaction. Humans are capable of expressing and detecting this intention by many subtle and mainly non-verbal social cues, such as touch, gestures, and body posture. Gaze is one of the most important social signals, as it is often involved in initiation of social contact and engagement [112].

Humanoid robots in contrast are to date still severely limited in this respect. It is perhaps mostly due to the subtleties of these signals that make their interpretation based on information obtained through visual, auditory and tactile sensors so challenging. This issue is

---

particularly thwarting the applicability of humanoid robots in areas where social interaction with humans is crucial, such as in elderly- and healthcare, household robotics, and social robotics in general [113, 114].

3.1.1. Related work on passive BCIs for HRI

Passive BCIs [35, 36, 37, 38] propose the utilization of implicit measures from electro- or psychophysiological signals for augmenting HRI (or HCI) with information about the current user's state (emotional responses, intentions, motivations). A series of performance-relevant variables with associated neuronal correlates, potentially useful for such passive BCIs, have been explored and reported so far and culminated in three main strands: attention level / task engagement [67, 115], mental workload [116, 117, 118], and affect [119, 120] (see [121] for an overview on affective BCIs). These correlates are mostly based on characteristic oscillatory modulations of the EEG frequency spectrum or changes in hemodynamic activity recorded via fNIRS. While most works presented so far were mainly focused on exploring the principle usability of these neuronal correlates in offline (open-loop) experimental studies (e.g. [119, 116, 122, 120, 118]), research works, employing passive BCIs in online closed-loop experimental studies are less frequent. Most such works employing closed-loop BCIs are in the field of gaming and entertainment as well as e-learning: For instance, Ewing and colleagues [115] showed in 2016 that the level of game demand could be adapted based on modulations of EEG frontal theta and parietal alpha power to maximize the player engagement. Similarly, Szafir and Mutlu proposed in 2013 the "adaptive content review" technique, which adapts the educational content presented to the students based on their attention-levels measured via EEG and demonstrated improved student's recall abilities by 29% as compared to a controlled baseline [123]. Research works in the field of (social) HRI, using brain-activity based adaptation of real robotic agents or machines are even more scarce. Szafir and Mutlu presented in 2012 [67] a study in which a real humanoid robot appeared as a story narrator. Based on EEG-based measures of attention / task engagement, the robot adapted its level of gesticulation, mimics and gazing during storytelling, which in turn positively influenced the level of details of the story participants were able to recall after the experiment. Strait and Scheutz presented in 2014 a study [120] in which they showed fNIRS-based PFC hemodynamic changes reflecting a person's affect (e.g. aversion) towards a real humanoid robot. The results of their study however were too variable at the individual level to be useful as a real-time measure for adapting robot behavior.

In summary, while the concept of passive BCIs has been around for almost a decade, most research works were so far focused on validating EEG- and fNIRS-based neural correlates in open-loop experimental studies. Research works exploring their usability in closed-loop setups using passive BCIs are rather scarce. Similarly, the exploitation of passive BCIs for augmentation of HRI is widely acknowledged as a logical merge of BCI and HRI [36], but only a few practical examples demonstrating the approach were so far presented.

3.1.2. Aim of the work

This work explores the feasibility of estimating from ongoing EEG signals the participant's intention to engage in interaction with a real humanoid robot. This work shall serve as a precursor to the development of a passive BCI facilitating HRI by providing the robotic system additional information about when and how to engage in interaction with the human partner. This research is thus focused on two basic aspects of social engagement (see Figure 10):

- The *intention to initiate gaze-contact* for entering into social engagement with others, in our case the humanoid robot *iCub*. This measure is crucial for the natural and efficient onset of interaction [112]. The ability to sense this measure would give a humanoid robot a means of deciding *whether and when* to engage in social interaction with a human.

- The distinction between the observer being *initiator or responder* of the gaze contact. This distinction is crucial for the pleasantness and shape of the further course of interaction [124]. The ability to sense this measure would give a humanoid a means of estimating its *social role* during the interaction and adapt its behavior according to the expectations of the human interaction partner.
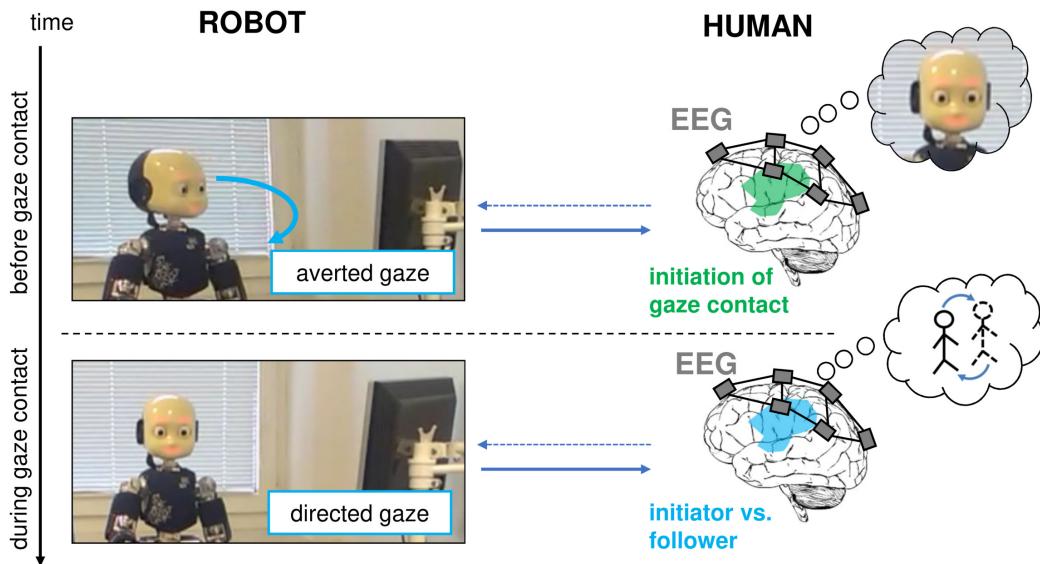
Figure 10 Two processes of social engagement: initiation of gaze-contact for onset of interaction (top) and distinction between initiator versus follower when gaze-contact is established (bottom).

This chapter is organized as follows: The experimental paradigm and design is explained in Section 3.2. The pilot study including experimental setup, data analysis, results and discussion are detailed in Section 3.3. Experimental setup, data analysis, results and discussion of the full study are detailed in Section 3.4. Section 3.5 summarizes the chapter.

3.2. Methods: experimental study 1

3.2.1. Experimental paradigm

The purpose of this experiment was to elicit in human participants electrophysiological activity related to two aspects of social engagement: the *intention to initiate gaze-contact* and the distinction between the observer being *initiator or responder* of the gaze contact. In order to evoke and capture these responses, our experiment made use of a *belief manipulation* that intended to make participants believe that they were able to willfully influence the robot's behavior (provoking the robot to engage in social interaction with them). Once captured, the respective EEG patterns can be identified and translated into predictive models ultimately applicable in social HRI in line with the passive BCI approach. Note, that this experiment was not a passive BCI approach, but a preliminary study to collect data with the purpose of developing a passive BCI.

3.2.2. Experimental design

In the experiment, participants were asked to interact with a humanoid robot in form of gaze-contact events (further referred as trial). Participants were seated in front of a humanoid robot and asked to keep their gaze on the robot and were not requested to shift their gaze towards or away from the robot (see Figure 13). Gaze shifts were performed by the robot, only, and conducted under two experimental conditions. These conditions differed only with regard to the instructions provided to the participants:

- "YOU INITIATE" interaction (condition **YI**): The participants were told that in this condition their *intention to initiate gaze-contact* with an interaction partner would be decoded from the EEG signals. They were also told that the decoding algorithm used worked in real-time and influenced the robot's behavior, depending on their willingness to engage in social interaction with the robot. Whenever the beep tone occurred in the YI blocks participants would be able to willfully influence the robot's behavior (provoking the robot to look at them).

- "ROBOT INITIATES" interaction (condition **RI**): Participants were instructed that in this condition the above-mentioned information would not be provided to the robot (connection was turned off). The robot would be rather entering the social interaction on its own *"intention"*. Participants were instructed to, upon the occurrence of the beep tone, wait for the robot to gaze at them.

Unbeknownst to the participants **no such an algorithm existed**, hence both conditions were technically identical; the gaze-contact always followed within a random time in between 5 to 8-s. With this belief-manipulation we aimed to evoke brain activity patterns specific for the intention to engage in social interaction.

A single trial (see Figure 11) consisted of an initial 5-s rest period, followed by a 5 to 8-s task period, in one of the two conditions "YOU INITIATE" (YI) / "ROBOT INITIATES" (RI), and then by a 3-s long gaze-contact event. The trial ended with another 2-s resting period. Each trial

has thus a duration of 15 to 18-s. The actions performed by the robot during each trial are further detailed in the next section.
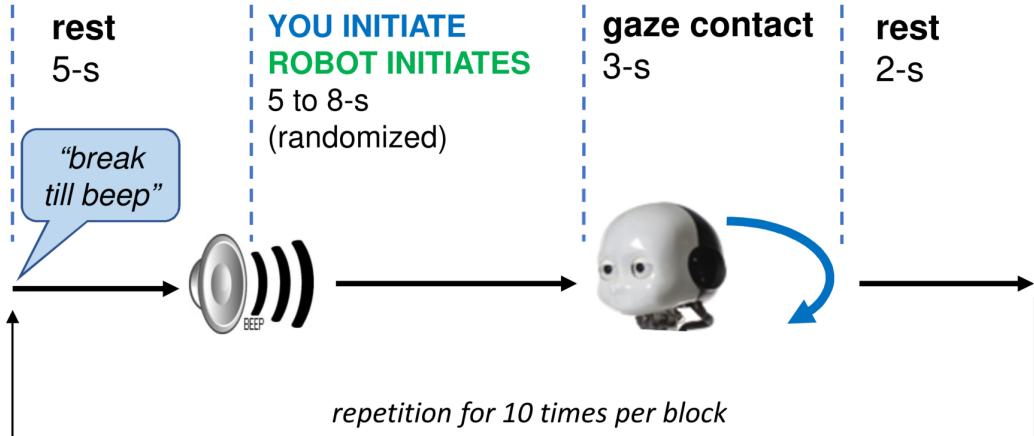
Figure 11 Engagement study trial structure: One trial consisted of an initial 5-s rest period, followed by a 5 to 8-s task period, in one of the two conditions "YOU INITIATE" (YI) / "ROBOT INITIATES" (RI), and then by a 3-s long gaze-contact event. The trial ended with another 2-s resting period. The trials were repeated 10 times per block.

Trials were arranged in 12 blocks of 10 trials each (10 gaze-contact events per block, see Figure 12 depicting the block design, and Figure 11 depicting the structure of a single trial). The two conditions (types of interaction with the robot) alternated randomly from block to block, such that no more than 2 consecutive blocks would belong to the same condition. Participants could self-pace the start of each new block: After pressing a specific key on a keyboard they first received verbal instruction from the robot speech synthesis system about the type of interaction in the next block ("The next block will be **your** turn!" for the YI condition; "The next block will be **my** turn!" for the RI condition). After pressing another key, the next block started. The total duration of the experiment was approximately 45 minutes.
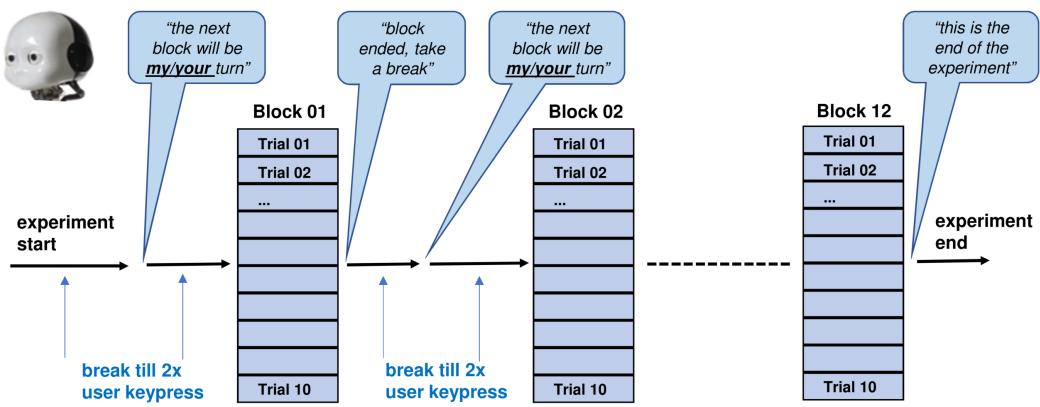
Figure 12 Experimental design and protocol: One block consisted of 10 gaze-contact events (trials) in one of the two conditions "YOU INITIATE" (YI) / "ROBOT INITIATES" (RI). Participants were informed verbally about the type of interaction prior to the start of the next block.

3.2.3. Stimuli and apparatus

The experiment took place in a quiet room which was partitioned into two sections by means of a movable wall. Experiments were conducted on the right side of the room, and the ex-

27

perimenter sat and monitored the session from the left side. Participants sat in a comfortable chair approximately 2 m away and facing the humanoid iCub (see Figure 13). iCub is a 53 degrees of freedom humanoid [125] which has an in-build control unit communicating with an external workstation via a local network based on TCP/IP (Transmission Control Protocol/Internet Protocol). The robot was arranged to be standing behind a table with its torso facing the participant; in the averted gaze state the robot was looking at a computer screen to its left (realized with a neck angle of $42°$, see Figure 13, top). The robot performed one action per trial, which consisted of two sequential head-movements: First, the robot turns its head to $0°$ neck angle with an angular velocity of $50°/s$ and furthermore lowers its gaze by $-5°$ in order to meet the gaze of the participant (parameters were calibrated prior to the start of the experiment and kept identical for all participants, both in the pilot and the full study). The robot keeps gazing at the participant for a 3-s period and subsequently moves its head back to the initial position. For the robot control and thus the implementation of the experiment protocol we used Yarp [126] and iCub [127] libraries. Furthermore, the robot was equipped with a speech synthesis system by including the package iSpeak, which acquires sentences over a Yarp port and lets the robot utter them. For the entire duration of the experiment, the robot's facial expressions were set to happy.

The workstation, besides controlling the robot, was further programmed to send, in specific moments of the experiment protocol, event-triggers via LPT (line printer terminal) to the EEG amplifier. These triggers appeared in the EEG data as event-codes facilitating later segmentation of the data. EEG data was acquired and transferred via USB (universal serial bus) to a separate PC (personal computer), where it was recorded. The EEG amplifier is battery-driven and was located on a tray near the participant. Participants were given earplugs for minimizing auditory distractions. Speech and beep indications were played back via Logitech stereo desktop speakers with an appropriate loudness. The participants were asked to sit still, but comfortably during the experiment and try to move as little as possible.

3.3. Pilot study, data analysis, and results

3.3.1. Participants and data collection

Six healthy participants took part in the experiment. During EEG data inspection we noticed large DC (direct current) drifts in participant s01 and therefore decided to remove this dataset from further analysis. Thus, data from 5 participants were analyzed (age: 26.0 ± 2.0, all males). Educational background was predominantly in engineering and computer science majors (4 out of 5 participants). Prior experience and familiarity with humanoid robots varied widely among the participants and scored 3.6 ± 1.9 on a scale of 1 "non-familiar" to 7 "familiar". All participants were informed about the experiment prior to its conduction, and agreed to participate and having their data acquired by signing a consent form. Each participant was paid an honorarium of 15 EUR.

EEG data was acquired with a Brain Products actiChamp amplifier equipped with 32 active
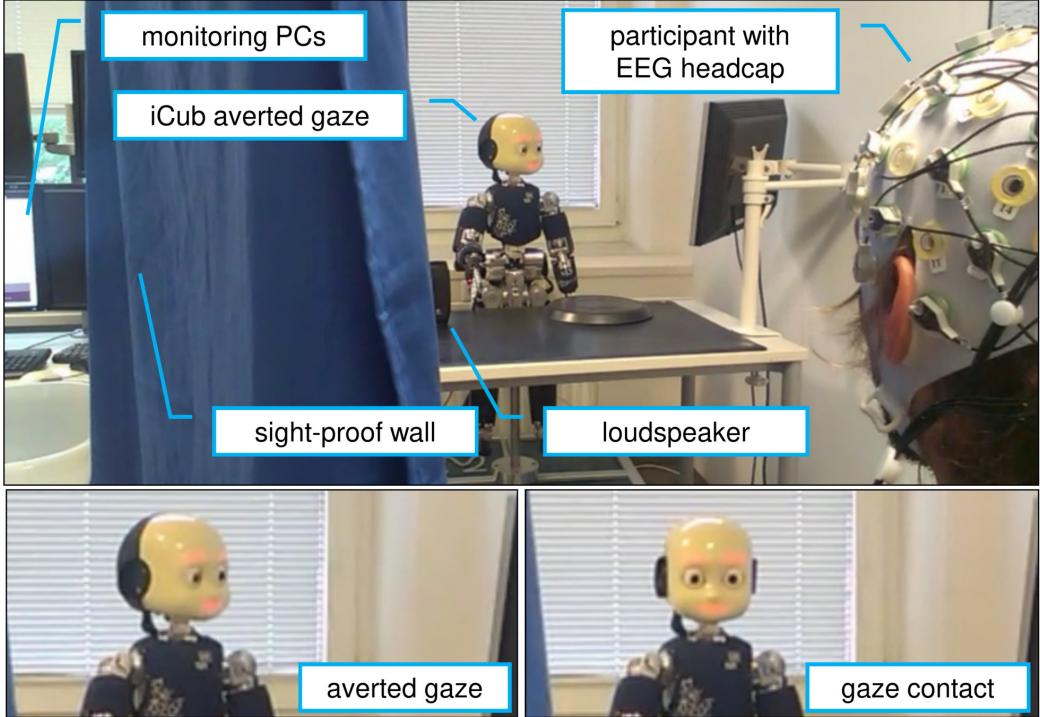
Figure 13 Top: experimental setup with participant. Bottom: iCub attending computer screen (averted gaze) versus iCub establishing gaze contact with participant.

EEG electrodes[1] arranged according to the international 10-20 system [128]. All leads were referenced to Cz and sampling rate was set to 500 Hz. The impedance levels of all leads were kept below 10 kΩ.

The participants were first instructed (verbally and by means of written instruction) about the experimental setup and handed a questionnaire about personal details. The participants were further instructed to sit still during the experiment, keep their gaze on the robot and avoid any unnecessary movements, such as head shifts. To test whether participants believed the experiment instructions they were given another questionnaire after the experiment that included questions e.g. about how well they thought the robot responded to their intention to engage in gaze-contact.

3.3.2. Analysis and modeling of EEG oscillatory activity

Data preprocessing: Data was preprocessed with a commercial software, the Brain Products Brain Vision Analyzer. Brain Vision Analyzer comes with a set of in-built signal processing functions particularly relevant for EEG data preprocessing and visualization. The following processing steps were carried out: (1) Re-referencing to CAR (common average reference) to reduce common signal contamination due to external noise sources; afterwards the original reference channel Cz was reconstructed from interpolations of neighboring channels. (2)

---

[1] EEG channel labels: Fp1, Fz, F3, F7, FT9, FC5, FC1, C3, T7, TP9, CP5, CP1, Pz, P3, P7, O1, Oz, O2, P4, P8, TP10, CP6, CP2, C4, T8, FT10, FC6, FC2, F4, F8, Fp2, Cz.

Downsampling to 250 Hz to reduce processing time in all further steps. (3) Bandpass filtering[2] using a Brain Vision Analyzer in-built Butterworth zero phase IIR (infinite impulse response)-filter with a low cutoff around 0.5 Hz and a high cutoff around 70 Hz (12dB/octave), both to remove slow DC drifts and high frequency noise. (4) Manual removal of signals recorded between blocks and not containing relevant information. (5) The last steps had the purpose to remove eye-movement related EEG artifacts from the data using ICA (independent component analysis): The data was first transformed to the component-level using 512 steps Infomax extended ICA. Then, components which were most probably associated with eye-blinks were manually selected and removed. Finally, the data was back-transformed to channel-level using inverse ICA. Afterwards, the data was exported for further analysis in MATLAB®.

Data segmentation: According to the purpose of the experiment, we aimed at exploring and modeling the electrophysiological patterns of two specific processes of social engagement with a humanoid: (1) the *intention to initiate gaze-contact* and (2) the distinction between gaze contact based on whether the human was the *initiator* or the *responder* of gaze contact. For both analyses we extracted and compared different segments of the EEG data:

- *Intention to initiate gaze-contact:* Patterns associated with the intention to initiate gaze-contact were sought by comparing EEG signals recorded during trials against signals recorded during a baseline condition, when "gaze-intentions" were assumed to be absent. The baseline signal was determined to correspond to the last 3-s of the resting period of the RI condition. With this processing step we obtained 60 750-samples-long *baseline* segments per channel and participant. Based on our experimental design we assumed to find brain activity patterns for the intention to initiate gaze-contact directly after the beep tone in those trials which were related to the YI condition. In order to avoid biases in the data we decided to extract trial segments with the same duration as the baseline segments. We assumed the patterns to be best developed some time after the beep and decided to extract 3-s segments starting 2 s after the *intention* segments.

- *Initiator versus responder gaze contact:* For this analysis we contrasted the 3-s periods of gaze contact of both conditions. Thereby we obtained 60 *responder* segments from the RI condition and 60 *initiator* segments from the YI condition, both 750 samples long.

Feature extraction: Feature extraction was carried out identically for both above described comparisons. As this study aimed at exploring modulations in EEG oscillatory activity, each segment was filtered into five standard EEG frequency bands that have been associated with different cognitive processes [72]: theta (4-7 Hz), low alpha (7-10 Hz), high alpha (10-13 Hz), beta (14-30 Hz) and gamma (30-47 Hz), with 2nd order zero-phase Chebyshev IIR-bandpass filters. In preliminary analyses we noticed distinct differences within the alpha band and therefore subdivided it into two separate sub-bands (low and high alpha). For each filtered segment we then computed the log-variance as a measure of spectral power which is associated with

---

[2] Due to the use of active electrodes, which reduce environmental artifacts, such as power line interference [129], no additional notch-filter was employed for further temporal filtering.

band-specific cortical activation in EEG [74]. For each segment/trial we thereby obtained one value as a means of cortical activation in a specific frequency band and channel. The features were then concatenated into one vector resulting in 160 features (32 channels x 5 frequency bands). The trials were then labeled according to the above described comparisons.

Data modeling and classification: SVMs were employed to create models for pattern discrimination, and were implemented with the LIBSVM library by Chang and colleagues [130]. SVMs were developed for solving binary classification problems and model an optimal hyperplane for discriminating both classes. They are particularly powerful as they find the best tradeoff between good generalization by simultaneously maximizing the performance and minimizing the complexity of the model. Moreover, SVMs can be used with kernel functions that map the features into high-dimensional space in which non-linearly separable data can be discriminated by linear hyperplanes. We decided to employ a RBF kernel, which has been reported as the most suitable kernel for SVM-based EEG-signal classification problems [131]. Furthermore, we performed an exhaustive search to find optimal values for the learning parameter C and the kernel parameter γ. For SVM training we used all 160 features and evaluated the models in two ways:

- 5-times-5-fold participant-individual CV (cross-validation) (**CV**): This analysis allowed us to obtain an estimate of how well participant-specific classification models would perform in classifying unseen data when being calibrated with different data of that same session and participant. The data of one participant were partitioned into 5 folds, 4 folds were used for training the SVM model and the remaining fold was used for testing. The folds were then shuffled until each fold had once been used for testing. The whole procedure was repeated 5 times. The 5x5 results were averaged and reported as Acc according to Eq. (3.1). This procedure was carried out for each participant individually.

- Leave-one-participant-out validation (**L1O**): This analysis allowed us to obtain an estimate of how well a classification model would perform in classifying unseen data of one participant when being calibrated with data of all remaining participants. The data of 4 participants were concatenated and used to train the SVM model. The data from the remaining participant was used for testing. The procedure was repeated until the data from all individual participants had once been used for testing. Identical to CV, the results are reported as Acc according to Eq. (3.1).

$$Acc = \frac{\text{number of correctly predicted trials}}{\text{total number testing trials}} \times 100\% \qquad (3.1)$$

3.3.3. Results of electrophysiological modulations and single-trial classification

Modulations of oscillatory activity: To facilitate data interpretation, the grand average over all trials and participants was computed and visualized for all 160 features in topographic plots

(see Figure 14). For the comparison *intention versus baseline* the strongest effect was a decrease of anterior and posterior low alpha-power (see Figure 14(a)). In beta- and gamma-band we could observe a hemispheric lateralization with increased anterior gamma-power on the left side and decreased central beta-power (see Figure 14(b)). When comparing *gaze-contact initiator versus gaze-contact responder* an effect on low alpha-power was not observed, both conditions resulted in relatively similar alpha-activities. However, beta- and gamma-band were even more strongly pronounced than in the previous comparison (see Figure 14(c)). Moreover, we observed power-decreases over left and right motor cortex (central) in high alpha-band (see Figure 14(d)).
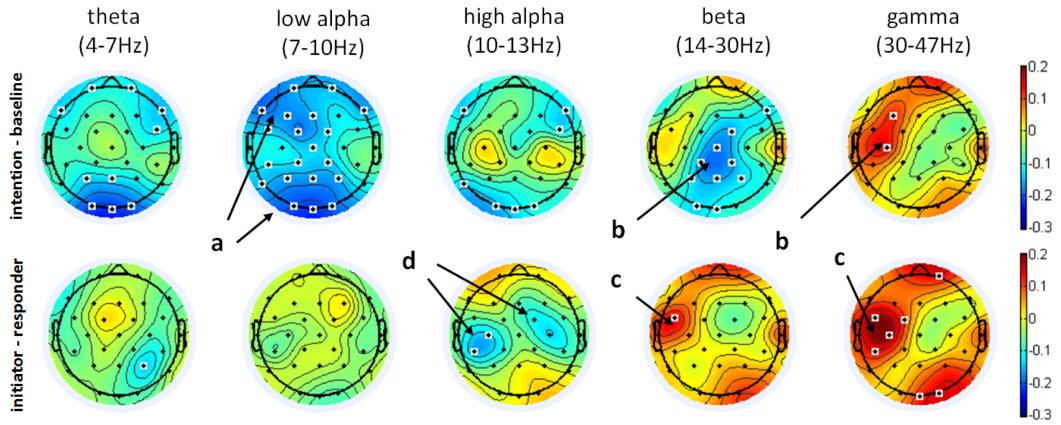
Figure 14 Results of modulations of oscillatory activity: Top - grand average ($n = 5$) topographical representation of band-power differences (*intention* minus *baseline*), strongest effects: anterior and posterior low alpha-power decrease (a); hemisphere lateralization for beta- and gamma-power (b). Bottom - grand average across all participants *initiator* minus *responder*, strongest effects: pronounced left hemisphere beta- and gamma-power increase (c); high alpha-power decrease over motor cortex (d). White markers indicate highest deviations between conditions.

Classification results: As reported in Table 3 a maximum individual classification accuracy of $Acc = 84.8\%$ (s04) when predicting the *intention to initiate gaze-contact* was obtained with the proposed approach. The CV-accuracy averaged across all participants was $Acc = 80.4 \pm 7.1\%$ (AVG (average)$\pm$SD (standard deviation)). For the L1O validation a maximum accuracy of $Acc = 67.5\%$ (s06) and an average accuracy of $Acc = 64.2 \pm 3.2\%$ (AVG$\pm$SD) across all participants was obtained.

Table 4 shows the results for predicting whether participants were initiating or following gaze-contact with the robot. We obtained a maximum CV-accuracy of $Acc = 84.3\%$ (s05) and on average $Acc = 77.0 \pm 5.2\%$ (AVG$\pm$SD) across all participants. L1O validation yielded a maximum accuracy of $Acc = 71.7$ (s04) and on average $Acc = 61.0 \pm 5.6\%$ (AVG$\pm$SD). We did not observe any tendency regarding optimal values for the parameters C and γ.

3.3.4. Discussion of electrophysiological and classification results

Despite the relatively small number of participants we observed plausible oscillatory modulations in the EEG patterns. The increase of anterior and posterior alpha power in the condition intention versus baseline could be related to increased alertness and levels of attention, an effect which has been consistently reported in the literature, such as in [132]. The differences

Table 3 Within-participant cross-validation (CV) and leave-one-participant out (L1O) classification accuracies for conditions *intention versus baseline*.

participant ID	CV			L1O		
	Acc [%]	C	γ	Acc [%]	C	γ
s02	83.8	100	0.001	65.0	0.1	0.005
s03	66.2	2	0.1	66.7	3	0.01
s04	84.8	3	0.01	58.3	0.001	2
s05	83.0	20	0.001	64.2	10	0.001
s06	84.0	10	0.005	67.5	0.001	0.005
AVG±SD	80.4±7.1			64.2±3.2		

Table 4 Within-participant cross-validation (CV) and leave-one-participant out (L1O) classification accuracies for conditions *initiator versus responder*.

participant ID	CV			L1O		
	Acc [%]	C	γ	Acc [%]	C	γ
s02	76.3	100	0.001	60.8	1	0.5
s03	69.2	3	0.1	55.8	0.001	0.005
s04	80.7	10	0.005	71.7	20	0.005
s05	84.3	10	0.01	58.3	20	0.001
s06	74.5	2	0.05	58.3	2	0.05
AVG±SD	77.0±5.2			61.0±.6		

in left hemispheric gamma band and central beta band power could be related to variations in approach and withdrawal motivation, which has been associated with differential activations of the medial prefrontal gyrus (MPF). For example, in 2005, Talati and colleagues [133] found greater left MPF activations for approach and greater right MPF for withdrawal tendencies. This would then suggest a pronounced approach motivation during the *intention* periods compared to the *baseline*. The power decreases over left and right motor cortex in high alpha band in the comparison *initiator versus responder* suggest increased engagement of motor areas and could be related to pronounced motor preparation / planning during the *initiator* as compared to the *responder* periods.

The offline classification accuracies obtained for participant-dependent models (within-participant

validation) are high and consistently close to or above 75% in all but one participant (s03, with comparably low classification accuracies in both comparisons). These results are promising in that they are close to typical within-participant classification accuracies in the domain of EEG-based BCIs, such as in SMR-based BCIs [134, 135]. Offline classification accuracies obtained for participant-independent models (across-participant validation) are much lower which are likely due to across participant variations in the observed patterns of oscillatory modulations. This was expected, since participant-to-participant transfer is a typical challenge in the domain of EEG decoding [136] with so far only moderate results.

Despite plausible patterns of oscillatory modulations and promising within-participant single-trial classification results, the findings of this study are limited in that it remains unclear if and to what extent the observed patterns can in fact be related to neuronal processes involved in engagement and social role evaluation. To address this, a follow-up study with a larger number of participants was conducted, and corresponding data analyzed in a complementary way.

3.4. Full study, data analysis, and results

The same experimental setup (see Section 3.2) was employed in a follow-up study with a larger number of participants including an extended EEG setup and additional measurement modalities. The main purpose of the full study was to complement preliminary findings obtained in the pilot study. Besides confirming previous findings on oscillatory modulations, we also performed an investigation of ERP which was not feasible based on the pilot data due to the low number of participants. For the sake of brevity, only results obtained for the latter analysis are reported in the remainder of this chapter. We hypothesized to observe ERPs in response to gaze-contact events with differential modulations when contrasting the conditions RI and YI which would allow a complementary interpretation of the data in light with previous findings.

3.4.1. Participants and data collection

Thirty-one healthy volunteers participated in this study. Due to technical problems, 6 of these datasets were not complete and hence excluded from further analysis. Three of the remaining datasets were heavily contaminated with artifacts and therefore also excluded. Furthermore, 7 of the remaining datasets were excluded since participants were doubtful about the experiment belief manipulation. Thus, data from 15 participants were analyzed (age: 27.5 ± 7.0; 5 females, 10 males; all had normal or corrected vision). Overall, prior experience and familiarity with humanoid robots scored lower than in the pilot study with 2.3 ± 1.6 on a scale of 1 "non-familiar" to 7 "familiar". Participants provided full written informed consent in regard to their participation, were offered full debriefing after one month, and were paid at the rate of 8 EUR / hour. The study was approved by the institutional ethics review board of the Technical University of Munich (reference number 236/15s).

EEG data was acquired with a Brain Products actiChamp amplifier equipped with 64 active EEG electrodes[3] arranged according to the international 10-20 system [128]. All leads were referenced to Cz and the sampling rate was set to 500 Hz. The impedance levels of all leads were kept below 10 kΩ. Other physiological measurements were also taken, including heart rate, respiration and galvanic skin response. Participants were also fitted with the SMI Eye Tracking Glasses by SensoMotoric Instruments, in order to record their gaze during the experiment and to control for those trials in which participants were completely inattentive to the robot. The glasses were wired to an adapted Samsung S5 mobile phone, which allowed the data to be recorded onto a memory card using the iViewETG controller software installed on the mobile device. Analysis and results based on gaze data and other physiological measurements are not reported in this chapter.

The participants were first instructed (verbally and by means of written instruction) about the experimental setup and handed a questionnaire about personal details. The participants were further instructed to sit still during the experiment, keep their gaze on the robot and avoid any unnecessary movements, such as head shifts. To test whether participants believed the experiment instructions they were given another questionnaire after the experiment that included questions e.g. about how well they thought the robot responded to their intention to engage in gaze-contact.

3.4.2. Analysis of event-related potentials (ERP)

EEG data processing: All EEG data preprocessing was carried out in MATLAB$^{\circledR}$, in part using functions provided by the EEGLAB toolbox [137]: (1) The data was first bandpass filtered using an EEGLAB in-built function (pop_eegfiltnew) which employs a zero phase Hamming windowed sinc FIR (finite impulse response) bandpass filter with cutoff frequencies of 1 Hz and 40 Hz in order to remove slow DC drifts, high frequency and power-line noise[4]. (2) Next, we identified and interpolated contaminated EEG channels using kurtosis with a threshold of 5% and reconstructed the original reference channel Cz by interpolation from neighboring channels. (4) Then, EEG signals were re-referenced to CAR to further reduce signal contamination due to external noise sources. (4) For reducing processing time of subsequent steps, the data was downsampled to 128 Hz. The last steps had the purpose to remove typical EEG artifacts from the data (eye-blinks, horizontal eye-movement, muscular artifacts, and data inconsistencies) using ICA: (5) The data was first decomposed using the ICA algorithm *runica* provided by the EEGLAB toolbox [137]. (6) ICA components which could be associated with above mentioned artifacts were subsequently removed using ADJUST, an EEGLAB plugin performing automatic identification of artifact components [138]. This automatic cleaning procedure resulted in the removal of an average of 18 out of 64 components per participant. (7) Afterwards the data were transformed back to channel-level using inverse ICA.

---

[3] EEG channel labels: Fp1, Fz, F3, F7, FT9, FC5, FC1, C3, T7, TP9, CP5, CP1, Pz, P3, P7, O1, Oz, O2, P4, P8, TP10, CP6, CP2, C4, T8, FT10, FC6, FC2, F4, F8, Fp2, AF7, AF3, AFz, F1, F5, FT7, FC3, FCz, C1, C5, TP7, CP3, P1, P5, PO7, PO3, POz, PO4, PO8, P6, P2, CPz, CP4, TP8, C6, C2, FC4, FT8, F6, F2, AF4, AF8, Cz.

[4] Due to the use of active electrodes, which reduce environmental artifacts, such as power line interference [129], no additional notch-filter was employed for further temporal filtering.

Data segmentation: After data pre-processing, data epochs ranging the period [-200...1200] ms relative to the gaze contact events were extracted for each of the two conditions (YI and RI). This period captured all relevant effects observed in the data. Each channel and trial were furthermore baseline corrected by subtracting the average of period [-200...0] ms relative to the gaze contact event from the time-course of the entire trial. Single-trial baseline correction is a common procedure in the analysis of ERP to remove unwanted fluctuations of DC levels between trials [139].

Computation of ERPs: Trials were afterwards averaged per condition (YI, RI) for each participant resulting in one average ERP activation pattern per participant and condition. Averaging across trials ensures that resulting activation patterns reflect brain activity that is both time- and phase-locked to the onset of the event/stimulus of interest. Individual average ERPs were then tested for significant differences across participants over both spatial and temporal dimensions (channels and sample time-points).

3.4.3. Results of ERP modulations

Results are depicted in Figure 15 and show the grand average ERP time-courses pooled over central channels FCz/C1/CPz/C2/Cz time-locked to the gaze-contact event for each condition (blue: RI "Robot Initiates", red: YI "You Initiate") and the difference grand average (black: YI minus RI). The difference grand average is furthermore depicted in form of spatio-temporal activation patterns across all channels and relevant time points above the plot. The ERP shows a slight negative deflection around 250 ms which is more pronounced in the YI condition compared to the RI condition. According to the latency and spatial distribution of the effect, this finding can be related to the N200 component which is expected at latencies within [180...325] ms over frontocentral channels [140]. Amplitude differences between conditions pooled across central channels within period [180...325] ms were significant at the $p < 0.05$ level across participants (paired Wilcoxon signed rank test, $n = 15$). Furthermore, a stronger positive reflection was observed around [300...600] ms which was also more pronounced in the YI condition compared to the RI condition. Specifically, the YI condition shows stronger activation after the main peak around 400 ms which is sustained until around 800 ms. Amplitude differences between conditions pooled across central channels within period [300...600] ms were significant at the $p < 0.05$ level across participants (paired Wilcoxon signed rank test, $n = 15$).

3.4.4. Discussion of ERP modulations

The difference of observed ERPs between the two conditions are spatio-temporally similar to difference waves of error-related potentials. The early negative deflection around 250 ms together with the later positive deflection around 500 ms, both spatially distributed over fronto-central channels resembled indeed the characteristic N2-P3 complex typical in ErrPs. In particular, the N200 component, implicated in *mismatch detection* [140], turned out significantly more pronounced in the YI condition compared to the RI condition. This suggests that in the YI condition participants seem to have experienced gaze-contact events as mismatch-
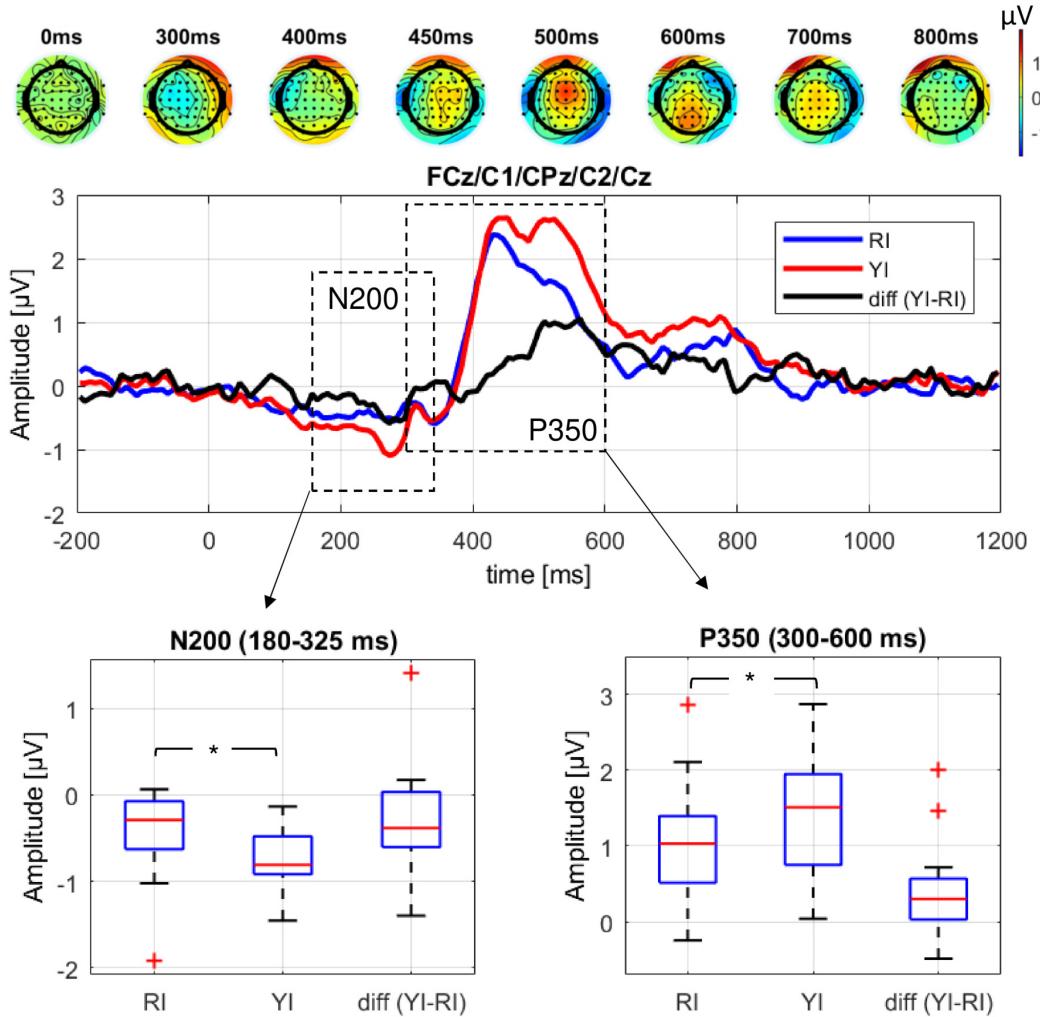
Figure 15 Results of modulations of event-related potentials (ERP): Grand average ERP time-courses pooled over central channels FCz/C1/CPz/C2 time-locked to the gaze-contact event for each condition (blue: RI "Robot Initiates", red: YI "You Initiate") and the difference grand average (black: YI minus RI). The difference grand average is furthermore depicted as topographic in form of spatio-temporal activation patterns across all channels and specific time points above the plot (Figure style adopted from [141]). Results show significant deviations between conditions during the period in which the N200 component is expected ([180...325] ms) and marginally significant deviations during the period in which the P350 component is expected ([300...600] ms). The bottom panel boxplots show the activity distribution across participants for each component pooled across channels FCz/C1/CPz/C2 and the respective period. Asterisks (*) denote significant differences with $p < 0.05$ between conditions (paired Wilcoxon signed rank test, $n = 15$).

ing/conflicting stimuli compared to the RI condition. The later positive deflection around 500 ms can be related to the P350 component, expected at latencies of 300-600ms over central channels. In fact, the findings of this study are in line with those observed in a study by Carrick in 2007 who suggested the P350 component to be a neural correlate of contextual evaluation of the social meaning of a gaze-contact event [142]. The qualitative resemblance between the observed ERP modulations with the spatio-temporal characteristics of the ErrP bring about the question as to what extent performance monitoring could have played a role in the experimental paradigm: During the experiment, participants were asked to engage with the robot during the YI condition by attempting to willfully control the robot to engage with them. In the RI condition, participants were passively waiting for the robot to engage with them without performing any particular task. The questionnaire results ensured that all participants whose

data was analyzed in fact believed they could control the robot. Nonetheless, since the mental task "influence the robot to engage with you" in the YI condition is quite abstract, it appears reasonable to assume that not all participants were in fact performing this task reliably within and across trials. For instance, participants could have tried to influence the robot, realized it was not responding, tried again or gave up. This could have created the following confounding scenarios:

Scenario A: Participant tries to influence the robot, and as a result expects the robot to respond within a certain period of time:

- **Outcome A1**: robot does not respond as expected, leading to an ERP in response to a non-occurring gaze contact event > expectation mismatch (error)

- **Outcome A2**: robot does respond as expected, leading to an ERP in response to an occurring expected gaze-contact event > expectation match (non-error)

Scenario B: Participant does *not* try to influence the robot, and as a result expects the robot *not* to respond:

- **Outcome B1**: robot does not respond as expected, leading to an ERP in response to an expected non-occurring gaze-contact event > expectation match (non-error)

- **Outcome B2**: robot does unexpectedly respond, leading to an ERP in response to an unexpected occurring gaze-contact event > expectation mismatch (error)

Which of the above situations occurred in a particular trial cannot be tracked back post-hoc due to the experimental design. It depended on how participants performed the task which was neither explicitly controlled nor measured in the experimental study. Following this alternative interpretation, we assume that feedback error trials were more frequent in the YI condition, since participants had expectations that could be violated, whereas in the RI conditions participants had no particular expectations that could be violated. This would explain the spatio-temporal resemblance between the observed difference wave (YI minus RI) and the typical characteristics of ErrPs, in particular, the more negative N200 in the YI condition compared to the RI condition. Further, we assume that condition YI would contain confounding ERPs from class non-error and class error, depending on which of above situations occurred in the respective trial. This confound - condition YI containing more mismatching gaze-contact events than RI, but mixed with matching gaze-contact events - could be the reason for the relative weak effect strength observed in the data.

3.5. Summary

These initial studies aimed at exploring the feasibilities of enhancing robot's abilities to sense when (timing) and how (social role) to engage in interaction with a human by means of EEG. We developed an experimental paradigm to evoke and capture electrophysiological data (EEG signals) associated with (1) the intention to initiate gaze-contact and (2) the distinction between initiator and responder role in established gaze-contact with the humanoid iCub. In a pilot study we focused on the analysis and modeling of oscillatory EEG activity. Results showed that intention to engage with the robot versus passively waiting for the robot to initiate contact can be discriminated from the brain oscillatory activity.

Within participant offline CV accuracies reached an average of $Acc = 80.4\%$. Also, discriminative oscillatory activity was found during gaze-contact periods when participants were initiating versus following the engagement process. Within participant offline CV accuracies reached an average of $Acc = 77.0\%$. Accuracies of participant-independent models were furthermore above chance level, with across-participant average accuracies of $Acc = 64.2\%$ and $Acc = 61.0\%$ for each of the analyzed conditions. In a follow-up study using the same experimental paradigm we focused on the analysis of ERPs time locked to the moment of gaze-contact between human and robot. Results showed different activation pattern for the two conditions: when the human initiated (YI) and when the robot initiated (RI) the interaction. Differences in activation occurred at latencies typical for the N200 and the P350 components, and effects were significantly more pronounced in the YI rather than in the RI condition. These findings indicated varying processing of expectation mismatch as well as differential evaluation of social context depending on the participant's belief about their social role during the a gaze-contact event. The observed spatio-temporal activation patterns resembled those typical for ErrPs which suggested the involvement of performance monitoring processes. Indeed, an alternative interpretation of the data, in perspective with the experimental paradigm used, relates the main effect to error-/performance monitoring. This re-consideration of the experimental paradigm gave rise to focusing on ErrPs in follow-up studies, and triggered the interest in using ErrPs as a quantitative measure about participants' belief and expectations towards robot behavior in HRI.

4. A feasibility study for validating robot actions using EEG-based error-related potentials

Validating HRI can be a challenging task, especially in cases in which the robot designer is interested in the assessment of individual robot actions within an ongoing interaction, without interrupting the interaction for questions/surveys. In this chapter, we propose a method for real-time quantitative assessment of robot actions and evaluate the usefulness as a complement to existing methods for validating HRI. The method is based on the decoding of ErrPs from the EEG of a human partner during interaction with a robot. To demonstrate the usability of the approach, we conducted a study investigating whether EEG-based ErrPs are elicited in response to a real humanoid robot displaying incorrect actions in a simplistic HRI task. Furthermore, we conducted a procedurally identical control experiment with computer screen-based symbolic cursor action. The results of our study confirmed decodability of ErrPs in response to incorrect robot actions with an average accuracy of $Acc = 69.0 \pm 7.9\%$ across 11 participants. Cross-comparison of ErrPs between experimental tasks and control conditions revealed high temporal and topographical similarity, but more distinct signals and, as a result, better decodability with a mean accuracy of $Acc = 90.6 \pm 3.9\%$, in the control experiment. This demonstrated that ErrPs can be sensitive to the stimulus eliciting them despite procedurally identical protocols. This implies that re-using ErrP-decoders across experimental tasks without re-calibration is accompanied by significant performance losses and therefore not recommended. Overall, the outcomes of our study confirm feasibility of ErrP decoding for human-robot validation, but also highlight challenges to overcome in order to enhance usability of the proposed method.

4.1. Introduction

More than a decade of research on HRI [143] has been dedicated to the question of how to make interaction with robots more intuitive and "natural" for the human user [144]. In this regard, the assessment and validation of robot behavior during interaction with humans is crucial for successfully directing technical improvements towards more widespread and effective integration of robots in society. Validating robot behavior during HRI can be a challenging task, particularly in the domain of humanoid and social robotics. Typical scenarios include collaboration tasks in shared environments [145, 146] and game-based or dialogue social interaction tasks [147, 67]. The human's subjective experience of the robot's behavior is usually assessed

---

with survey-based methods, such as questionnaires [148, 12, 149] or interviews [150]. The quality of interaction is often additionally assessed with objective performance measures, such as "time to complete the task", as in [145, 146]. To avoid interruption of the interaction flow, these measures are often taken at the end of a task and as such constitute an average assessment of the performed sequence of actions or events. However, the assessment of individual robot actions may be beneficial or, in specific cases, even required to effectively pinpoint factors that have influenced the interaction. Here, we propose a method that addresses these limitations by providing quantifiable information about the human partner's immediate assessment of individual robot actions. The proposed method is based on single-trial decoding of ErrPs, time-locked to the occurrence and the human partner's observation of individual robot actions. In line with the contemporary understanding of performance monitoring and the effect of ErrPs (see Chapter 2, Sections 2.3 and 2.4), we expect the occurrence of deviating brain responses in case the human observes incorrect robot actions compared to the observation of correct robot actions. Classifying these responses from the ongoing EEG signals allows for implicit and real-time binary labeling of single robot actions immediate to their occurrence (see Figure 16). In a collaborative assembly task, such an incorrect robot action may be the robot providing the human a wrong object for the next step in the assembly. In a game-based or dialogue interaction task, an incorrect robot action may be the robot performing a social cue, such as gaze contact with the human partner, in an unexpected, contextually inappropriate moment.

4.1.1. Related work on single-trial decoding of ErrPs for assessment of HMI

Decoding ErrPs from multichannel EEG signals is a binary classification task. The two classes are commonly denoted as *non-error* (ERP in response to correct or matching events, or positive feedback) and *error* (ERP in response to erroneous or mismatching events, or negative feedback). *ErrP decoding* refers to the prediction of the type of event (non-error, error) from the ERP resulting from the human's observation of that event. Over the last two decades, a multitude of works repeatedly showed that ErrPs can be reliably decoded from EEG signals with single trial classification accuracies around 70%–80% in a variety of experimental contexts [103]. Classifiers have furthermore been shown to be robust across recording sessions with up to 600 days between sessions [151]. Due to large amounts of channels (usually 32-64 channels) and sample time-points (approximately 1 s) that can be considered, the feature space is high-dimensional. Most research works employed temporal features extracted from sample time-points of individual channels, such as signal amplitude from a single channel at specific latencies time-locked to the event [103]. A few works explored also other features, such as frequency domain-based features [152, 153, 154, 42] or connectivity features [155]. In terms of classifiers used for binary discrimination, most works employ simple statistical methods, such as Gaussian Mixture Models, LDA (linear discriminant analysis) classifiers, or SVMs. Feature dimensionality reduction or regularization techniques are commonly used to reduce the risk of overfitting and to obtain robust and widely applicable classifiers. Integrated decoding approaches, such as those based on convolutional neural networks [156] are currently still underrepresented. ErrPs are typically decoded in an event-based fashion,

meaning, the signal processing and classification method is precisely synchronized with the moment of event occurrence. As such information may not be available in more natural HMI, several works explored the feasibility to decode ErrPs in a continuous, asynchronous fashion [157, 158, 159, 104]. Overall, they confirmed feasibility for asynchronous decoding, however with accuracies lower than for event-based classification (around 65%).

While ErrPs have been shown to be reliable across recording sessions [151], they can vary in spatiotemporal shape across participants and experimental tasks. This fact forces the experimenter to (re-)calibrate ErrP-decoders for individual participants and experimental task, a procedure which can be highly time- and resource consuming. Therefore, several works proposed methods for reducing calibration time [160, 161, 162, 163, 164, 165] or transferring classifiers between tasks [141, 166, 167, 168]. Despite these efforts, ErrP decoding still remains a tedious procedure with most research works proposing their own proprietary methods tailored to specific task. Most previous studies on ErrPs used computer screen-based interaction with perceptually simple symbolic stimuli with single-trial decoding performance of 70%–80% [169, 102, 170, 141, 104]. Others have studied ErrPs in the context of HRI [171, 172, 105, 173, 43, 174]. While some of the latter works reported high and thus practically useful single-trial decoding performances of >75%, e.g. [171, 105, 43], others reported lower performance: For instance, Welke and colleagues presented in 2017 a study in which they investigated the feasibility of decoding ErrPs in response to participants watching videos of robots performing erroneous actions [174] and reported decoding performances which where just slightly above chance-level. Similarly, Salazar-Gomez and colleagues reported ErrP decoding performances of around 65% in response to the human observation of erroneous actions performed by a real robot.

In summary, this indicates that the transfer of the ErrP decoding to real robotic system and across different types of tasks is not straightforward and points towards the need for a systematic study of observability and decodability of ErrPs in the context of varying stimuli.

4.1.2. Aim of the work

The present study addresses the question of whether and to what extend the observability and decodability of ErrPs are transferrable towards the assessment of real machines and robotic systems. Thus, we examined the observability and decodability of EEG-based ErrPs in response to a real humanoid robot displaying incorrect actions in a simplistic HRI task, where robot actions either conform (congruent, correct action) or disconform (incongruent, wrong action) to a selection the human partner made. Furthermore, we conducted a procedurally identical control experiment with computer screen-based symbolic cursor action. With this comparative experimental design, we address the following fundamental aspects with practical implications on the usability of deploying EEG-based ErrP decoding for validating robotic systems and HRI:
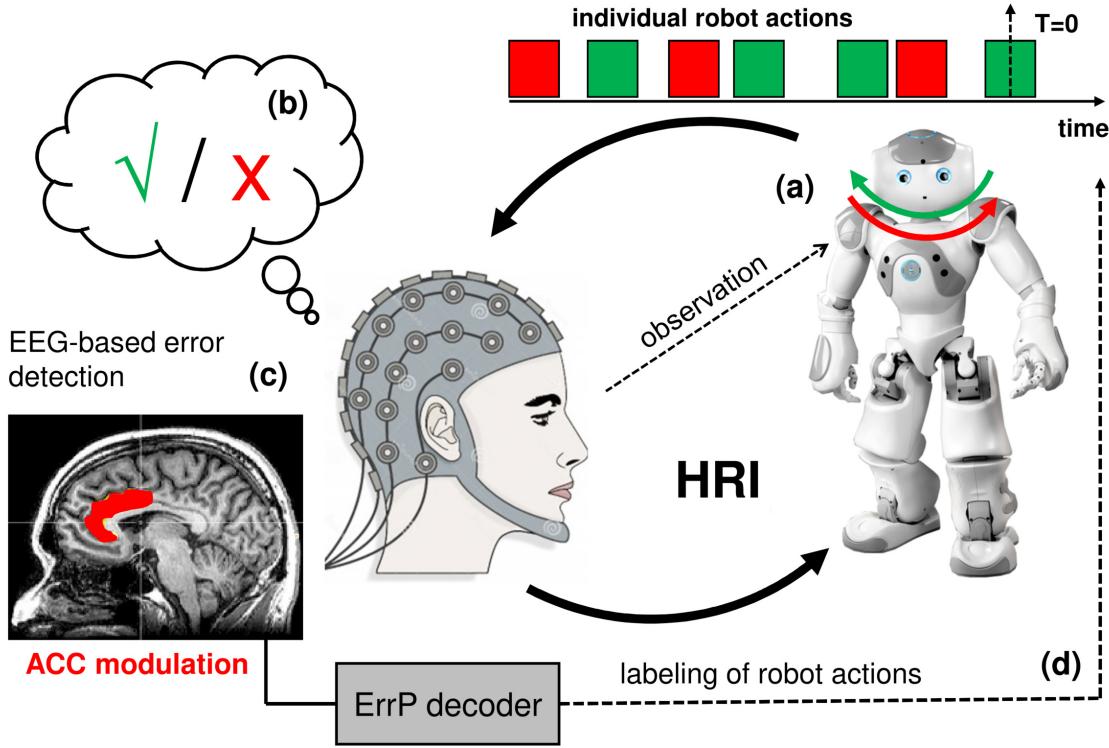
Figure 16 Conceptual illustration of using ErrPs for assessment of robotic systems and HRI: (a) During interaction with a robot, a human observes the behaving robot. Meanwhile, the brain activity is recorded via EEG and responses to single robot actions are captured and analyzed. (b) Brain responses (e.g. ErrPs) associated to the observation of incorrect robot action are expected to deviate from those associated to the observation of correct actions. (c) Classifying these EEG responses allows implicit real-time labeling of single robot actions immediate to their occurrence, usable for (d) post-hoc validation of the robot behavior.

- Verification of the observability and decodability of EEG-based ErrPs in response to the human observation of incorrect actions performed by a real humanoid robot.

- Demonstration of the extent as to which EEG signals and consequently the observability and decodability of ErrPs, vary with the type of entity performing the action (cursor, robot), despite procedurally identical experimental protocols and the use of the same decoding method.

- Investigation of the feasibility of deploying ErrP-decoders across experimental tasks, such as calibrating the decoder based on one task and applying it for decoding ErrPs based on the other task.

By addressing the first aspect, we aim to confirm feasibility of detecting ErrPs in response to human observation of incorrect robot actions. The second aspect concerns the understanding of whether and to what extent the type of stimulus can affect the observability of ErrPs and if this results in altered efficiency for decoding ErrPs in response to robot actions. By addressing the third aspect, we test if it is possible to re-use ErrP-decoders across experimental tasks without the need for re-calibration.

The remainder of this chapter is structured as follows: In Section 4.2, the experimental paradigm (Section 4.2.1), design (Section 4.2.2) and implementation (Section 4.2.3) of the

study as well as data analysis (Section 4.2.4) are presented. The corresponding results are reported in Section 4.3 and interpreted and discussed in Section 4.4. Section 4.5 summarizes the chapter.

4.2. Methods: experimental study 2 and data analysis

4.2.1. Experimental paradigm

Traditionally, the study of ErrPs was largely performed using CRTs and variants [92, 93]. To allow for systematic comparisons with the current body of literature on ErrPs and draw from a well-established paradigm, the experimental tasks were designed based on the principle of a CRT task with identical protocols in both tasks (Figure 17): One out of three possible target stimuli appeared on a computer screen and participants were requested to respond, with a corresponding key press, as quickly and precisely as possible. Feedback to the participant key press was one of two possible outcomes: congruent or incongruent response to the target stimulus. The two experimental tasks differed only in the type of feedback presented to the participant: In experimental task 1 (*cursor scenario*), a cursor, centrally placed on the computer screen, would either move towards (i.e. congruent) or away (i.e. incongruent) from the target stimulus (Figure 17(a)); in experimental task 2 (*robot scenario*), the head of a humanoid robot would either turn towards (i.e. congruent) or away (i.e. incongruent) from the target stimulus (Figure 17(b)). Robot head turns as the corresponding counterpart to the cursor actions were chosen for having the robot performing actions which can be understood in the sense of gaze cues [175]. As such, we realized an experimental setting with two tasks of identical procedure, but with different connotation: while the cursor scenario follows traditional protocols used in the study of ErrPs, the robot scenario deploys an embodied humanoid robotic presence performing actions that can be useful in real-world HRI [176]. In both experimental tasks, we expected to observe in the participant's EEG ERPs with different characteristics when evoked by congruent (semantically correct) versus incongruent (semantically incorrect) feedback. By varying the type/form of the feedback only, our design allowed us to test whether and how the observability and decodability of ErrPs evoked by simplistic symbolic feedback (experimental task 1) can be transferred to a scenario involving the execution of actions by a humanoid robot (experimental task 2) in a procedurally identical protocol.

4.2.2. Experimental design

4.2.2.1 Participants

Thirteen healthy participants took part in the experiment. The data of two participants were excluded from further analysis: participant s01 due to technical problems during the experiment, and participant s12 due to having been on medication during the experiment. The remaining 11 participants were 6 males and 5 females with average age: 29.4±7.4 years. Prior experience and familiarity with humanoid robots scored 2.8±1.8 on a scale of 1 "unfamiliar" to 7 "familiar". All participants were equally instructed about the experiment protocol and provided informed consent regarding participation in the experiment. Each participant was paid at a
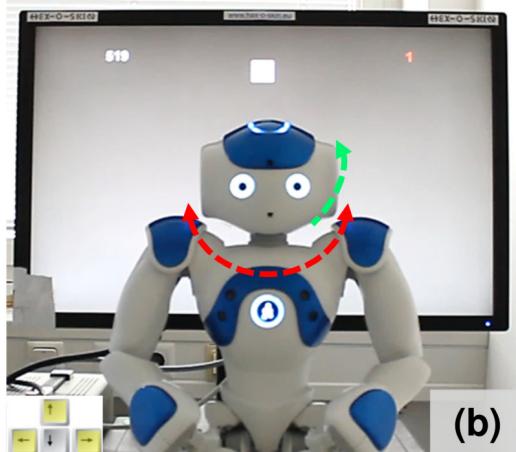
Figure 17 Illustration of the two experimental tasks; cursor task (left) and robot task (right). Upper panels: Participants were requested to respond to one out of three possible target stimuli (left, right, up) appearing on the screen with corresponding arrow key presses. In response, either a cursor (a) or the robot head (b) would move towards or away from the given target. The yellow dashed arrows show the directions of possible cursor movements and robot head turns. Lower panels: experimental setup showing a participant performing the cursor scenario (c) and the robot scenario (d).

rate of 8 EUR / hour. The study was approved by the institutional ethics review board of the Technical University of Munich (reference number 236/15s).

4.2.2.2 Experimental protocol:

The experiment was divided into two recording sessions (one for each scenario) taking place one after the other. About half of the participants (6 out of 11, see Supplementary Table 11, Appendix A) started with the cursor scenario (order: C/R) and the others with the robot scenario (order: R/C). In total, 500 trials were collected per scenario; each scenario was further divided into 10 blocks of 50 trials each; the duration of one block was approximately 2.5 min. Thus, the total duration of the experiment was approximately 60 min. After each block, the participant would take a rest and decide when to continue with the next block in a self-paced fashion. The participants were first instructed (verbally and by written instruction) about the experimental setup including the recording modalities and handed a questionnaire about personal details. Participants were informed about the approximate duration of the experiment, but not about the specific number of trials per block. Participants were instructed to react as quickly and precisely as possible to the appearing target stimuli.

4.2.2.3 Trial structure:

Each trial (see Figure 18) started with a pause of random duration between 500-2000 ms in order to avoid habituation to timing of appearance of the target stimuli. After the initial pause, one out of three possible target stimuli appeared on the screen, followed by the participant's self-paced key press. Feedback was presented in form of a cursor movement (cursor scenario) or a robot head turn (robot scenario) towards or away from the target stimulus. The feedback ended after 130 ms: cursor reached target stimulus / robot head movement reached end location. A second feedback was presented 200 ms afterwards in form of a colored frame around the target stimulus (green frame = correct, red frame = incorrect). The framed target stimulus disappeared 300 ms later; this initiated the robot head turning back to the initial location and the cursor re-appearing in the center of the screen 600 ms later.
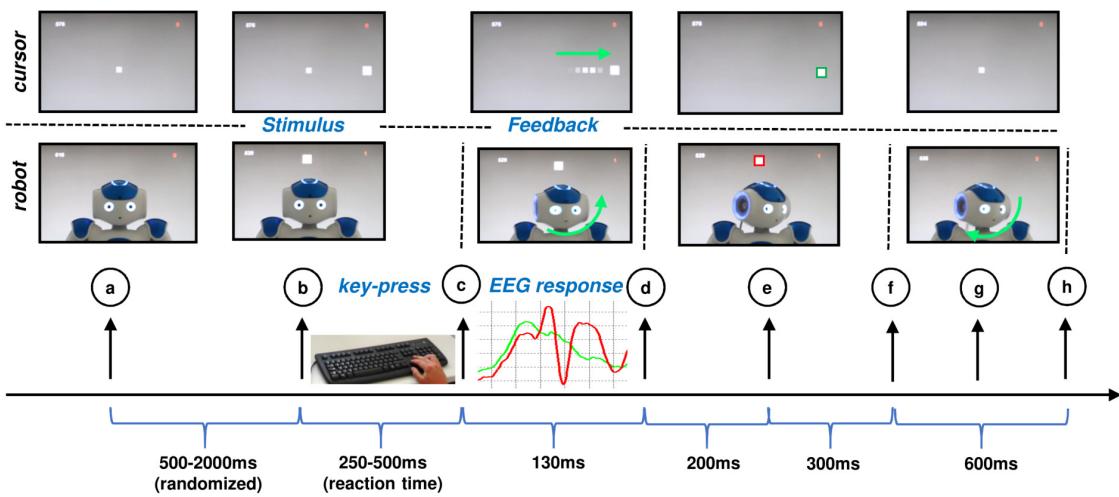
Figure 18 Trial sequence with exemplary illustration of the cursor scenario (top panel, example for a congruent/correct trial) and the robot scenario (middle panel, example for an incongruent/incorrect trial). (a) Trial start and break of random duration between 500-2000 ms, (b) appearance of target stimulus, (c) participant response in form of arrow key press, start cursor/robot head movement (d) end cursor/robot head movement (e) target border feedback presentation (correct: green; incorrect: red), (f) disappearance target, disappearance cursor / start robot head turning back (g) re-appearance cursor/ongoing robot head turning back (h) re-appearance cursor, end of robot head turning back, updating average reaction time and error count.

4.2.3. Stimuli and apparatus

4.2.3.1 Experimental setting

The experiment took place in a quiet room which was partitioned into two sections by means of a sight-proof wall. The experiment took place in the right side of the room and the participant sat approximately 150 cm away and facing the computer screen / the humanoid robot. Participant responses were registered with a computer keyboard located near the participant (see Figure 17(c and d)). All but one participant (s07) performed the key presses with their right hand. The experimenter was monitoring the session from the left side of the room.

4.2.3.2 Stimuli

Stimuli were presented on a 24-inch flat screen LCD computer monitor with 60 Hz refresh rate (see Figure 17). Participant responses were registered with the arrow keys of an ordinary computer keyboard. The experiment was programmed with Python using the Psychopy library [177] and executed on an Intel$^{®}$ Core i5 CPU 750@2.67 GHz. The target stimuli were realized as white squares of size 3x3 cm appearing in three possible locations on the computer screen (left, right, or up). Per trial, one out of the three possible target stimuli appeared. Participants were requested to respond with a corresponding arrow key (left target = left arrow key, right target = right arrow key, upper target = upper arrow key). Upon participant key press, feedback was initiated in form of a cursor (cursor scenario) or a robot head movement (robot scenario). False feedback events were introduced in a uniform random fashion after a correct response by the participant (i.e. key press congruent with target location), and with a pre-defined probability p_{err}. These events are termed "machine-errors" for the rest of this chapter. Machine-errors were manifested as cursor movement or robot head turns towards the wrong direction / incongruent to the target stimulus location. The wrong direction was selected in a uniform random fashion among the two remaining non-target directions. To avoid habituation to the machine-error probability, half of the blocks were executed with a probability of p_{err}=20% and the other half with p_{err}=50%. The order was pseudo-randomized such that no more than two subsequent blocks would belong to the same error probability category. The first block of each scenario was however always executed with p_{err}=20% to avoid any confusion in the beginning of each scenario. The cursor feedback was realized as a white square of size 2x2 cm, initially located in the center of the screen (see Figure 17(a and c)). Upon participant key press in response to the appearance of the target stimulus, the cursor would start moving towards or away from the target with uniform speed until reaching the target position after 130 ms. In the robot scenario, the cursor was substituted with a NAO humanoid robot located in a crouched posture in front of the computer screen such that the head position matched the center of the screen (see Figure 17(b and d)). NAO is a 58 cm tall humanoid robot with 21-25 degrees of freedom [178] which was controlled by the experiment program via LAN (local area network) using the Python-based NAOqi library. In the initial head position ($\Psi = 0°$ (pitch), $\Theta = 0°$ (yaw)), the robot was gazing directly towards the participant. Upon participant key press in response to the target stimulus, the robot head would turn towards or away from the target (left: $\Psi = 0°$, $\Theta = -40°$; right: $\Psi = 0°$, $\Theta = +40°$, up: $\Psi = -20°$, $\Theta = 0°$) and reach the end position after 130 ms, keep the position for 500 ms and move back to the initial head position. The average reaction time per block in ms was given in the left upper corner of the screen, whilst the number of errors per block, with no distinction of errors committed by the participant or the machine, were given in the right upper corner. This additional feedback was intended to keep participants' engagement in the task by self-monitoring their own performance throughout the experiment.

4.2.3.3 EEG recording and data pre-processing

EEG data were acquired with a Brain Products actiChamp amplifier equipped with 32 active EEG and EOG (electrooculogram) electrodes[1] arranged according to the extended international 10-20 system [128]. All leads were referenced to the average of TP9 and TP10 (average mastoids referencing) and the sampling rate was set to 1000 Hz. The impedance levels of all leads were kept below 10 $k\Omega$. Three channels were used for capturing electrooculogram (EOG1-3) signals in three locations of the participant's face (forehead, left and right outer canthi) according to a method suggested by Schlögl and colleagues [179]. The EEG amplifier was battery-driven and located on a tray near the participant. The data was transferred via USB to a separate recording PC (Intel® Core i5 CPU 750@2.67 GHz). The amplifier was connected to the PC executing the experiment protocol via LPT over which event triggers were sent to be stored synchronously with the EEG signals. All EEG data preprocessing was carried out in MATLAB®, in part using functions provided by the EEGLAB toolbox [137]. The subsequent processing steps were applied to each dataset (11 participants x 2 sessions) separately in the following order: First, the EEG and EOG signals were filtered with a zero phase Hamming windowed sinc[2] FIR band-pass filter with cutoff frequencies of 1 Hz and 20 Hz, to remove high-frequency and power-line noise. EEG channels showing signs of contamination with artifacts were identified using kurtosis with a threshold of 5% and replaced by interpolations from neighboring channels. EOG activity in the EEG signals (eye-blink and lateral eye movements) was corrected using the method suggested by Schlögl et al. [179]. Finally, EEG signals were re-referenced to CAR to further reduce signal contamination due to external noise sources.

4.2.3.4 Synchronization of stimuli onset with EEG recordings:

In order to ensure precise information about the moments of presentation of the feedback, the robot head was equipped with a LED (light emitting diode) and a photodiode to record the onset of head movements synchronously with the recording of the EEG signals. The computer screen was also equipped with a separate photodiode to record the timing of the cursor movement. None of the photodiode setups were directly visible and thus not distracting to the participants. Ground truth timing of onset of both cursor and robot head movement was obtained by analyzing the signals captured by the two photodiodes and introducing additional event markers into the EEG recordings.

4.2.4. Data analysis
4.2.4.1 Analysis of behavioral data

Our experimental design featured identical protocols but different feedback signals in the two scenarios. Therefore, we did not expect behavioral differences regarding *RT (reaction time)* and number of errors committed by the participants ($nErr$) across scenarios. Behavioral data

---

[1] Channel labels: FP1, FP2, F3, F4, F7, F8, FC1, FC2, FC5, FC6, C3, C4, T7, T8, CP5, CP6, P3, P4, P7, P8, TP9, TP10, O1, O2, Fz, Cz, Pz, EOG1, EOG2, EOG3.

[2] A sinc filter is a filter with impulse response being the sinc function; this results in an idealized filter with rectangular frequency response and linear phase response.

was analyzed to verify the absence of such behavioral differences. Several statistical tests were done for that: we tested whether the distributions of mean reaction times ($\overline{RT}$) and $nErr$ differed across the two scenarios, and tested whether the distributions of $\overline{RT}$ and $nErr$ differed across first and second performed scenario, irrespective of scenario order (C/R or R/C).

4.2.4.2 Analysis of event-related potentials and their stimulus-dependent variations

The data was segmented into epochs by extracting time intervals of [-500...1500] ms relative to the presentation of the feedback (onset of cursor movement / robot head turning, $t = 0ms$). These segments were further separated into three categories: (1) correct trials (non-error), (2) false feedback trials (machine-error), (3) human error trials. Per participant and recording session, we extracted on average approximately 325 non-error trials, 159 machine-error trials, and 16 human error trials. Since human-errors were not the focus of our investigation, the corresponding epochs were discarded from further analyses. Analysis of the shape and timing of the potentials in each scenario was carried out through (1) the computation of the time-locked average potentials for the machine-error and non-error potentials in channel Cz, (2) through the difference average (machine-error minus non-error averages) and (3) by computing the coefficient of determination r^2 [48]. Before averaging across trials, each channel and trial was baseline corrected by subtracting the average amplitude of the period [-200...0] ms from the entire signal epoch. Spatial ERP patterns were compared by computing the topographic interpolation of the potentials at the time of the main peaks of the difference average. Finally, we assessed the similarity of ERP time courses within the period [0...800] ms *per participant* by computing the 2D correlation coefficients between the difference average of the cursor and the robot scenario. Furthermore, we computed the 2D correlation coefficient between the difference *grand average* within the period [0...800] ms of the cursor and the robot scenario according to Eq. (4.1); with C and R being the difference average ERP of cursor C and robot R data (machine-error minus non-error); with $\overline{C}$ and $\overline{R}$ being the means of all elements in C and R; c being the spatial dimension (channel), and t being the temporal dimension (sample time point).

$$r = \frac{\sum\limits_{c}\sum\limits_{t}(C_{ct} - \overline{C})(R_{ct} - \overline{R})}{\sqrt{\sum\limits_{c}\sum\limits_{t}(C_{ct} - \overline{C})^2 \sum\limits_{c}\sum\limits_{t}(R_{ct} - \overline{R})^2}} \tag{4.1}$$

4.2.4.3 Single-trial classification and analysis of impact of stimulus-dependent variations

Computation of participant-specific ErrP-decoders, to be used to classify EEG signals into responses due to the observation of non-error or machine-error events, included the extraction of relevant features from the EEG signals and the training of a classification model. The

testing/validation step included feature extraction and application of that classification model on unseen data within and across session.

Feature extraction: In the context of single-trial classification of error-related potentials, different types of features have been used and reported in previous works. Temporal features extracted from the time series [102, 171] have been used in most cases, being reported as stable and reliable even across recording sessions [151]. Furthermore, we performed a feature cross-comparison in one of our earlier works [42] where we found temporal features being superior over spectral features in the context of decoding ErrPs. Therefore, temporal features were used in this work. For each trial and each channel, the signal amplitude was averaged within 9 overlapping 100 ms-long windows[3], relative to the occurrence of machine-error / non-error events for each channel and concatenated into the single feature vector of length 243 (27 channels x 9 windows).

Classification: The classifier used in the analysis was a regularized version of the linear discriminant analysis (rLDA) [180]. The rLDA classifier has been established as a robust method to discriminate mental states based on EEG signals in the field of brain-computer interfaces [181]. The LDA discriminant function is the hyperplane discriminating the feature space corresponding to two classes: $y(x) = sign(w^T x + b)$, with x being the feature vector, w being the normal vector to the hyperplane (or weight vector), b the corresponding bias, and $y(x) \in \{-1, 1\}$ the classifier decision. The weight vector and bias were computed by $w = (\hat{\mu_2} - \hat{\mu_1})(\tilde{\Sigma}_1 + \tilde{\Sigma}_2)^{-1}$ and $b = -w^T(\hat{\mu_1} + \hat{\mu_2})$, with $\hat{\mu_j}$ being the class-wise sample means, and $\tilde{\Sigma}_j$ being the class-wise regularized covariance matrices. Regularization aims at minimizing the covariance estimation error by penalizing very small and large eigenvalues. This leads to robust covariance estimates even for high dimensional feature spaces [181] as in our case. The regularized covariance matrices were computed by $\tilde{\Sigma}_j = (1 - \lambda)\Sigma_j + \lambda\nu I$, with $\lambda \in [0, 1] \subset \mathbb{R}$ being the shrinkage parameter, ν the trace (sum of diagonal elements) of Σ_j divided by the number of features, and I the identity matrix. The optimal shrinkage parameter λ was determined automatically based on the given training data using an analytic method proposed by Schäfer & Strimmer in 2005 [182].

Within-session validation: We validated the above described modeling approach within session (cursor, robot) using a 10-times-10-fold CV scheme. Per session and participant, the trials were randomly split in 10 folds, 9 folds were used for model calibration and the remaining fold was used for testing. This procedure was repeated until all folds were once used for testing. The entire procedure was furthermore repeated for 10 times. Each time and fold, the number of trials per class of the calibration data was balanced by random pick and replace (please note that the number of trials per class was initially unbalanced with ~65% non-error and ~35% machine-error trials). This analysis provides an estimate of how well participant-specific ErrP-decoders would perform in classifying unseen data when being calibrated with different data of that same session. Individual classification results per time and fold were

---

[3] Windows: [100...200] ms, [150...250] ms, [200...300] ms, [250...350] ms, [300...400] ms, [350...450] ms, [400...500] ms, [450...550] ms, [500...600] ms.

averaged and reported per session and participant as percentage of correctly classified trials (overall accuracy, $Acc = (TP + TN)/(TP + FP + FN + TN))$[4]; percentage of correctly classified machine-error trials (TPR (true positive rate), $TPR = TP/(TP + FP)$); percentage of correctly classified non-error trials (TNR (true negative rate), $TNR = TN/(TN + FN)$); and in addition, as the AUC (area under curve) of the receiver operator curve (ROC) since this measure can be more informative when reporting classification results based on data with imbalanced number of observations per class.

Cross-session validation: Furthermore, we performed a cross-session validation for each participant in two steps: calibration with the cursor data and testing on the robot data, calibration with the robot data and testing on the cursor data. This analysis allowed us to obtain an estimate of how well a participant-specific ErrP-decoder would perform in classifying data of one session if it was calibrated with data of the other session. The number of trials per class was balanced in the calibration data by random pick and replace. To increase the likelihood that most of the trials were used for calibration at least once, this procedure was repeated 100 times. The weight vectors w and biases b of the resulting 100 individual rLDAs were averaged to obtain a final rLDA. Testing was performed on the unbalanced data of the test session. Results are reported identical to the within-session validation as Acc, TPR, TNR, and AUC.

4.3. Results of behavioral data, ERPs, and single-trial classification

4.3.1. Consistency of behavioral data

Participants' individual data are summarized in Supplementary Tables 11 and 12. Mean reaction times did not differ significantly between scenarios ($\overline{RT}_{cursor} = 421 \pm 75ms$, $\overline{RT}_{robot} = 403 \pm 43ms$, AVG±SD, Wilcoxon's signed-rank test, $p = 0.123$). Number of human-committed errors did not differ significantly between scenarios ($\overline{nErr}_{cursor} = 17 \pm 14$, $\overline{nErr}_{robot} = 16 \pm 8$, AVG±SD, Wilcoxon's signed-rank test, $p = 0.916$). Mean reaction times did not significantly vary between the first and second performed scenario irrespective of type ($\overline{RT}_{first} = 415 \pm 52ms$, $\overline{RT}_{second} = 410 \pm 71ms$, AVG±SD, Wilcoxon's signed-rank test, $p = 0.240$), similarly with number of human-committed errors ($\overline{nErr}_{first} = 13 \pm 10$, $\overline{nErr}_{second} = 19 \pm 13ms$, AVG±SD, Wilcoxon's signed-rank test, $p = 0.095$). We conclude therefore that there is no effect of type of scenario or scenario order on reaction times and number of human-committed errors.

4.3.2. Modulations of ERPs

The results of the analysis of ErrPs and stimulus-dependent signal variations are depicted in Figure 19. The left and middle panels show the grand average ERPs for each category

---

[4] True Positive (TP) = correctly classified machine-error trial, True Negative (TN) = correctly classified non-error trial, False Positive (FP) = machine-error trial falsely classified as non-error trial, False Negative (FN) = non-error trial falsely classified as machine-error trial.

(blue: non-error, red: machine-error) and the difference average (dashed black: machine-error minus non-error) for each scenario (left panel: cursor, middle panel: robot). In both scenarios, the shape of the difference grand averages were similar to those previously reported [102, 103, 104] with regard to the timing and topographical distribution of the components N200 and P300. The N400 component was not observable in our data. Instead, we observed another late positive component with a latency of approximately 500 ms. This effect might be related to the P600 component which has been reported in the context of error processing as well as in response to syntactic and semantic anomalies [107, 108]. These findings indicate that the observed effects originated from error-/performance monitoring processes. The difference ERPs of the robot scenario were significantly less pronounced than those in the cursor scenario in that the grand average time courses of the robot scenario were attenuated in peak amplitudes and topologically less clearly distinguished than in the cursor scenario. This is also reflected in the results of the analysis of the coefficient of determination r^2 based on channel Cz. In both scenarios, grand average r^2-values were highest at similar time points (cursor: $r^2_{max} = 0.14$ at $t = 261ms$, robot: $r^2_{max} = 0.04$ at $t = 263ms$). The similarity analysis revealed a correlation coefficient of $r = 0.70$ between spatio-temporal difference grand averages of both scenarios. Computation of the 2D correlation coefficient between spatio-temporal difference averages for each participant resulted in a median $r = 0.48$ and exclusively all participants with a positive 2D correlation coefficient (participant-individual results are reported in Supplementary Table 13, Appendix A). We conclude that the ErrPs observable in the two scenarios were qualitatively similar in terms of shape, timing, and topographical distribution. This indicates that the observed effects originated from the same underlying neuronal process.

4.3.3. Single-trial classification results

The within-session single-trial classification results are depicted in Figure 20 (left panel) and detailed in Supplementary Tables 14 and 15 (Appendix A). Across-participant average classification performance for the cursor scenario resulted in $Acc = 90.6 \pm 3.9\%$ with $TPR = 87.3 \pm 4.3\%$, $TNR = 92.2 \pm 3.8\%$ and $AUC = 0.95 \pm 0.03$ (AVG±SD). For the data of the robot scenario, we observed significantly reduced and more variant within-session classification performance of $Acc = 69.0 \pm 7.9\%$ with $TPR = 66.1 \pm 6.5\%$, $TNR = 70.6 \pm 9.1\%$, and $AUC = 0.73 \pm 0.1$ (AVG±SD). The reduced classification performance based on the data of the robot scenario was consistent across participants ($Acc = -21.6 \pm 7.8\%$, $TPR = -21.2 \pm 8.1\%$, $TNR = -21.7 \pm 8.2\%$, $AUC = -0.21 \pm 0.1$, AVG±SD), however, with non-significant correlation between classification results of each session ($r_{Acc} = 0.27$, $p = 0.41$; $r_{AUC} = 0.17$, $p = 0.61$, Pearson's correlation, $n = 11$). All participant-individual classification accuracies were above the sample size adapted chance-level of $Acc_{chance} = 53.6\%$ ($p < 0.05$) for binary classification according to [183].
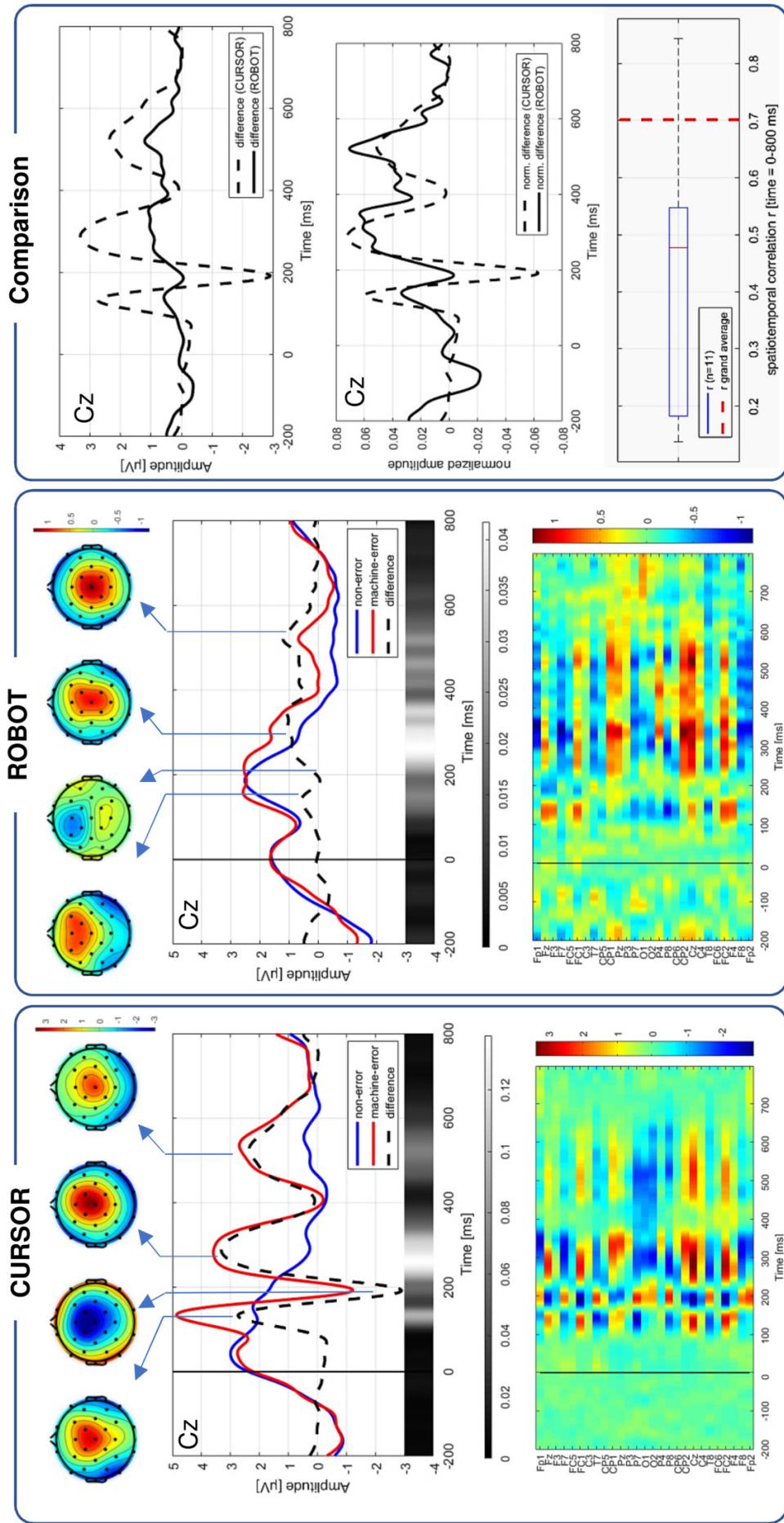
Figure 19 Grand average signals over Cz time-locked to the onset of feedback (cursor movement, robot head turn) for each category (blue: non-error, red: machine-error) and the difference grand average (dashed black: machine-error minus non-error), for both scenarios (left panel: cursor, middle panel: robot). The r^2-values for between non-error and machine-error are depicted below each plot, where bright colors indicate high values. The difference grand average is furthermore depicted as topographic plots for the main peaks above each plot and in form of a spatio-temporal activity matrix across all channels and time points below each plot. The right panel shows the resemblance of difference (and normalized difference) grand averages of channel Cz of both scenario as well as the across-participant distribution of 2D correlation coefficients between difference average of cursor and robot scenario with median $r = 0.42$. The dashed red line depicts the difference grand average 2D correlation coefficient $r = 0.70$. Please note that different axes scaling was used to facilitate qualitative comparisons.

The cross-session single-trial classification results are depicted in Figure 20 (right upper and lower panels) and detailed in Supplementary Tables 16 and 17 (Appendix A). For calibration with the cursor data and testing with the robot data, across-participant average classification performance resulted in $Acc = 68.3 \pm 6.1\%$ with $TPR = 34.9 \pm 13.0\%$, $TNR = 86.4 \pm 10.8\%$, and $AUC = 0.68 \pm 0.11$, AVG±SD (Figure 20, right upper panel). Cross-session classification performance showed high correlation with participant-individual spatio-temporal ERP similarity measures ($r_{Acc} = 0.80$, $p = 0.003$; $r_{AUC} = 0.79$, $p = 0.003$, Pearson's correlation, $n = 11$). Except for participant s10, classification accuracies were above chance-level, however, with a systematic bias between TPR and TNR: across all participants, we observed consistent low classification rates for class machine-error and high classification rates for class non-error. This is most likely related to shifts in the distributions of features favored by the rLDA classifier for separating the cursor data, causing the decision boundary to favor one class over the other in the robot data.

For calibration with the robot data and testing with the cursor data, average classification performance resulted in $Acc = 73.1 \pm 13.0\%$ with $TPR = 70.3 \pm 24.8\%$, $TNR = 74.7 \pm 11.6\%$, and $AUC = 0.78 \pm 0.18$, AVG±SD (Figure 20, right lower panel). No classification bias was observed for the robot-to-cursor transfer, but a decrease of accuracy compared to the within-cursor validation accuracies. This is most likely related to the rLDA favoring features in the robot data which have less discriminative power in the cursor data. Cross-session classification performance showed high correlation with participant-individual spatio-temporal ERP similarity measures ($r_{Acc} = 0.57$, $p = 0.06$; $r_{AUC} = 0.61$, $p = 0.05$, Pearson's correlation, $n = 11$). In all but two participants (s08, s10), classification accuracies were close to or above 70% with an $AUC > 0.73$ and no systematic bias between TPR and TNR. The fact that s08 and s10 revealed also very low accuracies in the within-session validation of the robot data may explain why their cross-session classification results turned out to be low as well.
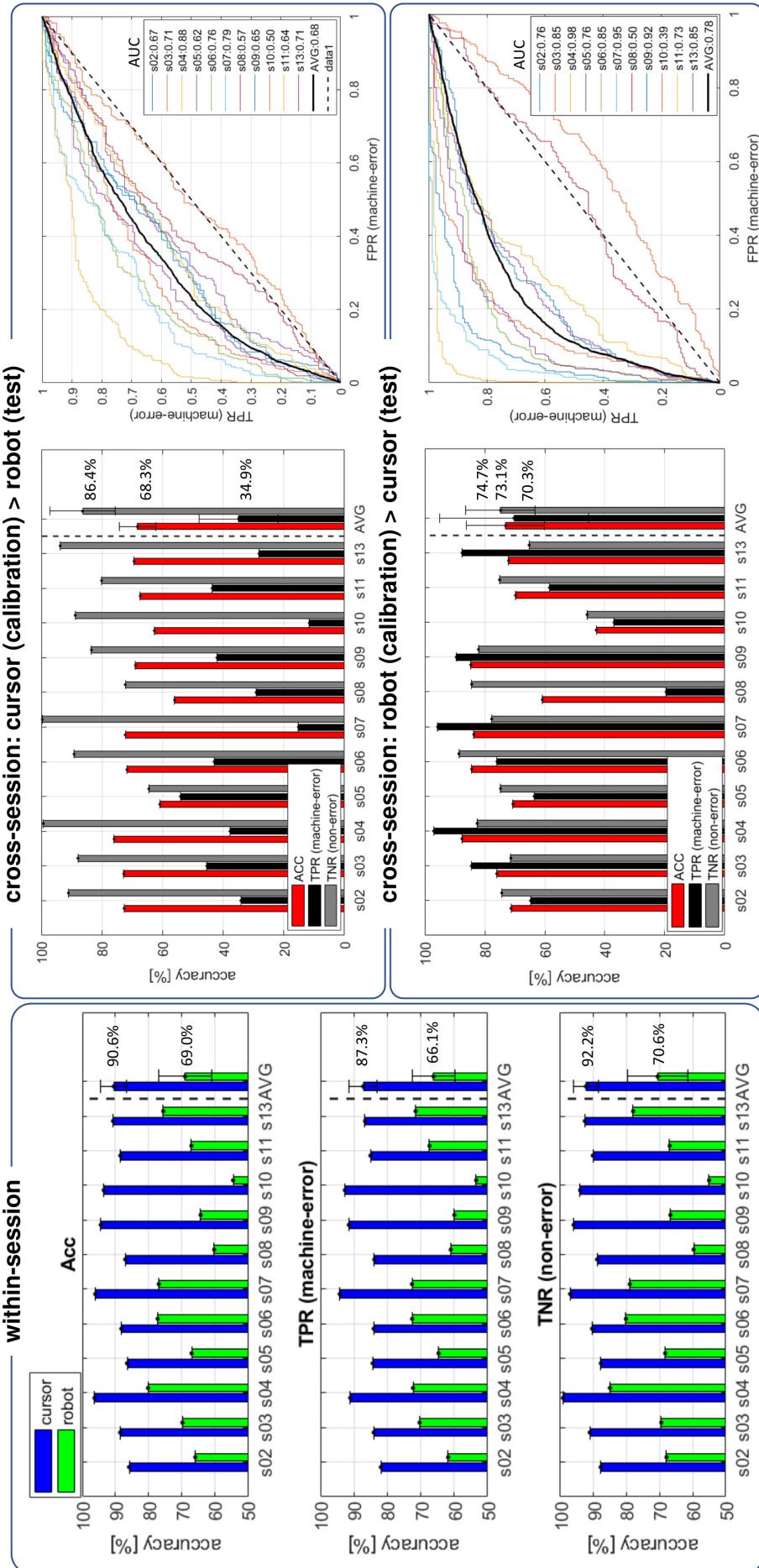
Figure 20 Single-trial classification results: Per participant and average within-session classification results (left panel). Per participant and average cross-session classification results with calibration based on cursor data and validation on robot data (right upper panel) and calibration based on robot data and validation on cursor data (right lower panel). Classification results are reported as percentage of correctly classified machine-error events (TPR), percentage of correctly classified non-error events (TNR), total percentage of correctly classified events (Acc), and area under receiver operator curve (AUC). Average results across participants are highlighted in separate bars (AVG) and are numerically highlighted; error bars represent the single standard deviation across participants. Participant-individual results are further detailed numerically in Supplementary Tables 14-17 (Appendix A).

4.4. Discussion of ERP modulations and single-trial classification results

ErrPs are decodable in response to incorrect robot actions, but decoding performance is sensitive to the stimulus:

The study confirmed feasibility of decoding ErrPs in response to the human observation of incorrect robot action. We obtained a classification accuracy of average $Acc_{robot} = 69.0 \pm 7.9\%$ across 11 participants, which is comparable to results in response to robot actions obtained by others [105, 173, 174]. For the cursor task, we obtained an average classification accuracy of $Acc_{cursor} = 90.6 \pm 3.9\%$ which is comparable or higher than previously reported single-trial classification results based on ErrPs in response to screen-based stimuli [102, 103, 104]. From these results, we conclude that the observability and decodability of ErrPs is sensitive to type of stimulus eliciting them, even in procedurally simplistic tasks. Besides the lower accuracy in the robot task, we observed higher variations of decoding accuracy indicating that participants seemed to have reacted more differently to the robot than to the cursor, e.g. s04, s06, s07, s13 with $Acc_{robot} > 75\%$ and s08, s09, s10 with $Acc_{robot} < 65\%$, and s10 close to chance level (see Supplementary Table 15, Appendix A).

The observed quantitative variations of ErrPs and the respective differences of decoding performance across experimental tasks denote an interesting outcome of our study. Since the realization of both the cursor and robot actions were implemented to be as similar as possible (see Section 4.2.3), technical factors can be largely excluded. Beyond technical factors, the observed differences may have resulted from an across-task specific perceptual jitter, e.g. greater trial-by-trial latency variations of the observed ErrPs in response to the robot actions as compared to the cursor actions. However, no evidence of such variations could be found in a post-hoc analysis in which we attempted to rectify the phases of single-trial ErrPs without notable success. Given these insights, it appears more likely that the observed variations resulted from systematic differences in the perception and/or evaluation of the two types of actions across participants. Whether and to what extent such perceptual or cognitive factors affect the observability and decodability of ErrPs in HRI requires further attention in follow-up studies and are discussed in more detail in the outlook of this thesis (see Chapter 7).

Despite the observed quantitative differences, the ErrPs extracted from the two scenarios were qualitatively similar in terms of shape, timing, and topographical distribution. This indicated that the observed effects originated from the same underlying neural process. Based on this finding, we tested whether ErrP-decoders calibrated with the data of one task can be re-used to classify responses based on the other task. The cross-session single-trial classification analysis resulted in a classification bias for the cursor-to-robot transfer (high classification rate for non-error and low classification rate for machine-error events) consistent across participants, and a significant drop in classification accuracy for the robot-to-cursor transfer compared to the respective within-session results. Based on these results, we do not recommend

a straightforward re-use of the ErrP-decoders across tasks without re-calibration.

Despite the observed variation between experimental tasks, our results demonstrated and confirmed feasibility of decoding ErrPs in response to the observation of semantically incorrect robot actions given that ErrP-decoders are calibrated based on data of the same task.

ErrPs can be used in more complex HRI scenarios:

We purposely designed a relatively simple HRI scenario to demonstrate principal feasibility of decoding the human perception of semantically incorrect robot actions. Furthermore, the chosen experimental paradigm allowed us to compare the ErrP responses in the robot task to those obtained in response to a simplistic stimulus in a procedurally identical task. The robot scenario in our study resembles most that of a real-world HRI task, in which the robot designer is interested in validating the congruency of robot gaze cues with the human partner's expectations. In a follow-up study (see Chapter 5) we investigated to which extent the results of the present study are transferable to such a more realistic HRI task. The results of the study showed that ErrPs, online decoded from the human interaction partner, can be used to both validate the congruency of robot gaze cues with the human partner's expectations and likewise, to successfully adapt the robot's gaze behavior during interaction. Whether and to what extent the outcomes of our present study scale to even more complex HRI scenarios with different types of robot actions remains open for future work. In light of the findings we can, however, state that the brain responses we observed in both experimental tasks were related to error / performance monitoring due to their temporal and topographical similarity with previously reported ErrPs in different experimental paradigms [102, 104]. This supports the notion that the effects observed derive from a high-level "generic" neuronal process, that is understood to be largely unrelated to the situational context or the associated stimulus [98, 184].

Practicality in ErrP-based validation of HRI denotes challenges to overcome:

A few aspects render the method of ErrP decoding for robot validation laborious and impractical to be readily deployable: Firstly, the cumbersome, expensive, and sensitive EEG setup, and secondly, the necessity for participant-specific (re-)calibration of the ErrP-decoder. The need for inexpensive and easy-to-use EEG systems with sufficient signal quality has already been recognized by the BCI community [185]. Along this line of research, we made a contribution in the form of the development of a simple, mobile, and comparably inexpensive (~800USD) EEG system. Our device was initially deployed in a study investigating sensorimotor rhythms in patients with CP (cerebral palsy) [186]; the usability for measuring and decoding ErrPs has also been successfully validated in a follow-up comparative study (see Appendix B). Participant-specific (re-)calibration is a generally recognized issue thwarting practicality of brain-computer interfaces [136]. Non-stationarities and dissimilarities of EEG signals across recording sessions and participants generally render (re-)calibration a necessity to assure suf-

ficient functionality of BCIs. ErrPs have, however, been shown to be stable across recording sessions within the same participant and experimental task up to 600 days [102, 151]. This indicates that re-using ErrP-decoders without re-calibration is principally possible if participant and task remain the same across experimental sessions. With regard to participant-to-participant transfer we recently investigated the feasibility for participant-independent ErrP-decoders and demonstrated the effectiveness of supervised and unsupervised online adaptation of such generalized ErrP-decoders [164, 165].

4.5. Summary

In this chapter we presented a neuronal-based method for quantitatively assessing in real-time the human perception of robot actions during HRI. The method is based on the decoding of ErrPs from EEG signals. The usability for the assessment of robot actions was tested by comparing the ErrPs generated in response to incorrect robot actions with incorrect computer screen-based cursor actions. The results confirmed the decodability of ErrPs upon observation of incorrect robot actions (average classification accuracy of $Acc = 69.0 \pm 7.9\%$), although with a lower accuracy than upon observation of incorrect cursor actions (average classification accuracy of $Acc = 90.6 \pm 3.9\%$). This demonstrated that ErrPs can be sensitive to the stimulus eliciting them despite procedurally identical protocols. The qualitative similarity of ErrPs generated in the two scenarios indicates nevertheless a common neuronal origin of the signals. Overall, the study demonstrated that assessment of incorrect robot actions via ErrPs is feasible, but the re-use of ErrP-decoders across experimental tasks was not supported.

5. Human-agent co-adaptation using error-related potentials

ErrPs decoded from the ongoing EEG of a human observer have been proposed as an intuitive feedback signal for improving HMI (e.g HCI or HRI). While the use of ErrPs as a teaching signal for robot skill learning has been successfully demonstrated by recent studies by others, so far, no efforts have been made towards HRI scenarios where mutual adaptations between human and robot are expected or required. These are collaborative or social interactive scenarios without predefined dominancy of the human partner and where robots are perceived as intentional agents. Here we explore whether ErrPs can be used as a feedback signal from the human for mediating co-adaptation in HRI. We experimentally demonstrate ErrP-based mediation of co-adaptation in a closed-loop HRI study where successful interaction depended on co-adaptive convergence to a consensus between them. While participants adapted to the robot by reflecting upon its behavior, the robot adapted its behavior based on ErrPs decoded online from the human partner's ongoing EEG. ErrPs were decoded online in a single trial with an average accuracy of $Acc = 81.8 \pm 8.0\%$ across 13 participants, which was sufficient for effective adaptation of robot behavior. Successful co-adaptation was demonstrated by significant improvements in interaction efficacy and efficiency, and by the robot behavior that emerged during co-adaptation. These results indicate the potential of ErrPs as a useful feedback signal for mediating co-adaptation in HRI as demonstrated in a practical example.

5.1. Introduction

Recently, ErrPs have been proposed as a feedback signal from the human for guided adaptations of physical robotic systems [171, 172, 105, 42, 174, 173, 43]. The basic concept is to harvest ErrP responses from a human observer upon recognition of erroneous or inappropriate robot actions in order to adapt or improve the robotic system post-hoc or on-the-fly. This approach is particularly promising as a complementary method for adapting robotic systems and HRI, because: (1) ErrPs are naïve responses which require no mental effort from the human observer. (2) ErrPs can be decoded in real-time, allowing for online adaptations of the robotic device without interruption of ongoing interaction with the human partner. (3) ErrPs are understood to be sensitive to violations of expectations [187, 188] and as such comprise an implicit and immediate feedback, informative (3a) for improving the robotic system to better align with the observer's expectation, and (3b) possibly informative with regard to the observer's overall assessment of the robotic system and/or the quality of interaction. The following section provides an overview of research works that made use of this principle. Here, a distinction is made between the deployment of ErrP decoding as an additional modality for

---

adaptation of BCI decoders (hybrid BCI approach) and the deployment of ErrP decoding as single modality for adaptation of robotic systems and HRI.

5.1.1. Related work on ErrP-based adaptation in HMI

ErrP-based adaptation of hybrid BCIs:

Schalk and colleagues were among the first in 2000 to propose the use of ErrPs for online improvements of BCIs. They demonstrated that ErrPs occur in response to the participant's observation of the BCI delivering wrong output, e.g. mismatching the participant's intended command the BCI was ought to execute [189]. Since then, several works have demonstrated improvements in the information transfer rate of BCIs, via simultaneous ErrP decoding, both in healthy as well as in patient populations. One can distinguish here between BCIs for target selection (e.g. based on the P300 oddball paradigm or VEPs), and BCIs for continuous control (e.g. based on SCPs, SMRs, or attention-based paradigms). An overview of common BCI paradigms is provided in Chapter 2, Table 2. While early works mainly focused on correcting or reversing erroneous BCI output, later works used ErrP-based feedback to adapt BCI decoding models during operation. The simultaneous decoding of two or several modalities, for instance, sensorimotor rhythms for BCI control and ErrPs for correcting output of such sensorimotor decoding, has been termed *hybrid BCIs* (for an overview, the reader is referred to [190]).

Parra and colleagues demonstrated in 2003 improved performance of 21% in participants executing a binary forced choice visual discrimination task, when errors were automatically corrected via online decoded ErrPs [191]. Blankertz and colleagues demonstrated in 2003 improved information transfer rates when correcting decoding of motor intent from pre-movement ERPs by means of post-movement ErrPs [192]. Similarly, in their work of 2005 and later, Ferrez and Millán demonstrated improved information transfer rates in sensorimotor BCIs by correcting erroneous BCI output using ErrPs [169, 102, 193, 194]. Schmidt and colleagues demonstrated a multi-target selection BCI based on the P300 oddball paradigm and showed that ErrPs could be used to detect wrong target selection, as such increasing spelling speed on average by 49% across eleven participants [195]. A similar approach was presented by Spueler and colleagues in 2012, in this study including patients with severe motor impairment due to ALS [196]. In a recent work by Cruz and colleagues this approach was even shown to work in a tetraplegic patient [197]. An integrated approach for EEG-driven position control of a robot arm for neuroprosthetic and rehabilitation purposes was presented by Bhattacharyya and colleagues in 2014 [198]. By employing three modalities (sensorimotor rhythms for controlling, P300 for stopping, and ErrPs for correcting robot arm movements) they showed efficient and precise robot arm control in five healthy participants. One of the first works demonstrating ErrP-based online adaptation of BCI decoders was presented by Llera and colleagues in 2011 [153]. Their BCI decoded covert attention based on lateralized alpha activity for binary control in a forced choice task. ErrPs were shown to be useful to improve the decoding model for covert attention during operation. Similarly, Spüler and colleagues

showed in 2012 that also decoding models for code-modulated VEP based BCIs for character spelling can be online adapted using ErrPs [170]. With this integrated approach, they yielded the highest information transfer rates achieved with non-invasive BCIs to date (144 bit/s). More recent work was presented by Mousavi and colleagues in 2017 [199]. Instead of employing separate decoding models for sensorimotor rhythms and ErrPs, they proposed an integrated model capturing both processes and showed significant improvement of classification accuracy across 10 participants. While the approach of using simultaneous decoding of ErrPs for improving BCI performance is highly convincing, recent works in this area are scarce. A reason for this may be that the improvements achieved by above mentioned works do not entirely justify the increased complexity of hybrid BCI systems.

ErrP-based adaptation of robotic systems and HRI:

In 2008, Chavarriaga, Ferrez and Millán proposed that ErrPs may constitute a learning signal to infer the participant's intended strategy ([200], see also [151]). Their experimental protocol comprised of a screen-based cursor that could move in discrete steps either towards or away from a target (similar to the setup shown in Figure 21(a), left). By using online decoded ErrPs generated whilst the participant monitored the cursor (feedback signal for reinforcement learning), they showed that it is possible to learn a policy that effectively moves the cursor towards the target. This same principle was demonstrated also in more complex scenarios. For instance, Iturrate and colleagues used an experimental protocol with a simulated robotic gripper displayed on a computer screen that could deliver objects into 5 locations [171]. Upon wrong delivery of an object (mismatching the participant's intended location), an ErrP could be detected and used for reinforcement learning of the correct location. Similarly, Chavarriaga and Foerster showed in 2010 that a pre-trained gesture recognition system based on hand motion sensors could be improved using online decoded ErrPs [201, 202]. The first known work employing online decoded ErrPs in an HRI scenario involving a real robotic system (instead of virtual) was presented by Iturrate and colleagues in 2010 [171]. They used a robotic manipulator (Katana 300) that could be moved horizontally in 5 pre-defined positions (similar to the simulated gripper in [41]) and showed the presence of reliably decodable ErrPs when movements did not correspond to those intended by the participants. This work was extended in 2013 to include movements to occur in 2D (instead of 1D horizontal movements) with a 5x5 grid world setup similar to the setup in Figure 21(a), middle ([203], see also [204]), and to work also in a real robotic manipulator [105] (see Figure 21(a), right). In 2017 Salazar-Gomez and colleagues [173] showed that a robot could learn a binary sorting task by means of ErrPs online decoded from a human observer (see Figure 21(b)). The closeness of the experimental setup to a real-world HRI made this work an outstanding example, despite the principle not being novel. In a similar line, Kim and colleagues showed in 2017 [43] that a robot could learn, based on online decoded ErrPs, to recognize and replicate human gestures (see Figure 21(c)). In their work, they particularly emphasized the potential of ErrP-based HRI as a set-up for "intrinsic interactive reinforcement learning of robot skills" without particular effort required from the human interaction partner.

Besides robotic systems, online decoded ErrPs have also been employed in car driving tasks by Zhang and colleagues, for example, to infer the driver's turning direction, both in a simulated [205] and a real driving situation [206]. In a similar line, a recent work by Choi and colleagues demonstrated successful decoding of ErrPs when the car moved differently from how the user intended to drive [207].
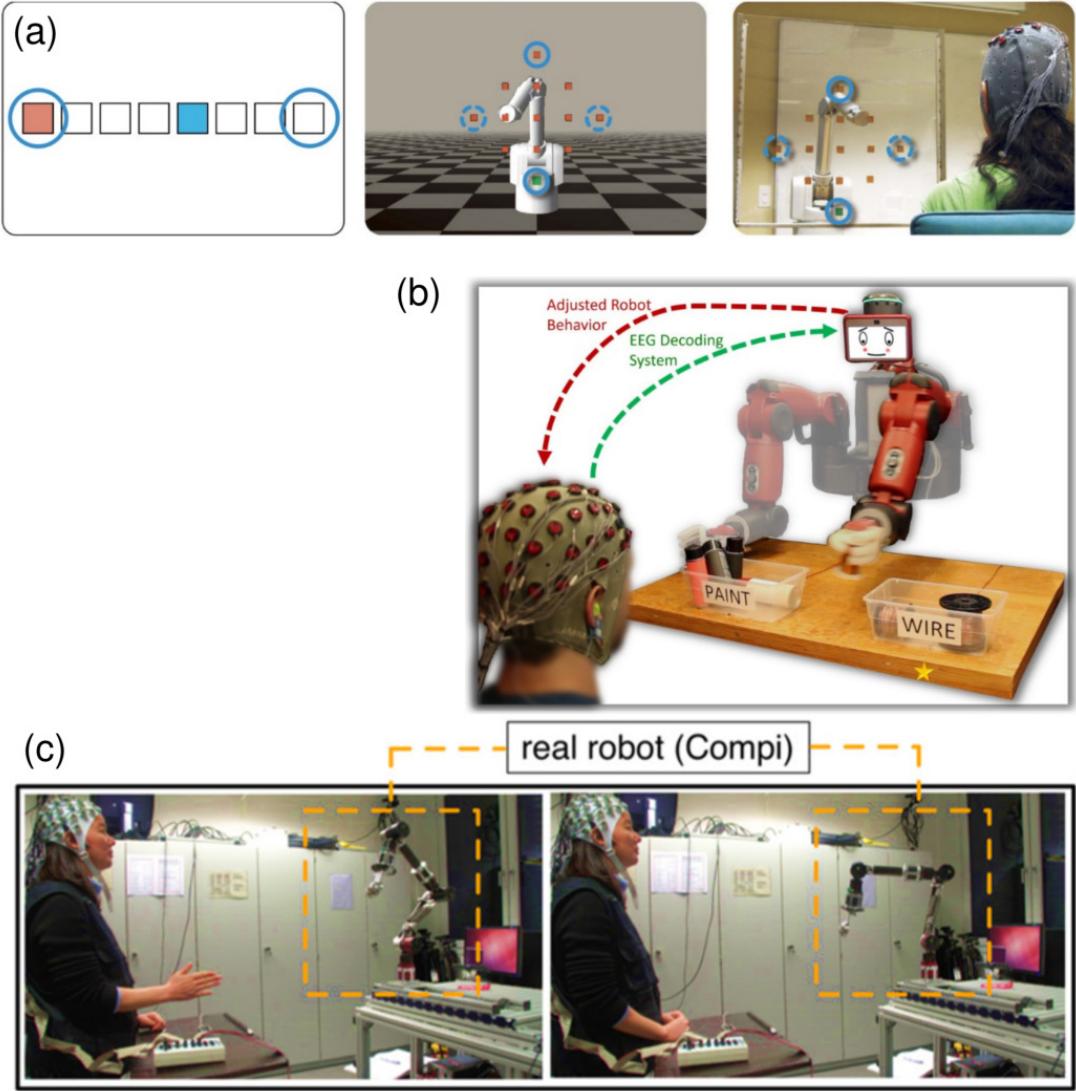
Figure 21 ErrP-based adaptation of robotic systems and HRI: (a) ErrP-based learning of robot end-effector trajectories: (left) screen-based cursor movement towards target, (middle) 2-dimensional reaching task in simulated robot environment, and (right) 2-dimensional reaching task with real robotic system [Adapted from Iturrate et al., 2015 [105]]. (b) ErrP-based learning of robot performing a binary sorting task: ErrPs, online decoded from a human observer are used as feedback to learn the correct assignment of objects to corresponding boxes. [Reprinted from Salazar-Gomez et al., 2017 [173]]. (c) ErrP-based learning of robot recognizing and replicating human gestures: ErrPs, online decoded from a human interaction partner are used as feedback for robot learning of the correct assignment of three human gestures and the correct replication of these gestures [Reprinted from Kim et al., 2017 [43]].

In summary, while above mentioned works showed promising results of this highly innovative approach, they were primarily concentrated on using ErrPs as a teaching signal for robot skill learning. A question that remained unexplored is whether this ErrP-based feedback signal can also be useful in situations where both human and robot are required to adapt to each other to converge to a consensus in the given joint task. That is, situations in which there is

no explicit "right" behavior (policy) the robot is supposed to be taught, but the human partner may as well adapt to the robot. By providing feedback, e.g. for reinforcement learning of robot skills, the human participant represents a critic and as such the reward function for the learning system. In all previous works, the reward function was either defined externally (by task instructions) or by employing designs involving unambiguous tasks (single explicitly correct reward function). In HRI/HMI scenarios where machines appear as autonomous agents, human interaction partners may also adapt to or entirely rely on the system. This may result in reward functions being volatile and changing dynamically. How to formalize and deal with such volatile reward functions denote a challenge implicating both technical, but also human factors. Approaching this question is important regarding human interaction with systems that have a form of intentional agency, such as HRI in the context of collaborative or social interactive scenarios. This contrasts to interaction scenarios in which robotic systems are used as tools supposed to fulfill an explicit function (explicit "right" policy).

5.1.2. Aim of the work

The current study explored the usability of ErrPs as a feedback signal in the context of human-agent co-adaptation. The approach used, schematically described in Figure 22, conceptually assumes the interaction between two partners: (1) An intentional artificial agent with a policy π determining its behavior based on a set of behavioral states S, actions A, and goals / intentions G. (2) A human partner, interacting with that agent based on a belief of the agent's policy π'. While the agent is provided feedback through online decoded ErrPs to gradually adapt its policy π to the human's belief π', the human partner may gradually adapt his/her belief π' to the agent's policy π by reflecting upon its behavior. As such, both systems (human and agent) are adaptive, allowing for mutual adaptation with the aim to converge to a consensus in form of an alignment of the human's belief and the agent's actual policy: $\pi_{final} \approx \pi'_{final}$.

The conceptual approach (Figure 22) was implemented in form of a human-robot collaborative guessing game where the human partner had to guess, from a humanoid robot's gazing behavior, which of three available objects was selected by the robot. While the human's task was to learn to infer the robot's intentions / goals by observing and interpreting its gazing behavior, the robot's task was to learn to convey its intentions / goals via gazing behavior to the human partner; efficient interaction required their convergence to a consensus by co-adaptive learning of both parties. Crucially, the sole modality of information provided to the robot was ErrPs, online decoded from the ongoing EEG of the human partner. This approach was explored in a closed-loop experimental study and, as being demonstrated in the following sections, enabled such co-adaptive human-robot learning, indicated by significant improvements in HRI performance. This extends previous accounts on ErrP-based adaptation of robotic systems and HRI by addressing the following aspects:

- Demonstration of the usability of ErrPs for mediating co-adaptation in HRI. This relaxation of interaction constraints - permitting mutual adaptation - is particularly important with regard to HRI scenarios where the human partner does not have a predefined dominant role
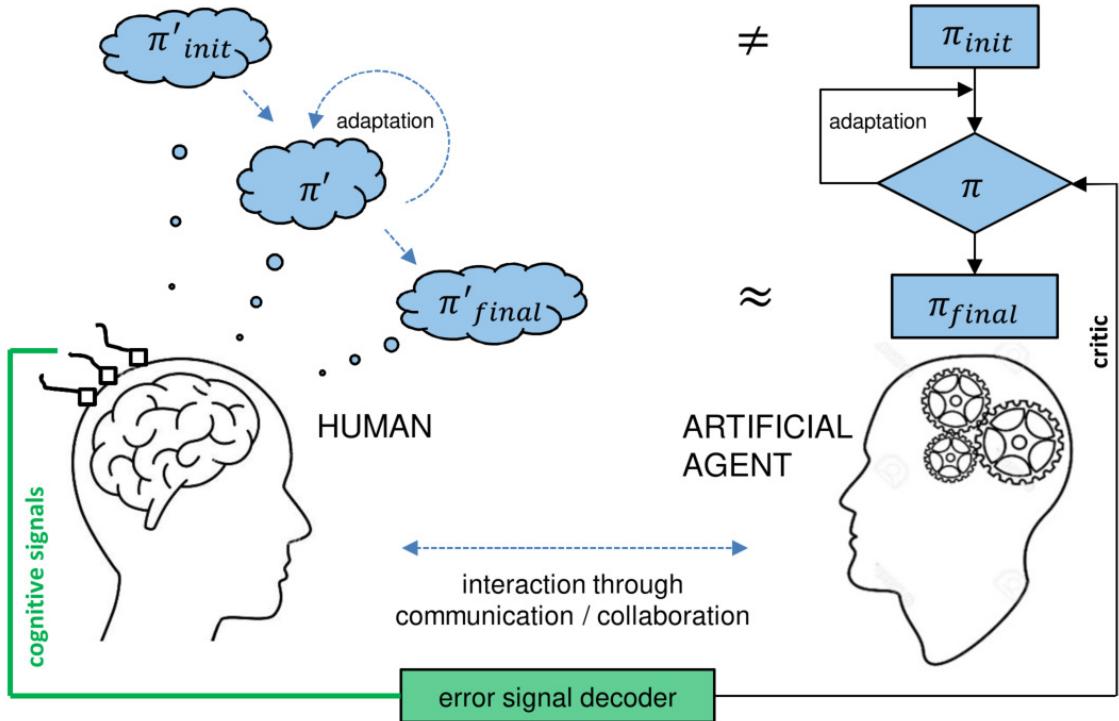
Figure 22 Conceptual approach: in interaction with an agent, the human holds a mental model (belief) of the agent's policy π' to predict its future behavior, which can be based on prior expectations π'_{init} and is further adapted during interaction. ErrPs, online decoded from neural activity of the human partner are provide as a critic for guided adaptation of the agent's actual model π. This creates a two-party co-adaptive system allowing both human and agent seeking consensus in form of an alignment of the human's belief and the agent's actual policy: $\pi_{final} \approx \pi'_{final}$.

(principal or teacher role). These are scenarios in which adaptations of the human to the robot are expected or even necessary for successful interaction, such as in collaborative or social interactive HRI.

- Previous works adapted the robot's behavior using ErrP feedback based on single, explicitly erroneous robot actions [105, 173, 43]. In more complex robot behavior, however, individual robot actions are more likely to occur in rapid succession and not to be temporally well isolated, the latter being a prerequisite for reliable ErrP decoding. Here we explore robot adaptation based on ErrPs arising from and reflecting the human's interpretation of the robot's intention / goal, with the latter being inferred from a sequence of robot actions instead of a single occurrence. Along this line, this work proposes and successfully employs an ErrP-based episode update strategy with delayed reward for online adaptation of the past sequence of robot actions.

This chapter is structured as follows: in the subsequent sections the experimental paradigm (Section 5.2.1), design and tasks (Section 5.2.2) are described in detail, followed by a thorough description of the implementation of the technical components of our approach (EEG-based online decoding of ErrPs and corresponding adaptation of robot behavior) in Section 5.2.3. The main results of efficacy of online ErrP decoding and human-robot co-adaptation are reported in Section 5.3, followed by a discussion of the results in Section 5.4. Section 5.5 summarizes the chapter.

5.2. Methods: closed-loop BCI setup and experimental study 3

5.2.1. Experimental paradigm

The experimental paradigm is schematically described in Figure 23(a). Three objects were located between the participant and the robot. The robot would select one among the three objects (unknown to the participant) which denoted its covert intention / goal g and subsequently started executing a gaze pattern (action sequence) by turning its head (actions: A) towards the objects and the participant (states: S). The participant's task was to guess the robot's initially selected object from observing its gaze behavior: The participant may for instance consider which object the robot fixated more often or for the longer duration. Eventually, the robot would reveal the actual object it had initially selected, resulting in the participant experiencing a match or mismatch with the object he/she believed the robot selected. In that moment the participant's ongoing EEG signals would change, reflecting the robot's revelation as an error- (mismatch) or non-error (match) response and used as a negative or positive reward for adaptations of the robot's gaze behavior policy π. Meanwhile, the participant may update his/her prior belief π' about the robot's gaze behavior to improve guessing in subsequent games. We hypothesize that by feedback the ErrP to the robot this would iteratively adapt its gazing behavior towards a pattern that facilitates the participant to correctly infer the robot's selected object. To what extend this convergence is driven by the human adapting to the robot or the robot adapting to the human is deliberately kept flexible to investigate the feasibility of ErrP-based mediation of co-adaptation as outlined above.

5.2.2. Experimental design

5.2.2.1 Participants

Eighteen healthy participants took part in the study. The data of the first two participants were discarded due to technical problems during the experiment. The remaining sixteen participants were 7 females and 9 males of average age: 29.2 ± 5.0, and all right-handed. All participants were informed about the tests before its conduction. They participated voluntarily and gave written consent. Participants were paid an honorarium of 8 EUR/h for their participation. The study was approved by the institutional ethics review board of the Technical University of Munich (reference number 236/15s).
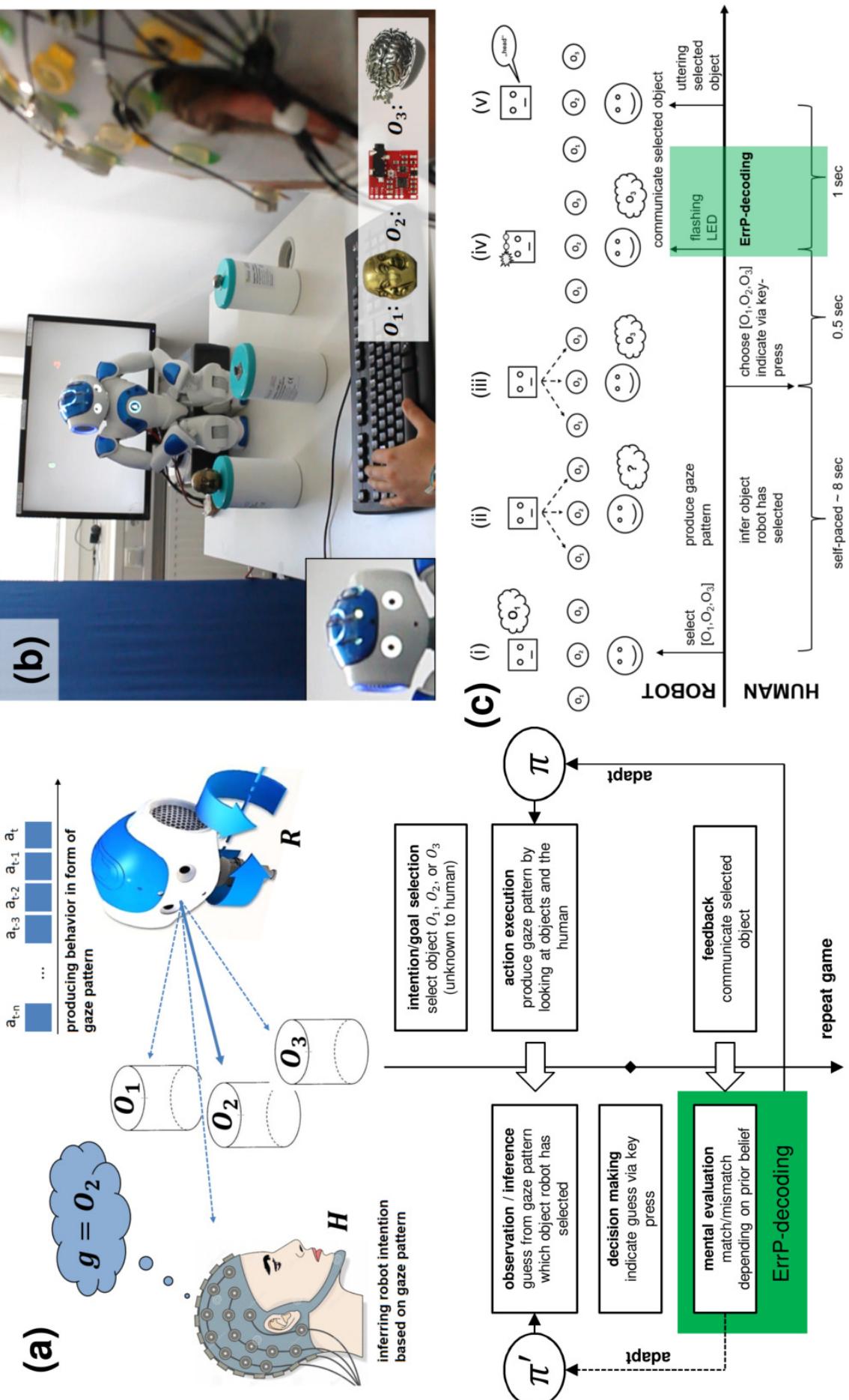
Figure 23 (a) Experimental paradigm and task: Human participant and robot play a guessing game in which the robot produces a gaze pattern based on which the participant has to guess the secret object. The participant's brain responses are measured (marked in green) and used as a feedback signal to adapt the robot's gaze behavior policy π, while the participant may adapt the prior belief π' about the robot's gaze behavior policy. (b) Experimental setup from the perspective of a participant. (c) Trial structure of a single guessing game with the corresponding moment of ErrP decoding marked in green.

5.2.2.2 Experimental tasks and protocol

An overview of the experimental tasks is provided in Table 5. Each experiment consisted of two parts conducted in the following order: (1) open-loop calibration session (CALIB), (2) four separate closed-loop co-adaptation sessions (CORL). In both parts, the participant was asked to repeatedly play the guessing game together with the robot (for the remainder of this chapter a single game is also called trial; for technical details about the trial structure the reader is referred to Section 5.2.3.1 and Figure 23(c)). The first part of the experiment (CALIB) had the purpose of collecting EEG data for calibrating participant-specific ErrP-decoders utilized afterwards for online ErrP decoding during the co-adaptation sessions (CORL). During CALIB, the robot's gaze behavior policy was pre-programmed and not adapting; during CORL, the robot's gaze behavior policy was online adapted based on the ErrPs decoded from the participant's EEG signals while interacting with the robot. We opted for an experimental design integrating calibration and online application in a single recording session per participant to avoid the possibility of day-to-day data variability. Furthermore, the number of trials for CALIB and CORL were defined such as to keep the duration of the experiment around one hour maximum to reduce and avoid concentration lapses resulting in data quality degradation.

Open-loop calibration session (CALIB): During this experimental task, the robot's gaze patterns followed a deterministic behavior in which the robot tended to look at the selected object more often or remained gazing at it (for details about the technical implementation the reader is referred to Section 5.2.3.2). With this gaze behavior, participants achieved high accuracies in guessing the robot's selected object ($>$95%). Participants performed in total 150 guessing games (trials) during the calibration session in three blocks of 50 trials each, resulting in a total duration of 20.4$\pm$6.7 minutes. Pilot experiments showed that participants could perform 50 trials in a row without reporting notable concentration drops. The breaks in between blocks allowed the participants to relax and prepare for the next block. To control the number of error events, false feedback was introduced with a probability of $p_{err} = 0.3$. Usually, false feedback rates are chosen around 20% [169, 103, 105]. Here, an increased false feedback rate was used to obtain a higher number of error observations given the limited number of 150 trials. False feedback was realized as wrong robot feedback irrespective of whether the participant guessed correctly. This resulted in approximately 35% error (mismatch) events combining false feedback and participant guessing mistakes. The EEG data recorded during the calibration session was afterwards used to build a participant-specific ErrP-decoder to be employed subsequently for online ErrP decoding in the co-adaptation sessions. During CALIB, participants indicated their guess by key press and received feedback about their correct and wrong number of guesses displayed on a computer screen. Since the calibration session had a comparably long duration and was rather monotonous (non-adaptive robot behavior), this feedback was introduced as a means for self-monitoring to encourage participants to improve and maintain their performance throughout the session.

Closed-loop co-adaptation sessions (CORL): After the calibration session, each participant performed four separate online co-adaptation runs, each consisting of 50 trials, with an aver-

age duration of approximately 6-8 minutes per run (corresponding to the duration of one block during CALIB). We opted for the conduction of several independent runs per participant, as co-adaptation may fail due to uncontrollable and random factors (e.g. stochastic sampling of robot actions, self-paced participant response; further details are described in Section 5.2.3). Furthermore, a single CORL was kept below 10 minutes to restrict participant frustration if co-learning would turn out unsuccessful. In the beginning of each run the robot gaze behavior policy was reinitialized such that gazing was uniformly random distributed among the three objects and therefore allowed for no informed guesses of the selected object (chance-level $p = 1/3$). In each trial, the participant's ErrP response to the robot revealing its initially selected object was decoded online from the ongoing EEG signals and utilized to update/adapt the robot's gazing behavior policy. In three of the four runs (CORL-I, CORL-II, CORL-IV), the procedure was identical to the open-loop calibration session, in that participants indicated their guess by key press. In CORL-III, participants were asked to observe the robot's behavior for a pre-defined time only, without communicating their guesses via key press. We introduced this additional run to investigate whether explicit actions linked to the decisions (consolidated belief) of the participant were required or whether covert beliefs/decisions (sole mental reflection without explicit actions) are sufficient for successful co-adaptation. This modification was introduced to test the usability of the ErrP-based approach in a situation where no explicit information from the participant is available as it is likely the case in more naturalistic HRI scenarios. Unlike in CALIB, in CORL, no feedback on the number of correct guesses was provided to the participants. This had no notable effect on the observability and decodability of ErrPs (see Section 5.3.1, and Figure 24(b)).

Instructions: The participants were asked to guess the robot's selected object by inferring information from its gazing behavior. They were, however, not given any specific hints on what to focus in particular. After the calibration session, the participants were furthermore informed that during the online co-adaptation runs the robot's gaze behavior may change as it is subject to adaptations based on their ongoing brain activity. No further details about the implementation of the experiment were provided.

5.2.3. Stimuli, apparatus, and closed-loop BCI

5.2.3.1 Systems overview

Figure 23(b) shows the experimental setup from the perspective of a participant sitting approximately 150 cm in front of a humanoid robot. The robotic platform chosen for the experiment was the humanoid robot NAO. NAO is a commercially available (SoftBank Robotics) 58 cm tall humanoid robot with 21-25 degrees of freedom [178] that was controlled in this experiment by a program running on an external PC connected to the robot via LAN. The only body part of the robot used was the head with pitch- and yaw-movements. Three arbitrarily chosen physical objects were located on top of cylindrical containers in fixed positions between the participant and the robot (O_1: a metal Buddha head - left, O_2: an electric circuit board - middle, O_3: a metal brain keychains - right). The robot's forehead was equipped with three identical green LEDs geometrically aligned with the three objects and placed in a distance

Table 5 Overview of the experimental tasks and corresponding purpose in the order of conduction, covering part 1 (open-loop calibration session: CALIB), and part 2 (four separate closed-loop co-adaptation sessions: CORL-I,-II,-III, and -IV). The duration of the entire experiment was approximately 1 h per participant, including breaks.

part	short name	short description and purpose	duration
1 Open-loop calibration session	CALIB	**Task**: To guess the robot's selected object, indicating guess via key press. **ntrials**: 150 (3 blocks of 50 trials each) **Purpose**: Collect EEG data for subsequent calibration of participant-specific ErrP-decoders. **Robot gaze policy**: pre-programmed and non-adaptive. Elicitation of ErrPs with random occurrences of false-feedback events with a probability of $p_{err} = 0.3$.	15-25 min.
	ErrP-decoder calibration	Automatic calibration of participant-specific ErrP-decoder based on data collected during CALIB.	5 min.
2 Closed-loop co-adaptation sessions	CORL-I	**Task**: To guess the robot's selected object, indicating guess via key press. **ntrials**: 50 **Purpose**: Online application of ErrP-decoder for mediating human-robot co-adaptation. **Robot gaze policy**: Initial uniformly random gaze behavior; updated after each trial based on the classified outcome of the corresponding online decoded ErrP.	6-8 min.
	CORL-II	Same as CORL-I with reinitialization of gaze policy.	6-8 min.
	CORL-III	**Task**: To guess the robot's selected object <u>without</u> overtly indicating guesses via key press (compared to CORL-I,-II, and -IV). Robot performed gaze behavior for a pre-defined fixed duration **ntrials**: 50 **Purpose**: Online application of ErrP-decoder for mediating human-robot co-adaptation <u>without explicit decisions</u> from the human partner (sole observation and mental reflection of the robot's gaze behavior). **Robot gaze policy**: same as in CORL-I,-II, and -IV with reinitialization of gaze policy.	6-8 min.
	CORL-IV	Same as CORL-I with reinitialization of gaze policy.	6-8 min.

of approximately 1.5 cm from each other (see Figure 23(b), bottom left corner). The LEDs were controlled via a 4-channel digital analog converter (Phidget). We used the LEDs in the experiment as the visual feedback to communicate the robot's initially selected object to the participant (Figure 23(a): "feedback") with a fixed representation (O_1: left LED, O_2: middle LED, O_3: right LED). Subsequently, participants received an additional auditory feedback in form of the robot speaking out the name of the chosen object (O_1: "The head", O_2: "The circuit", O_3: "The brain"). The robot's choice was communicated in two ways for the following reasons: LED-based feedback was introduced because of its saliency and perceptual simplicity, expected to result in more distinct brain-responses as compared perceptually complex or gradually unfolding stimuli which have been reported to result in attenuated ErrP responses [159, 174, 45] (see Chapter 4). The subsequent additional robot speech feedback was introduced to increase participant's engagement in the experiment, as robot talking has been reported to foster engagement in HRI [208]. The participants were instructed to particularly attend the LED feedback. Behind the robot, a computer screen was located which served to provide the participants with additional information about the number of correct (left, green) versus incorrect (right, red) guesses during the calibration session (CALIB). During the co-adaptation runs (CORL) this feedback was not provided. Participant responses (Figure 23(a): "decision making") were performed with the left hand and registered with the following keys of an ordinary computer keyboard (O_1: key "1", O_2: key "2", O_3: key "3"). The experiment was realized in a single program using the Python-based NAOqi-library (robot control), the

Phidgets-library (LED control), Psychopy library (keyboard and screen control) [177] and executed on an Intel® Core i5 CPU 750@2.67 GHz. Furthermore, this program received input from the ErrP-decoder during the co-adaptation runs which was executed on a different PC via the TCP/IP-based *labstreaminglayer* protocol [209].

Trial (guessing game) structure: Figure 23(c) shows the structure of a single trial: (a) the trial started with the robot gazing at the human, and performing a uniform random selection of either one of the objects $g \in G$, as well as a uniform random selection of an initial gaze state $s_{init} \in S$ and (b) subsequently started alternately gazing in a fixed pace at the objects and the participant based on the current policy π, starting from the initial state s_{init}. Meanwhile, the participant's task was to guess the robot's choice from its gaze behavior and (c) indicate the guess with a corresponding left-hand key press in a self-paced fashion. Upon key press response, the robot stopped action execution and turned its head back at the participant. (d) After a delay of half a second (for avoiding any superposition of event-related brain activity in response to the preceding robot head movement), the robot announced the selected object to the participant by lightning up the corresponding LED attached to its head. This stimulus was delivered for one second, during which the ErrP response was decoded (the duration of a typical ErrP is typically no longer than 600-800 ms [103]) and (e) afterwards, the additional auditory robot speech feedback was provided to the participant. This trial structure applied to CALIB and CORL-I, -II, and -IV. In CORL-III participants were not required to indicate their guesses via key press responses (no explicit decision making), ending the robot's action execution phase, but only to observe the robot's behavior. In CORL-III, the duration of robot action execution was therefore fixed to 15 gaze transitions. This parameter was determined empirically based on the average number of gaze transitions until participant decision during the calibration sessions of a series of pilot experiments. Other than step (b) and (c), no other parts of the trial were affected by the modifications in CORL-III.

5.2.3.2 Robot gaze policy and behavior

Intention / goal selection: The possible selections the robot can choose from defined the agent's set of goals / intentions $G = \{g_{O1}, g_{O2}, g_{O3}\}$. The robot's selection always followed a uniform random choice among the three options.

Gaze policy: The robot's internal gaze policy was realized as a discrete state-space model with four states, $S_\pi = \{s_{objInt}, s_{othObjx}, s_{othObjy}, s_{human}\}$, with s_{objInt}: gazing at selected object; $s_{othObjx}$, $s_{othObjy}$: gazing at one of the other objects; and s_{human}: gazing at human. An action is considered a transition from one gaze state to another or remaining in the current gaze state, leading to 16 possible state-action pairs: $A_{i,j}, i, j = S_\pi$. The policy $\pi(a_{i,j}|s_i) \in [0,1] \subset \mathbb{R}$ determined the gaze behavior described by the probability of taking action $a_{i,j}$ in state s_i (gaze transition from state s_i to s_j). The decision for the next action was always performed by a weighted random selection among the four possible actions in the current state (remain in current state or transit to one of the three other states). This way of realizing

action selection implicitly introduced an exploration-exploitation behavior, an approach used to foster successful policy convergence [210]. In our case, equiprobable distributions among the set of actions to be selected from, drive exploratory behavior and divergent probabilities drive exploitation behavior.

Action (gaze pattern) execution: The robot gaze behavior resulted from a fixed mapping between the covert policy-states S_π and the overt action-execution states S_{act} with corresponding pre-defined robot head angles, pitch Ψ and yaw Θ: $S_{act} = \{s_{O1}, s_{O2}, s_{O3}, s_H\}$, with s_{O1}: gazing at $O1$, $\Psi_{O1} = 25°$, $\Theta_{O1} = -20°$; s_{O2}: gazing at $O2$, $\Psi_{O2} = 25°$, $\Theta_{O2} = 0°$; s_{O3}: gazing at $O3$, $\Psi_{O3} = 25°$, $\Theta_{O3} = 20°$; and s_H: gazing at participant, $\Psi_H = 0°$, $\Theta_H = 0°$. The mapping depended on the selected object and was realized as in Eq. (5.1), (5.2), and (5.3). During action execution, the robot performed one action (state-transition) per 400 ms (2.5 Hz); each gaze shift was executed with a fixed speed of 15% and 10% of the maximum joint speed for pitch- and yaw-movements, respectively. These parameters were manually tuned such that the frequency of gaze shifts was maximized while preserving smooth, non-jerky gazing behavior and approximately conforming to human timing of head posture shifts during conversation [211].

$$S_{\pi \to act}(g_{O1}) = \{s_{objInt \to O1}, s_{othObjx \to O2}, s_{othObjy \to O3}, s_{human \to H}\} \tag{5.1}$$

$$S_{\pi \to act}(g_{O2}) = \{s_{objInt \to O2}, s_{othObjx \to O3}, s_{othObjy \to O1}, s_{human \to H}\} \tag{5.2}$$

$$S_{\pi \to act}(g_{O3}) = \{s_{objInt \to O3}, s_{othObjx \to O1}, s_{othObjy \to O2}, s_{human \to H}\} \tag{5.3}$$

CALIB gaze policy: During the calibration session, the gaze policy was pre-programmed with high probabilities for the following state-action pairs, such that the probabilities of all four possible actions in each state summed up to one: $p(s_{objInt}|s_{human}) = 0.85$, $p(s_{objInt}|s_{othObjx}) = 0.85$, $p(s_{objInt}|s_{othObjy}) = 0.85$, $p(s_{ObjInt}|s_{ObjInt}) = 0.85$ and low probabilities of $p = 0.05$ for all remaining state-action pairs. This resulted in gaze behavior in which the robot tended to fixate the selected object or gaze at it more often.

CORL gaze policy initialization: At the beginning of each co-adaptation run, all state-action pairs of the robot's gaze policy were initialized with $p(s_j|s_i) = 0.25$. This resulted in uniform random gaze behavior which allowed no informed guesses about the selected object (chance-level $p = 1/3$). As the co-adaptation runs proceeded, the gaze policy underwent one update per trial based on the outcome of online decoded ErrPs. The procedures for ErrP decoding and gaze policy update are described in detail in the following sections.

5.2.3.3 Single-trial decoding of error-related potentials

EEG data recording: In all parts of the experiment, EEG and EOG data were acquired with a Brain Products actiChamp amplifier equipped with 32 active EEG electrodes[1] arranged according to the international 10-20 system [128]. All leads were referenced to the average of TP9 and TP10 (average mastoids referencing) and the sampling rate was set to 1024 Hz. The impedance levels of all leads were kept below 10 kΩ. Three channels were used for capturing electrooculogram (EOG1-3) signals in three locations of the participant's face (forehead, left and right outer canthi) according to a method suggested by Schlögl and colleagues [179]. The data was transferred via USB to a separate recording PC (Intel$^{\circledR}$ Core i5 CPU 750@2.67 GHz). Data recording, pre-processing, and ErrP decoding was performed using OpenViBe [212] together with customized processing functions implemented in MATLAB$^{\circledR}$.

Offline modeling of ErrP-decoder: For each participant, an individual ErrP-decoder was trained based on the data collected during the calibration session, and using the following automatic procedure (taking about 5 minutes): The data was first filtered with a causal first-order Butterworth FIR bandpass filter with cutoff frequencies 0.5 and 20 Hz. Then, EOG activity (horizontal and vertical) was reduced by using a regression method proposed by Schlögl and colleagues [179]. The data was re-referenced to CAR and then segmented into data epochs [0...1000] ms relative to the moment of presentation of LED feedback. Segments corresponding to trials in which the participant's guess did not match the robot's feedback were labeled as error-events, without distinguishing between mismatches resulting from a human incorrect guess or a robot false feedback; segments in which the guess matched the feedback were labeled as non-error-events. All data segments were then normalized by subtracting their individual means from each segment time series. Temporal features extracted from the time series were used in this work for its proven higher performance relative to other types of features [102, 171, 105, 42]. The arithmetic means of the signal amplitude in pre-defined windows[2] relative to the moment of feedback presentation was computed, such that all relevant components of the ErrP event-related potential were covered, resulting in a total of 189 temporal features per epoch (27 channels x 7 windows). The features were then used to train a regularized version of the linear discriminant analysis classifier (rLDA) [180] similar to the classification approach used in the previous chapter (see Chapter 4, Section 4.2.4.3). The regularized covariance matrices were here computed by $\tilde{\Sigma}_j = (1-\lambda)\Sigma_j + \lambda I$, with $\lambda \in [0,1] \subset \mathbb{R}$ being the shrinkage parameter and I the identity matrix [182]. The optimal shrinkage parameter was determined using 10-times-10-fold CV based grid search for $\lambda = [0,1]$ in steps of 0.05. To avoid the classifier favoring one class over the other, for each time and fold, the number of trials per class was balanced by random pick and replace (please note that the number of trials per class was initially unbalanced with ~65% non-error and ~35% error trials). The λ with the highest cumulative accuracy of non-error (TNR) and error (TPR) recognition was selected and used to train the final rLDA classifier based on all trials of the calibration

---

[1] EEG channel labels: FP1, FP2, F3, F4, F7, F8, FC1, FC2, FC5, FC6, C3, C4, T7, T8, CP5, CP6, P3, P4, P7, P8, TP9, TP10, O1, O2, Fz, Cz, Pz, EOG1, EOG2, EOG3.
[2] Windows: [150...250] ms, [200...300] ms, [250...350] ms, [300...400] ms, [350...450] ms, [400...500] ms, [450...550] ms.

data. Also, in this final step, the numbers of trials per class were balanced by random pick and replace. To increase the likelihood that most of the calibration trials were used at least once, this procedure was repeated 1000 times and the weights w and bias b of individual models were averaged to obtain a single final rLDA classification model. The chance-level threshold for this binary classification problem is 60.37% for both TNR and TPR, given the manually balanced number of trials for decoder calibration[3]. Participants for which either TNR or TPR did not exceed the above threshold were excluded from all further analyses.

Online ErrP decoding: The ErrP-decoder trained with the calibration data was then used to decode ErrPs from the participant's ongoing EEG during the co-adaptation runs. The signals were continuously bandpass filtered using a causal first-order Butterworth FIR bandpass filter with cutoff frequencies 0.5 and 20 Hz (identical filter parameters as used during offline modeling). EOG activity was continuously reduced by applying the EOG decorrelation matrix obtained from the calibration data. Finally, the continuous signals were re-referenced to CAR. Upon occurrence of a feedback event (robot communicating decision via flashing one out of three LEDs), the respective data segment ([0...1000] ms time-locked to the event) was processed in the same way as in offline modeling: (1) single-trial normalization, (2) temporal features extraction, and (3) classification into non-error or error events using the rLDA classifier trained on the calibration data.

5.2.3.4 ErrP-based agent policy adaptation

For the ErrP-based policy adaptation we decided to employ a learning paradigm based on policy gradient methods, a subform of RL (reinforcement learning) [210]. Among others, the main advantages of policy gradient methods compared to other RL methods, such as Q-learning are as follows: first, policy gradient methods act on-policy directly which facilitates the interpretation of policy adaptations in contrast to value function-based approaches. Second, policy gradient learning has been proposed as the method of choice in RL for humanoid robots as they can deal with complex learning tasks involving many degrees of freedom [213, 214]. Although not in the focus of this work, this property favors the generalizability and scalablity of our approach to more complex robot behavior and interaction scenarios. The policy update function is given in Eq. (5.4) which was executed at the end of each trial during the co-adaptation runs, starting with the initial policy π_{init} with equal probabilities for all state-action pairs $p(s_j|s_i) = 0.25$. In short, for the computation of the parameters of the new policy π^{t+1},

---

[3] inverse cumulative binomial distribution with number of trials $nTrials = 150 * p_{err}$ and $p_{err} = 0.35$; probability of success $p_{success} = 0.5$; confidence threshold $p = 0.05$

to be employed in the next trial, the parameters of the old policy π^t were merged in a weighted fashion with the empirical distribution of the observed state-action pairs during the current trial:

$$\pi^{t+1}(a_j|s_i) = \pi^t(a_j|s_i) + \alpha \tilde{R} \sum_{t=1}^{n} a_{i,j}^k \tag{5.4}$$

with t being the count of the current trial, $\tilde{R}$ being the reward estimate derived from the ErrP-decoder class decision, with negative reward $\tilde{R} = -1$ for a classified error event and positive reward $\tilde{R} = +1$ for a classified non-error event, α being the learning rate, and $\sum_{t=1}^{n} a_{i,j}^k$ the occurrence count of action $a_{i,j}$ (transition from state s_i to s_j) of the action sequence $k = (1, \ldots, n)$ executed by the robot in the current trial, with n depending on the participant's self-paced decision. Truncation and normalization were performed after adding the policy gradient $\alpha \tilde{R} \sum_{t=1}^{n} a_{i,j}^k$ to the parameters of the old policy π^t: parameter updates of π^{t+1} which exceeded the range $\{0, 1\} \in \mathbb{R}$ were truncated to 0 and 1, respectively, and all actions per state were then normalized to sum up to one. Based on pilot experiments, the learning rate was empirically set to $\alpha = 0.1$ such that convergence could be reached relatively fast within a few policy updates. Fast convergence was preferred given the limited number of 50 trials per co-adaptation run. The rationale behind including the empirical distribution of observed state-action pairs into the policy gradient was based on the assumption that more prominent state-action pairs are likely to contribute more to the participant's false or correct guess compared to less prominent state-action pairs. State-action pairs which occurred and hence were observed more often than others were as such more strongly reinforced (increase or decrease of corresponding state-action probability depending on $\tilde{R}$) compared to state-action pairs which occurred less prominently or never during the trial. This way, the policy was updated to promote correct guessing or in other words, to fit to the participant's belief by quantitatively taking into account the characteristic of the gazing behavior the participant has observed.

5.3. Data analysis and results

Three out of 16 participants did not meet the inclusion criterion defined in Section 5.2.3.3: Offline decoder performance did not exceed the chance-level of 60.37% in either TNR, TPR or both in participants s05, s10, and s13 (see Supplementary Table 18, Appendix A). The data of these participants were therefore excluded from subsequent analyses. For the sake of full disclosure of the obtained data, individual results of excluded participants are nevertheless reported and discussed separately.

5.3.1. Offline and online ErrP decoding performance

Figure 24(a) shows the grand average ERP time-courses recorded at the Cz, time-locked to the onset of LED-feedback presentation by the robot. Also illustrated is their topographical distribution at specific time-points. The grand average difference (black line in Figure 24(a)) showed the typical N2-P3-complex which has been reported consistently in the context of ErrPs [102, 103, 104, 105]. The negative deflection (N2-component, expected around 180-325 ms [140]) was mostly pronounced fronto-centrally around 300 ms post stimulus and the positive deflection (P3-component, expected around 250-500 ms [215]) was mostly pronounced fronto-centrally around 400ms. The coefficient of determination based on channel Cz reached highest values of $r^2 = 0.09$ for 288 ms and $r^2 = 0.11$ for 394 ms averaged across all participants ($n = 13$) which speaks in favor for a good overall separability of the data. Figure 24(b) shows a comparison of the grand average difference ERPs over Cz across calibration session (CALIB) and co-adaptation runs I, II, IV (CORL)[4] with high temporal resemblance between the experimental conditions. These observations were reflected in the single-trial classification performances (see Figure 24(c) and Supplementary Table 18, Appendix A). The average offline ErrP-decoder performance based on calibration data CV was $Acc = 80.2 \pm 7.5\%$, with $TNR = 81.2 \pm 7.7\%$ and $TPR = 79.2 \pm 7.5\%$ (AVG±SD). ErrP online decoding performances were comparably high in accuracy (see Figure 24(c) and Supplementary Table 18, Appendix A) with $Acc = 84.2 \pm 7.4\%$ for CORL-I, $Acc = 77.1 \pm 12.1\%$ for CORL-II, and $Acc = 84.0 \pm 10.6\%$ for CORL-IV. The main difference observed in comparison to the offline CV results was a higher decoding performance for non-error events ($TNR = 86.5 \pm 11.7\%$, $TNR = 82.5 \pm 18.3\%$, $TNR = 90.4 \pm 5.8\%$ for CORL-I, II, IV, respectively) and a lower performance for error events ($TPR = 75.3 \pm 12.0\%$, $TPR = 70.4 \pm 15.4\%$, $TPR = 74.4 \pm 17.0\%$ for CORL-I, II, IV, respectively). This performance bias was significant across participants for all co-adaptation runs ($p = 0.026$, $p = 0.023$, $p = 0.002$, for CORL-I, II, IV, respectively; paired Wilcoxon signed rank test, $n = 13$). Online decoding accuracies were on average lower in CORL-II compared to CORL-I and CORL-IV. This was consistent across participants as decoding accuracies differed significantly between CORL-I and CORL-II, and between CORL-II and CORL-IV; no statistically significant difference was found between CORL-I and CORL-IV ($p_{I-II} = 0.031$, $p_{II-IV} = 0.046$, $p_{I-IV} = 0.600$; paired Wilcoxon signed rank test, $n = 13$). Possible explanations are discussed in Section 5.4. The theoretical chance-level for online ErrP decoding per CORL is 62.0%[5] which was exceeded in all but three cases: s06/CORL-IV, s11/CORL-II, s14/CORL-II (see Supplementary Table 18, Appendix A).

Table 6 shows the overview of individual participant results. The separability of the ErrPs are expressed in form of the maximum coefficient of determination across all channels r^2_{max}, separately for CALIB and the CORL data (all trials of CORL-I,-II, and –IV). The results show comparably low values of $r^2_{max} \approx 0.11$ for the three participants in which offline decoding

---

[4] Please note that no results are reported for CORL-III, since no participant key press responses (validation ground truth) were captured during this part of the experiment.

[5] inverse cumulative binomial distribution with number of trials $nTrials = 50$; probability of success $p_{success} = 0.5$; confidence threshold $p = 0.05$.

Table 6 Overview of individual results per participant in the order of columns from left to right: maximum coefficient of determination r^2 across all channels within period 150-550 ms for CALIB and CORL. Offline cross-validation ErrP-decoder accuracies based on CALIB data. Average online ErrP-decoder accuracies during co-adaptation runs CORL-I,-II, and –IV. Within-participant Pearson's spatiotemporal correlation coefficients between average difference ERP time courses of all channels (error minus non-error) of CALIB and CORL (average of all trials of I, II, IV) within period 150-550 ms.

	r^2_{max} CALIB	$\overline{r^2_{max}}$ CORL	Acc CALIB (offline CV)	$\overline{Acc}$ CORL (online acc.)	$corr2$ (CALIB,CORL)
s03	0.20	0.31	69.3%	86.7%	0.67
s04	0.34	0.41	86.3%	91.3%	0.72
s06	0.36	0.20	82.0%	72.7%	0.74
s07	0.25	0.31	81.4%	82.7%	0.58
s08	0.43	0.48	85.7%	87.3%	0.87
s09	0.46	0.29	92.8%	86.7%	0.82
s11	0.17	0.09	68.9%	65.3%	0.29
s12	0.30	0.23	84.7%	90.0%	0.41
s14	0.39	0.27	88.7%	77.3%	0.72
s15	0.18	0.31	72.4%	85.3%	0.70
s16	0.17	0.31	73.3%	74.7%	0.49
s17	0.36	0.54	81.8%	88.0%	0.77
s18	0.39	0.24	78.9%	74.7%	0.53
AVG±SD	0.31±0.10	0.31±0.12	80.5±7.5%	81.8±8.0%	0.64±0.17
s05	0.12	0.12	50.5%	50.7%	-0.02
s10	0.10	0.19	64.4%	28.7%	0.13
s13	0.11	0.13	50.2%	67.3%	0.12

performance did not exceed the chance-level threshold (s05, s10, s13), whereas all other participants show higher r^2_{max} values. This indicates, that calibration failed in these participants, mainly due to their generally limited separability of ErrP responses (possible explanations are discussed in Section 5.4). As expected, the overall ErrP-decoder offline CV accuracies (Acc CALIB) and the online average decoding accuracies ($\overline{Acc}$ CORL) reflected the results obtained from the analysis of the coefficient of determination, with high separability resulting in higher decoding accuracies. Table 6 furthermore reports Pearson's spatiotemporal correlation coefficients between the average difference of ERP time courses of all channels (error minus non-error) within the period 150-550 ms (period in which the temporal features were ex-

tracted). The overall high correlation coefficients of average $corr2 = 0.64 \pm 0.17$ (AVG±SD) reflect high spatiotemporal resemblance and support the notion that the decoded ErrPs did not notably differ between CALIB and CORL experimental sessions, despite the different experimental conditions.

5.3.2. Results of ErrP-based co-adaptation

To investigate the extent of co-adaptation between participant and robot, we analyzed the development of two behavioral measures in conjunction with the development of the policies during the co-adaptation runs: (1) Guessing performance - the development of the accuracy of correct guesses. This measure was expected to increase if both participant and robot converge to a consensus. (2) Gaze transitions until participant's decision - number of gaze transitions performed by the robot until the participant made a decision. This measure was expected to decrease as participants and robot converge to a consensus. (3) Policy convergence – the policy change of trial-by-trial updates. This measure was expected to decrease if policies converge.

Efficacy - guessing performance: The participant has three objects to choose from and therefore chance-level was $p = 1/3$. At the beginning of each co-adaptation run the robot's gaze policy was initialized with equal probabilities for all actions. This guaranteed a random guess in the first trial of all co-adaptation runs. Hence, if during the co-adaption runs, the participant's guessing performance exceeded chance-level, the robot's gaze policy must have been updated such that it promoted the participants to guess correctly (e.g. the underlying object was more readily inferrable by participants from the gaze pattern generated by this policy); vice-versa, if guessing performance did not increase above chance-level during the run, then updates in the robot's gaze policy did not facilitate the participant's task and/or were misleading. To investigate whether the guessing performance depended on the ErrP-decoder performance during online operation, we computed Pearson's correlation coefficients between the overall guessing performance (percentage of correct guesses within one run) and the ErrP-decoder accuracy (percentage of correctly classified trials) across all participants for each co-adaptation run separately. Overall guessing performance correlated positively with the online decoding accuracies in all three co-adaptation runs with $r = 0.71$ ($p = 0.006$) for CORL-I, $r = 0.79$ ($p = 0.001$) for CORL-II, and $r = 0.47$ ($p = 0.1$) for CORL-IV (Pearson's correlation, $n = 13$). These results indicate that improvements in guessing performance depended on the ErrP-decoder performance during online operation, e.g. high ErrP-decoder performance fostering high guessing performance. As a result, those participants for which the ErrP-decoder calibration performance resulted in below chance-level accuracies (s05, s10, s13), also showed no notable improvements in guessing performance (see Supplementary Tables 18 and 19, Appendix A).
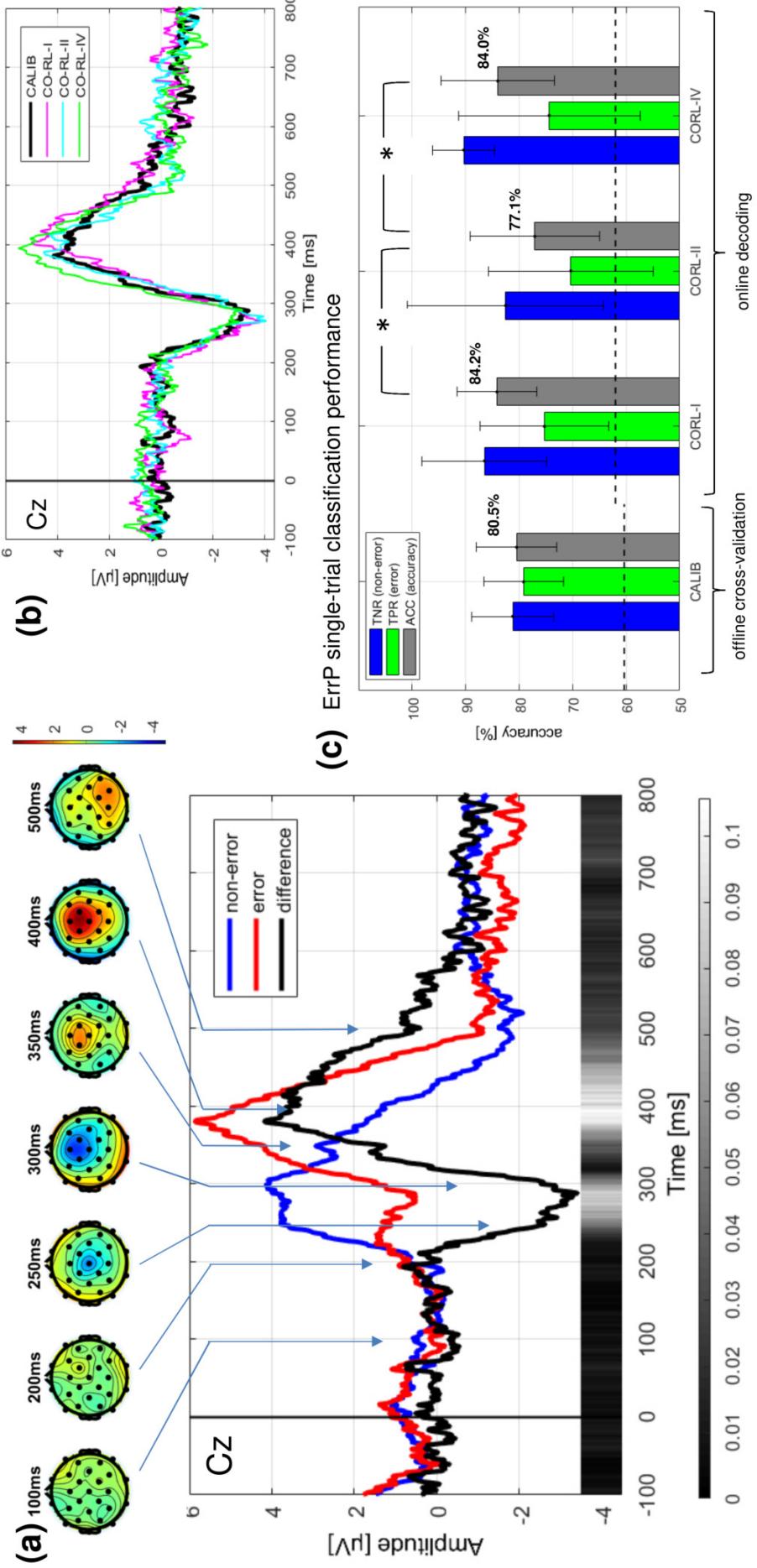
Figure 24 (a) Grand average ($n = 13$) ERP time-courses over channel Cz time-locked to the onset of feedback presentation for each class of events (blue: non-error, red: error) and the difference grand average (black: error minus non-error). The r^2-values for between non-error and machine-error are depicted below the plot, with brighter colors indicating higher values. The difference grand average is furthermore depicted as topographic plots for the main peaks above each plot and in form of a spatio-temporal activity matrix across all channels and specific time points above the plot (Figure style adopted from [141]). (b) Comparison of the difference grand average across the calibration session and the co-adaptation runs I, II, and IV. The comparison shows high resemblance of the difference ERPs across experimental sessions. (c) ErrP single trial classification performance for offline CV based on the calibration data and online decoding performance during the co-adaptation runs. The black dashed lines indicate the theoretical chance-level of 60.37% for CALIB and 62.0% for individual CORLs. For CALIB, chance-level was exceeded in all participants; for CORL, chance-level was exceeded in all but three cases (s11/CORL-II, s14/CORL-II, s06/CORL-IV). Online decoding accuracies were significantly different between CORL-I and CORL-II and between CORL-II and CORL-IV, indicated by the asterisks.

78

To investigate improvements of guessing performance over the course of co-adaptation runs, each run was partitioned into five segments of 10 trials each. Guessing performance was computed as percentage of correct guesses in each segment[6]. Figure 25(a) shows across participant distributions of guessing performance from the start (Trial: 1-10) until the end (Trial: 41-50) of each co-adaptation run. The results show an increase in median guessing performance from initial chance-level up to 90% in CORL-I, 70% in CORL-II, and 80% in CORL-IV. In all three runs, the majority of participants exceeded the threshold of the confidence interval (70%) at some point during the run: In CORL-I and CORL-II, in the fourth segment, and in CORL-IV already in the second segment. Despite the significant differences in ErrP decoding performance (see 5.3.1), no significant differences of overall guessing performance were observed between CORLs ($p_{I-II} = 0.528$, $p_{II-IV} = 0.250$, $p_{I-IV} = 0.104$; paired Wilcoxon signed rank test, $n = 13$). Assuming a co-adaptation run to be "successful" when guessing performance $\geq 70\%$ in three subsequent segments[7], then successful co-adaptation was achieved in 10 out of 13 participants in at least one out of the three runs. Two participants achieved 3/3 successful runs (s09, s12); two participants achieved 2/3 successful runs (s03, s15), six participants achieved 1/3 successful runs (s06, s07, s08, s14, s16, s18). An exemplary successful co-adaptation run (s09/CORL-I) is visualized in Figure 26. Individual results are detailed in Supplementary Table 19, Appendix A.

Efficiency - gaze transitions until participant decision: The absolute number of gaze transitions turned out to vary widely among participants, even during the calibration session, ranging between 10 to 50 transitions (corresponding to a duration of robot action execution between ~4-20 seconds per trial) with average 13.0±6.3 (CALIB), 15.7±7.0 (CORL-I), 15.0±6.9 (CORL-II), and 14.8±10.7 (CORL-IV). Therefore, the number of gaze transitions until participant decision was analyzed by partitioning each co-adaptation run into five segments of 10 trials each (following the same procedure as in the analysis of guessing performance) and counting the number of gaze transitions within each of these 10-trial segments relative to the number of transitions occurring during the first 10-trial segment. The results are depicted in Figure 25(b): in all three runs, the median number of gaze transitions decreased by 15-27% relative to the first segment with across participant significant decreases in some segments ($p < 0.05$, one-sample Wilcoxon signed rank test). In CORL-I the number of gaze transitions decreased by 27.6% (median of percent reduction calculated across participants); in CORL-II by 19.2% and in CORL-IV by 15.6%. This result illustrates that during the co-adaptation runs, participants not only became more precise in guessing, but also on average faster in deciding about the robot's selected object. This suggests that the robot's gaze behavior adapted in a way that was generally easier and quicker to understand by the participants. The absolute number of gaze transitions until participant decision is given in Supplementary Table 20, Appendix A.

---

[6] The 5% confidence threshold is exceeded if ≥ 7 out of 10 trials were correct, one-sided binomial test with chance level $p = 1/3$

[7] probability for exceeding by chance: $p \leq 7.6 * 10^{-6}$, one-sided binomial test with chance level $p = 1/3$
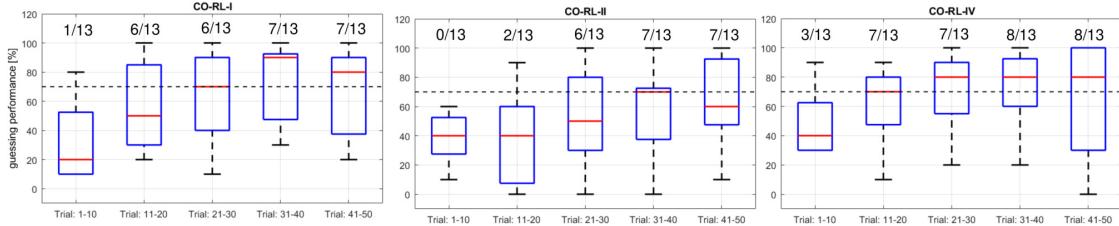
Policy convergence: To quantify policy convergence, the difference between subsequent policy iterations was computed for each participant and CORL individually. This was carried out by determining the value of the state-action pair with the maximum difference between subsequent policy iterations, termed as the policy change after trial t according to Eq. (5.5). Following the same procedure as in the analyses of behavioral measures, policy convergence was analyzed by partitioning each co-adaptation run into five segments of 10 trials each and averaging $\Delta\pi$ within each of these 10-trial segments. The results are depicted in Figure 25(c): in all three runs, the median of $\Delta\pi$ across participants decreased steadily from the first until the last segment from an initial median of $\Delta\pi \approx 0.2$ to a final median of $\Delta\pi \approx 0.1$. This indicates that policies were on average converging relative to an increasing guessing performance. The results further indicate that not all policies converged within 50 trials as the median policy change was still $\Delta\pi \approx 0.1$ in the last segment of all three runs. These findings are further discussed in Section 5.4. Policy convergence for an exemplary successful co-adaptation run (s09/CORL-I) is visualized in Figure 26.

$$\Delta\pi^t = \max |\pi^{t-1} - \pi^t| \tag{5.5}$$
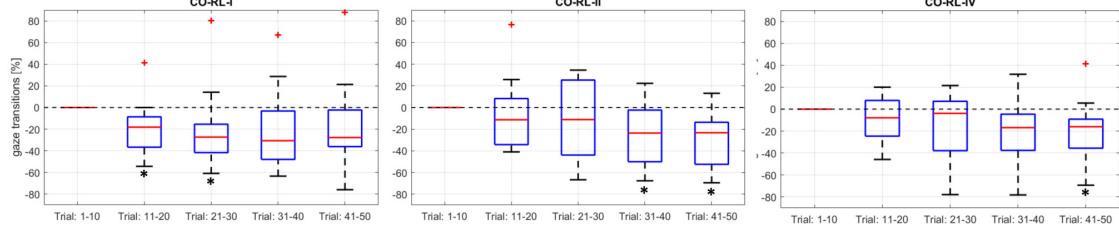
5.3.3. Emergence of gaze behavior

In addition to assessing the development of co-adaptation, we were also interested in the nature of the gaze behavior which emerged during the co-adaptation runs. This analysis allowed for a qualitative assessment of CORL-III in which participants were not requested to explicitly indicate their guesses via key press responses in comparison with the other co-adaptation runs. Since participants were not instructed to follow a particular strategy/policy, any gaze behavior was considered acceptable and denoted as useful if it helped the participant to perform better and faster in guessing the robot's selected object. Figure 27(a) shows an overview of the learned policies for the 13 participants and all co-adaptation runs, including CORL-III. Successful co-adaption was expected to be reflected in policy convergence towards the end of the run. Therefore, the average of the policies of the last 10 trials (trial: 41-50) were depicted, with thick red lines representing high probabilities and thin blue lines low probabilities. The policies which emerged from successful co-adaptation runs are highlighted with a blue frame; the guessing performances of the corresponding last 10 trials are furthermore depicted next to the average policy. By qualitative visual inspection, we identified two different recurring policies which are furthermore termed "fixation" and "nodding" behavior (see Figure 27(b)). The "fixation" policy led to gaze behavior in which the robot tended to fixate the selected object. Example cases are s03/CORL-I, s07/CORL-II, s15/CORL-I. In the "nodding" policy the robot was gazing alternatingly between the participant and the selected object in a nodding-type fashion. Examples are s09/CORL-I, s12/CORL-II, s18/CORL-IV. Also, in CORL-III, the robot's gaze behavior converged in a few cases to one of the two identified policies, e.g. "fixation" behavior in s08 and s18, and "nodding" behavior in s16 (Figure 27(a)). These cases indicate that participants explicitly indicating their decision (as in CORL-I, -II, -IV) was not required for successful co-adaptation, and suggests that the ErrP-based method presented here also worked based

(a) Efficacy: guessing performance

(b) Efficiency: gaze transitions until subject decision

(c) Policy convergence
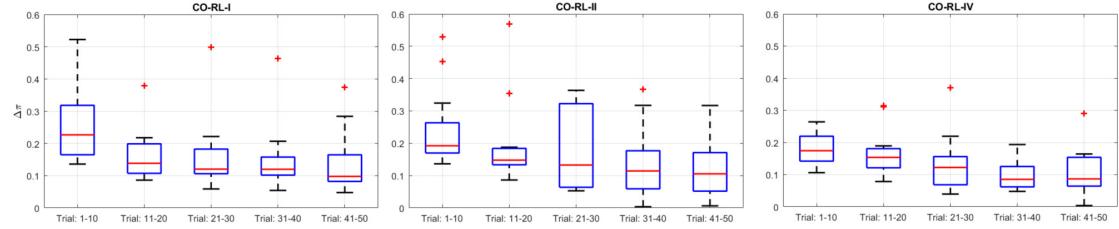
Figure 25 (a) Boxplot representation of guessing performance computed for consecutive segments of 10-trials, across participants ($n = 13$). Different panels correspond to different co-adaptation runs. In all three runs, the median guessing performance increased to 70-90%. The black dashed line indicates the threshold of the confidence interval of 70%. The numbers on top of the boxplots represent the ratio of participants which exceeded the threshold of the confidence interval in the corresponding segment. (b) Boxplot representation of gaze transitions until participant decision computed within consecutive 10-trial segments, relative to the number of transitions counted in the first segment (black dashed line). In all three co-adaptation runs, the median number of transitions decreased by 15-27% with significant across participant deviations in some segments (boxes marked with a black asterisk). (c) Boxplot representation of policy change computed within consecutive 10-trial segments. In all three co-adaptation runs, the median policy change decreased from $\Delta \pi \approx 0.2$ in the first segment to $\Delta \pi \approx 0.1$ in the last segment, indicating policy convergence relative to increasing guessing performance.

on covert beliefs/decisions without explicit actions linked to the decisions. Convergence to the "fixation" behavior was expected, since it is very similar to the pre-programmed policy used during the calibration task; participants likely could have used it as a proxy. The "nodding" behavior was however unexpected, since it had not occurred before during the calibration session and participants could therefore not use it as a proxy. Interestingly, the number of cases associated with convergence to the "nodding" policy (7) were approximately on par with those with convergence to the "fixation" policy (8). The "nodding" behavior may have emerged from participants gradually finding it useful and in result having adapted to and positively reinforced it. This observation retrospectively confirms co-adaptation between participant and robot, since if participants were instructed to teach the robot a specific behavior, different, previously unexpected behavior was less likely to emerge.
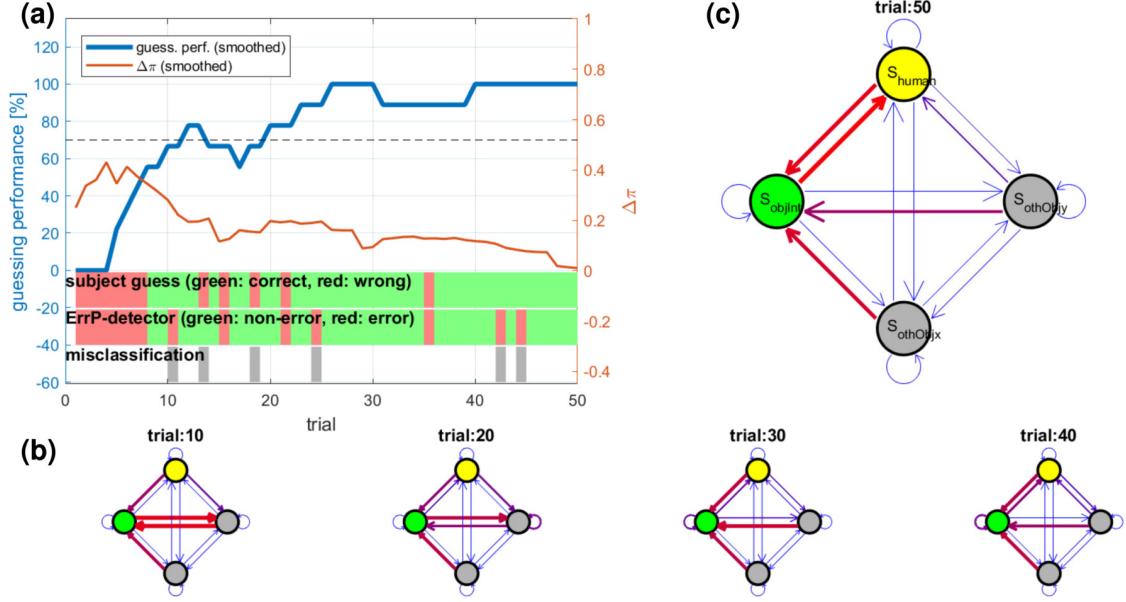
Figure 26 (a) The plots represent a smoothed representation (over 10 trials) of the development of guessing performance (blue line) together with the rate change of policy updates $\Delta\pi$ (orange line). The black dashed line represents the threshold of confidence (70%) for guessing performance. Single-trial participant guesses, corresponding ErrP-decoder classification decisions and misclassifications are illustrated below the plots. (b) Shows the gaze behavior policies after different numbers of iterations (after trial: 10, 20, 30, 40). (c) shows the final policy at the end of the co-adaptation run. The states of the gazing policy are color-coded as follows: s_{human} (yellow), s_{objInt} (green), $s_{othObjx}$, $s_{othObjy}$ (grey). State-transitions with high probabilities are represented with thick red lines, low probabilities with thin blue lines. This particular example shows convergence towards the "fixation" behavior around trial 30 and "nodding" behavior towards the end of the co-adaptation run (see Section 5.3.3).

5.4. Discussion of closed-loop BCI setup and experimental results on ErrP-based co-adaptation

ErrPs are a useful feedback signal for successful mediation of co-adaptation in HRI:

ErrPs, decoded from human participants' brain activity in real-time during HRI, might be useful in the future to adapt the behavior of artificial agents, such as robots, to better align with human expectations, needs and conventions. We understand our study as a logical extension of previous works [105, 173, 43] which demonstrated the potential of using ErrPs as a teaching signal for robot skill learning. Instead, our experimental paradigm featured a scenario in which there was no explicit "optimal" or "correct" behavior the robot was required to adapt to, but where mutual adaptation between human and robot was permitted; the "optimal" robot policy had to be negotiated between both parties in a co-adaptive fashion. This introduced a considerable level of uncertainty and complexity into the experimental setup as participants could not follow a specific task or proxy. With this relaxation of constraints in the experimental setup, we aimed at validating the usability of ErrPs as an implicit feedback signal to improve HRI where adaptation is possible from both interaction partners. Despite the uncertainty and complexity introduced, we observed significant improvements in interaction performance across participants over the course of individual co-adaptation runs, as indicated by behavioral measures of efficacy and efficiency: The average percentage of correct guesses (efficacy) increased
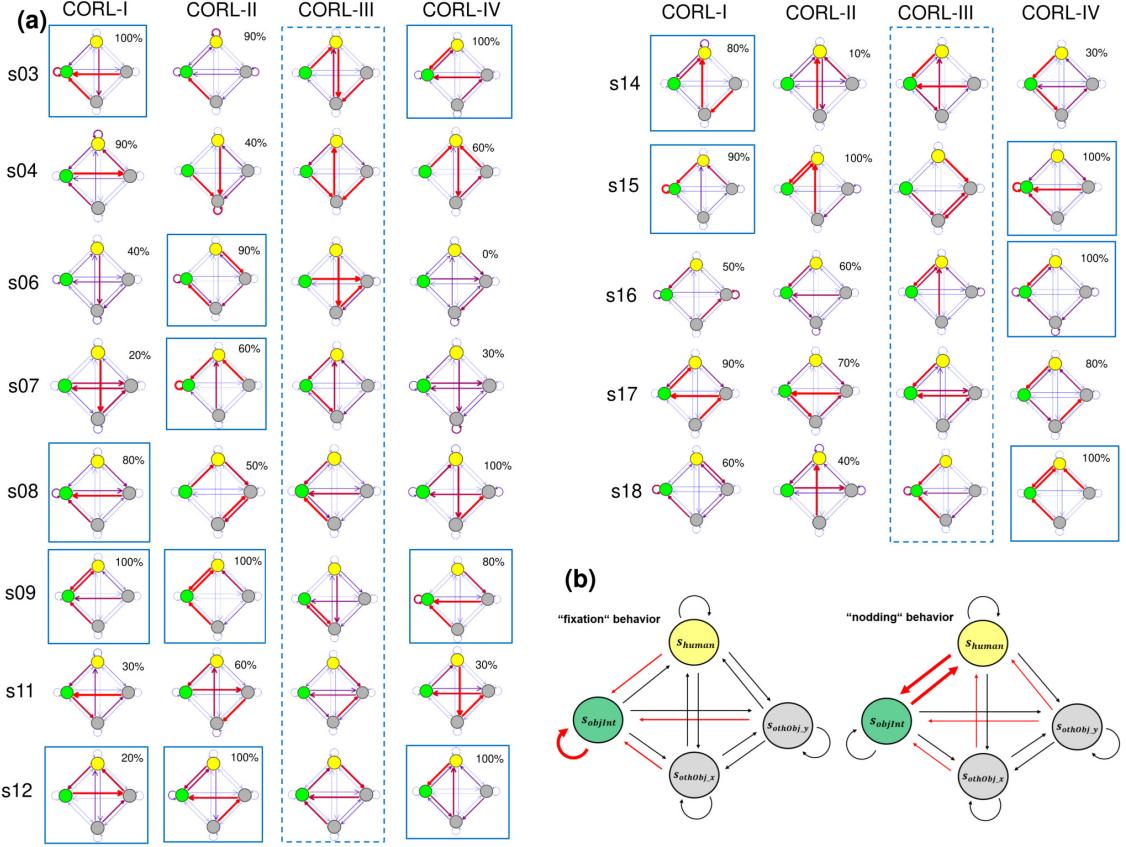
Figure 27 (a) Overview of gaze policies averaged across the last 10 trials for 13 participants and all co-adaptation runs. Successful co-adaptation runs are highlighted with blue frames. The guessing performance during the corresponding last 10 trials is depicted next to the policies. (b) – Identified recurring policies: "fixation" behavior in which the robot tended to fixate the selected object (examples: s03/CORL-I, s15/CORL-I, s08/CORL-III, s18/CORL-III); "nodding" behavior in which the robot alternatingly gazed at the participant and the selected object in a nodding-type fashion (examples: s09/CORL-II, s16/CORL-III, s18/CORL-IV).

from the initial chance-level (~33%) to 70-90% within 10-40 trials (corresponding to 1-4 minutes), median across participants. Additionally, the number of gaze transitions made by the robot before the participant indicating his/her guess (efficiency), relative to the corresponding number in the beginning of the co-adaptation run, decreased on average by 15-27% across the different runs. Hence, adaption of the robot's policy, based on the ErrPs collected from the human interaction partner, was accompanied by a higher performant and more efficient interaction. Comparison of learned gaze policies between CORL-III (no explicit human decision) and the other conditions CORL-I,-II,-IV (with explicit human decision) suggested furthermore that the explicit action (key press response used as validation ground truth) linked to the participant's decision was not a pre-requisite for successful co-adaptation, but can be achieved solely based on human's covert belief confirmation/violation.

Online single-trial ErrP decoding performance is comparable to previous works:

Online single-trial ErrP decoding performance was on average $Acc = 81.8 \pm 8.0\%$ (AVG$\pm$SD) across 13 participants which is comparable to previously reported ErrP classification performances used for closed-loop adaptation of robotic systems: Iturrate and colleagues (2015) ob-

tained online decoding accuracies around 74% across 12 participants using temporal features combined with LDA classification [105]. Salazar-Gomez and colleagues (2017) obtained online decoding accuracies around 65% across 4 participants using correlation-and covariance-based features after spatial filtering using the xDAWN algorithm [216]. Based on post-hoc offline analyses, they reported however the presence of a secondary ErrP for which they estimated a theoretical online decoding performance around 80% [173]. The most recent work by Kim and colleagues (2017) reported high online decoding performances of balanced accuracy around 90% across 7 participants using temporal features after xDAWN spatial filtering and classification using linear SVM. They explain their high performance being mainly a result of their data augmentation approach based on decoding ErrPs in two separate time windows (instead of just one) [43]. A limitation observed from the online single-trial classification results of the present study (Section 5.3.1) is a consistent and significant across participants bias towards the non-error class. This classification bias has been reported consistently in the context of ErrP decoding and was related to the design of calibration protocols with typically unbalanced number of samples per class [103]. As class-balancing was performed in the present study, the systematic bias was likely related to the limited number of 150 samples used for ErrP-decoder calibration. Although the use of a regularized LDA may have partially counteracted this [182], the use of a priori information from other participants [160] or upsampling the minority class, instead of downsampling the majority class may have helped improve online decoding performance and are recommended modifications for future works.

ErrP-decoder calibration resulted in chance-level performance in three participants (s05, s10, s13). These participants showed lower r^2_{max} values compared to the remaining participants (see Table 6), which points to a worse separability of their ErrP responses. A post-hoc visual inspection of the raw EEG data showed comparably strong artifact contamination in s10 (mainly noisy channels) and in s13 (mainly slow DC drifts); the data of s05 on the other hand was largely unaffected by artifacts. This suggests that the low calibration performance in s10 and s13 was mainly due to the technical setup and could have been resolved by repeating the experiment on a different day or by adding automatic artifact rejection to the modeling procedure. Why calibration failed in s05 is currently unexplained; one possibility could be that this participant had been insufficiently concentrated on or engaged in the task.

Across participants, ErrP online decoding performance was significantly lower in CORL-II than in CORL-I and CORL-IV. This temporary drop in performance in CORL-II is unlikely related to technical reasons, but rather to the participants' concentration / task engagement level, because it recovered toward the end of the experiment in CORL-IV. CORL-II was a direct repetition of CORL-I, which might have had a negative effect on the participant's motivation and engagement. On the other hand, CORL-IV followed CORL-III; the intermediate variation of the experimental protocol with CORL-III might have had a positive effect on the participant's engagement during CORL-IV. Despite noticeable differences in the median guessing performance (Figure 25(a)), the systematic drop in ErrP online decoding performance had no significant effect on the guessing performance in CORL-II.

ErrP-decoder performance plays an integral role in the success of co-adaptation:

The ErrP-decoder performance played an integral role in successful co-adaptation, as indicated by positive correlations between overall guessing performance and ErrP-decoder performance during online operation. Furthermore, on average, policies converged in relation to increasing guessing performance (see Figure 25(a) and 25(c)) as indicated by a median decrease of the policy change rate over the course of co-adaptation runs. This supports the functionality of the policy adaptation approach here adopted. We observed, however, several cases in which co-adaptation failed despite high ErrP-decoder performance ($Acc > 75\%$). These cases might be due to unknown human-related factors, such as variations in task engagement, attention variations, or variations in interpretation of the experiment. From a technical perspective, the policy adaptation approach used may have influenced the stability of co-adaptation as well. In some of these failed cases there were temporary increases in guessing performance followed by decreases (unlearning), indicating temporary, but unstable learning (exemplary cases are s14/CORL-IV, s17/CORL-II, see Supplementary Figures 40 and 41, Appendix A). The learning approach here adopted does not enforce convergence to an optimal policy. This has the advantage that while converging to one policy, bifurcations to other policies remain possible. This flexibility might be particularly important in the context of co-adaptation as changes of the human strategy are likely and imaginable in the sense that a policy which was previously preferred by the participant is neglected and replaced by a different preferred policy. Exemplary cases supposedly showing such policy re-adaptations are s08/CORL-IV, s09/CORL-II, and s14/CORL-I (see Supplementary Figures 42, 43, and 44, Appendix A). These cases show initial convergence interrupted by periods of increased policy changes and subsequent secondary convergence. On the other hand, this flexibility, in combination with a learning rate parametrized to promote quick learning, may have encouraged instabilities or quick unlearning, as the outcome of single trials interfered with the learning process (e.g. sensibility to ErrP-decoder misclassifications). One possible way of stabilizing the policy adaptions would be to use an adaptive learning rate based on the past rewards (ErrP-decoder decisions), e.g. decreasing the learning rate in case of increasing number of past non-error events. However, whether, and to what extent a systematic control of the learning process is recommendable in the context of ErrP-based human-agent co-adaptation remains an open question for future investigations. In our experiment the learning process was limited to 50 iterations (trials), which turned out insufficient for drawing definite conclusions about the co-adaptation process in the long run. Therefore, an entry point for follow-up studies is most importantly the investigation of the dynamical effects of co-adaptation for longer periods or during continuing interaction. Several of above considerations have been addressed in follow-up research in form of model-based simulations of the integrated human-agent co-adaptive system (see Chapter 6).

Emergence of different gaze patterns qualitatively support successful co-adaptation:

The analysis of emergence of gaze policies revealed that in most successful co-adaptation runs the robot's gaze behavior converged to either one of two different policies: "fixation" and "nodding" behavior. While the "fixation" behavior was expected as it closely resembles the gaze behavior during the calibration session, the "nodding" behavior, in contrast, was not expected. Although both behaviors are qualitatively different and may be interpreted as conveying different meanings, an alternative interpretation is that both are consistent in that the target object is attended to more often than others. In that sense, it is likely that the type of behavior to which the system converged to depended on whether the alternating transitions between s_{human} and s_{objInt} or the transition $s_{objInt \rightarrow objInt}$ were sampled more often in an early stage of the co-adaptation run. Either way of interpreting, the emergence of the two types of behaviors supports the hypothesis of co-adaptation, since for the participants both strategies seemed valid despite having had no explicit exposure to the "nodding" behavior before the start of the co-adaptation runs. On that note, one may argue about why just two different behaviors emerged from the interaction, given that the manifold of imaginable and possibly valid strategies is much bigger (e.g. a slightly more complicated gaze pattern or a consistent logical swap of the target object with one of the other objects). This observation may be related to constraints in human information processing and learning of more complex statistical patterns but remains an open question for future investigations. Further exploration of how robot behavioral policies, as in this case the robot's gazing policy, develop during such interaction will provide useful insights for improving the technical implementation of ErrP-based mediation of human-robot co-adaptation and may likewise provide insights about human information processing and learning.

5.5. Summary

In this chapter, we experimentally demonstrated the usability of EEG-based ErrPs as a feedback signal for mediating co-adaptation in HRI. Our study featured a simplified HRI scenario in which successful interaction depended on co-adaptive convergence to a consensus between participant and robot. ErrPs were decoded online from participants' ongoing EEG signals with an average accuracy of $Acc = 81.8 \pm 8.0\%$ and utilized for adaptations of the robot behavior, while the participant adapted to the robot by reflecting upon its behavior. Adaptations of the robot behavior were here implemented with an episode update strategy using ErrPs as a delayed reward feedback signal for the past sequence of robot actions, instead of performing adaptations on single occurrences of robot actions. This update strategy allows for adaptation of more complex robot behavior, where individual robot actions are more likely to occur in rapid succession and not to be temporally well isolated. Successful co-adaptation was demonstrated in the majority of participants (10 out of 13) by significant improvements in interaction efficacy and efficiency (more precise and faster interaction) and by plausible robot behavioral policies that emerged during the interaction.

6. Computational modelling of ErrP-based human-agent co-adaptation

Both recent works by others and our previous study have demonstrated the potential of using ErrPs online decoded from a human interaction partner for robot skill learning as well as the feasibility to utilize ErrPs for mediation of co-adaptation in HRI. While these studies showed empirical evidence of the utility of this approach, a systematic understanding of the dyadic interacting system (human and robot) remained unexplored. The study here presented addresses this by proposing a computational model of the human counterpart and simulating the integrated dyadic system including the robot. The basis of this approach is the empirical data obtained in our previous study as well as the corresponding experimental setup (see Chapter 5). This allowed both the successful fit and validation of the proposed computational model. A series of model-based simulations were furthermore conducted to systematically explore open questions related to ErrP-based human-robot co-adaptation. The main findings point to ways to improve co-adaptation: (1) by employing an alternative robot learning paradigm, and (2) by enhancing ErrP-decoder performance, with non-error decoding rate playing the more crucial role than error decoding rate. Furthermore, model-based simulations suggest that extensions to more complex robot behavior are possible under the constraint of practical timescales of successful convergence of co-adaptation. These, among other findings presented in this chapter, have practical implications for future steps along this line of research. The proposed computational model may be reused furthermore for simulating future studies prior to conductance and this way help to define testable hypothesis and suitable experimental designs.

6.1. Introduction

Recent studies by others have demonstrated the usability of ErrPs online decoded from a human participant's ongoing EEG signals as a teaching signal for learning and adaptation of robot behavior in HRI [105, 173, 43]. Following this line of research, our recent study extended this principle and showed evidence for the usability of ErrPs to mediate co-adaptation in human-agent interaction (see Chapter 5). While the above works provided evidence supporting this approach in the domain of HRI and HMI in general, they were largely focused on the technical aspects of it (tailored to specific HRI tasks), and are limited in providing an overarching understanding of the dyadic interacting system (human and machine). The involvement of human participants, who vary in their way of perceiving, interpreting, and coping with the overall task and robot stimulus makes the approach not straightforward to understand.

This may be a reason why the approach performed well in some participants and rather poorly in others (see results reported in Chapter 5, Section 5.3).

6.1.1. Related work on modeling human decision-making, error processing, and learning

While research on modeling human (and animal) decision-making has a longstanding history in neuroscience, cognitive psychology, and game theory, it appears rather scarce in the domain of BCI, with a few works over the last decades, such as [217]. Also, in the domain of HRI, modelling human behavior appears to be not yet common sense, although some recent research works demonstrated the potential of improving HRI by including models emulating human behavior (see [30] and [26] for an overview). At the intersection between HRI and BCI, e.g. in the relatively new field of ErrP-mediated collaborative HRI, computational modeling of human behavior appears to be yet entirely unexplored. Therefore, concepts and methods presented here mainly draw from recent and historical works from the domain of cognitive neuroscience.

The monitoring of information conflicts in the human (but also other animals') brain has been consistently associated with the ACC and projecting areas [218]. Over the last two decades the ACC has been implicated with many processes, among others with conflict and error coding, task preparation and response selection, social cognition, pain and effortful control [82]. These processes are subsumed under the term "performance-monitoring" (see Chapter 2, Section 2.3 for further details). A multitude of computational models have been proposed over the last two decades, describing the neurophysiological and behavioral effects related to performance monitoring (see [82] for an overview). Early seminal models were focused on a relatively limited set of tasks and largely based on EEG and fMRI data. Prominent examples are the *conflict-monitoring model* [219], the *error likelihood model* [220], the *RL-ERN model* [93], and the *volatility model* [221]. More recent models aim at providing a unifying perspective on ACC function and related behavioral effects and describe a larger set of empirical findings including data from lesion studies as well as single-cell recordings, both in humans and animals (primates and rodents). Prominent examples are the *PRO (predicted response outcome) model* [101], the *reward value and prediction model (RVPM)* [222], the *hierarchical reinforcement learning model (HRL)* [223], and the *hierarchical error representation model (HER)* [224]. In parallel to attempts for unification, ACC function and related behavioral effects have also been modelled and described in the context of effort control, i.e. allocation of mental resources during difficult or effortful tasks in the absence of error, conflict, and choice [225, 226]. Most of the proposed models employed a reinforcement learning-based computational framework, in particular, temporal-difference learning in actor-critic architectures. A few models, however, utilized different frameworks, such as based on Bayesian inference [221] and predictive coding [227]. Despite the considerable body of work presented so far and the general agreement on the underlying computational framework no single unified model has yet been proposed. The most promising candidate for laying out the basis of a unified account is perhaps the PRO model [101]. In essence, it ascribes the ACC monitoring mechanisms for

(un)predictable outcomes and their deviations from expectations. This provided a framework for understanding a large set of empirical findings, recently also integrating effort control [228]. A direct application of the so far proposed models to our work was challenged by the fact that they were largely instantiated to account for data based on experimental tasks and stimuli typical in the domain of neuroscience (e.g. CRTs, Stroop task, or gambling tasks). In contrast, our experiment encompassed a comparably complex stimulus (robot gaze transitions) in a much less constrained task setting: the stimulus unfolds in time at non-deterministic timescales (robot action sequence is terminated with the self-paced decision of the participant). This lets us to decide on developing a model tailored to our specific problem. An important prerequisite for our proposed model was to allow flexible extensions to different HRI scenarios in future work. The proposed model was mainly inspired by the PRO model, but also others, as acknowledged below.

6.1.2. Aim of the work

The goal of the present work is to obtain insight into the process of ErrP-based human-agent co-adaptation by describing the dyadic interacting system in form of an integrated model. Here, we distinguish between human factors and technical factors as possible contributors to (un-)successfull co-adaptation. Possible human factors are for instance learning/adaptation style or risk-taking attitude; technical factors are for instance experiment design, ErrP decoding performance, and the type of robot/agent learning paradigm. While human factors are naturally hard to control, insights about technical factors in relation to varying human factors allow the formulation of system design guidelines encouraging successful co-adaptation and in turn enhanced human-agent interaction performance. This research is divided into four steps: (1) proposition of a computational model of human decision-making and learning in the context of ErrP-based human-agent co-adaptation, (2) fitting the model according to the empirical data from our previous study (see Chapter 5), (3) embedding the model into an interaction environment with a simulated agent, and (4) performing a series of model-based simulations of the integrated model to elucidate potential factors influencing co-adaptation. This model-based approach contributes to a better understanding of ErrP-based human-agent co-adaptation in two ways:

- By proposing a computational model for human decision-making and learning, validated in the context of our example ErrP-based human-agent co-adaptation scenario. This model shall provide a platform to be reused for the simulation of future studies and this way help to define testable hypothesis and suitable experimental designs prior to conductance.

- By providing insight into ErrP-based human-agent co-adaptation in the context of our example HRI scenario, both from a technical and human behavioral perspective. More specifically, by employing the proposed computational model in model-based simulations to derive estimates about the influence of across-individual varying human factors, recommendations for future technical improvements for the design of ErrP-decoders and ErrP-based robot adaptation, recommendations for the design of future studies, and an estimate as to which extent the approach can be practically scaled to more complex HRI.

This chapter is organized as follows: Section 6.2 introduces the proposed computational model, together with the data fitting procedure in Section 6.2.3, its integration into the co-learning environment 6.2.4, and the final model validation 6.2.5. Section 6.3 presents a series of model-based simulations and results to elucidate factors influencing co-adaptation. Section 6.4 discusses the results and Section 6.5 summarizes the chapter.

6.2. Methods: Computational model of human decision-making and learning

6.2.1. Experimental paradigm and empirical data

This section briefly recapitulates the experimental setup and technical implementation of our previous study; for more detailed information, the reader is referred to Chapter 5. In the experiment, participants had to perform a collaborative task together with a real humanoid robot[1] (see Figure 28), which required co-adaptation from both sides (human and robot). The task was designed as a repeated guessing game in which participants were asked to infer from the robot's gaze behavior which one out of three available objects it had selected (underlying goal / intention). A single guessing game (further referred to as *trial*) started with the robot secretly deciding for one of the three objects and then proceed with alternatingly gazing at either of three objects or the participant in a fixed pace (one gaze transition in 400 ms). Meanwhile the participant would attempt to infer the correct object from observing the robot's ongoing gaze behavior and eventually take a decision for either of the three target objects. This means that participants could continue observing the behaving robot until they were certain about their decision. Subjects were asked to indicate their decision in a self-passed fashion via keypress responses (these were only used as ground truth for validation). Afterwards, the robot would communicate the true choice of object to the participant (feedback); time-locked to the moment of feedback presentation, the participant's EEG-based ERP would be classified online into non-error (match) or error (mismatch) and then administered to the robot for adapting its current gaze policy. This way, the robot's gaze behavior would gradually adapt to participants' expectations; likewise, participants may have adapted their expectations, for instance learned / adopted a gaze pattern "suggested" by the robot. A measure for successful co-adaptation was *guessing performance*, defined as the participant's accuracy in correctly inferring the robot's chosen object from its gaze behavior. Sixteen healthy participants took part in the experiment (age: 29.2±5.0, 7 females, 9 males). Each one performed first a calibration session (CALIB) followed by four co-adaptation sessions (CORL-I-IV). The calibration session consisted of 150 trials (guessing games) in which the robot gaze behavior followed a fixed (e.g. non-adaptive) policy. Data collected during this session was used to derive participant-specific ErrP-decoders. During the following four closed-loop co-adaptation sessions, each consisting of 50 trials, the previously built ErrP-decoder was employed for online adaptation

---

[1] The robotic platform chosen for the experiment was the humanoid robot NAO, which is a commercially available (SoftBank Robotics) 58 cm tall robot with 21-25 degrees of freedom.

of the robot's gaze policy. In each co-adaptation session, the gaze policy was (re-)initialized such that the robot would generate random gaze behavior in the first trial (starting without prior knowledge).
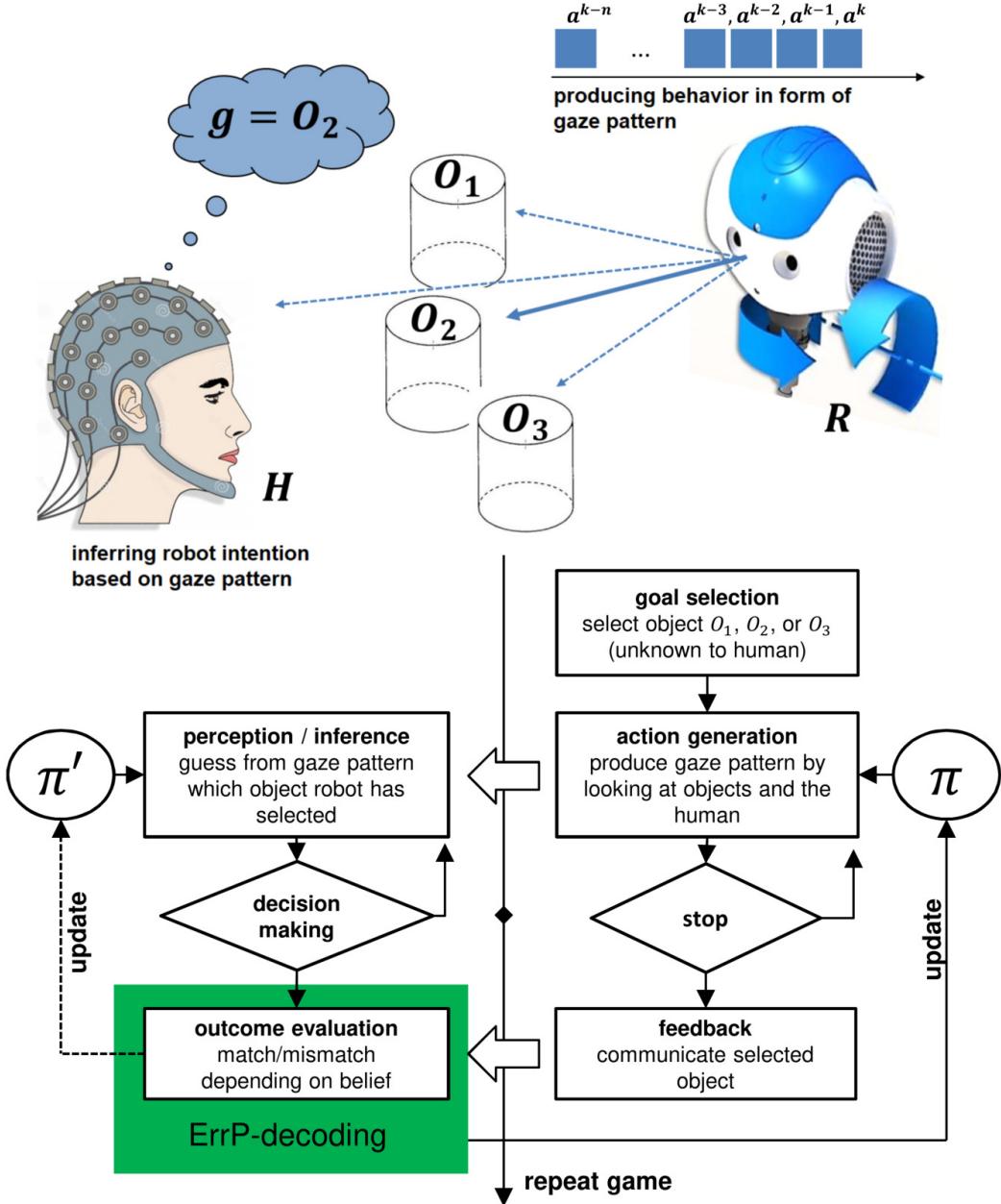
Figure 28 Revisiting the experimental paradigm: Human participant and robot play a guessing game in which the robot covertly selects one out of three objects. Subsequently the robot generates a gaze pattern based on which the participant had to guess the secret object. The participant's brain responses were measured (marked in green) and used as a feedback signal to adapt the robot's gaze policy π, while the participant may have likewise adapted expectations π' about the robot's gaze behavior.

6.2.2. Overview of proposed model

The proposed model consists of four modules (see Figure 29): (1) The *perception and inference* module is responsible for transferring observations of the agent's actions into a belief about the agent's goal. (2) The *decision-making* module is responsible for turning the belief into an explicit action stating the predicted agent's goal. As long as no decision was made,

the perception and inference module continues to observe the agent and further updates the belief. (3) The *outcome evaluation* module receives the agent's feedback and compares it with the predicted agent's goal, based on which it derives a trial outcome in form of a reward signal and a measure of expectation violation. (4) The *learning* module updates the current knowledge in the model based on the prediction error and the history of observed agent actions.
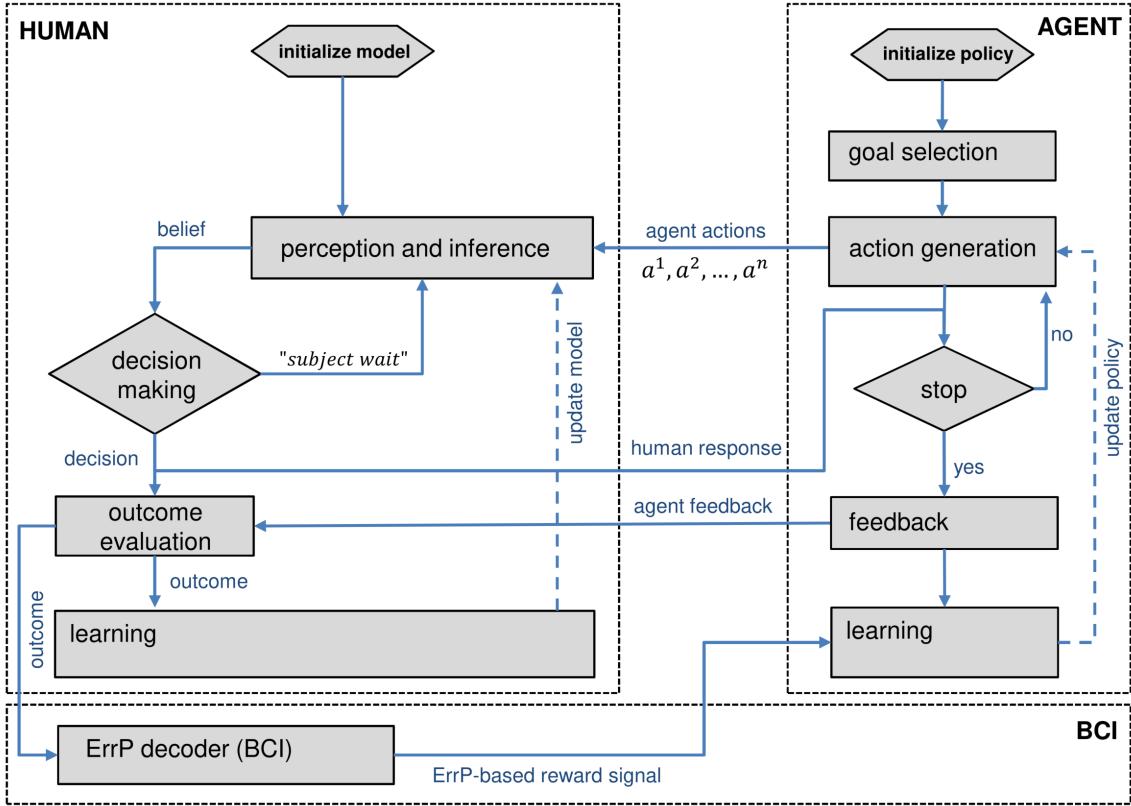
Figure 29 Model architecture: Overview of proposed architecture of the human computational model embedded in the interaction process with the agent. Agent adaptation is performed based on feedback from the evaluation module of the human computational model. Decoding and transmission of that feedback signal derived from online decoded ErrPs are simulated in the BCI block.

6.2.2.1 Perception and inference

Pisauro et al. (2017) recently provided empirical evidence from a combined EEG-fMRI study that humans accumulate evidence in favor of the different alternatives before committing to a decision [229]. To model this evidence accumulation process, we used Bayesian inference based on the formulation of Behrens et al. in 2007 for describing human behavioral data in a perceptual decision-making task [221]. Bayesian inference prescribes a standard computation resulting in a posterior belief $p(g_m|a_{i,j}^k)$ that alternative g_m is true given observation $a_{i,j}^k$. Here, g_m is the alternative among the set of agent goals $g = \{g_{O1}, g_{O2}, g_{O3}\}$, with M being the number of possible goals (or target objects: $M = 3$). $a_{i,j}^k$ is the observed agent action from gaze state s_i to s_j in time step k. The agent can be in four possible gaze states; either gazing at the human or any of the three target objects $S = \{s_{O1}, s_{O2}, s_{O3}, s_H\}$. Transitions are possible from one gaze state to another or staying in the current state. As observations

arrive sequentially over time, the posterior belief $p(g_m|a_{i,j}^{1:k})$ can be updated recursively using all observations $a_{i,j}^{1:k} = \{a_{i,j}^1, ..., a_{i,j}^k\}$ up to time step k. Eq. (6.1) describes the computation of the posterior belief after observation of the first agent action in the current trial t. Before observing any agent action, the belief about the agent's goal is uniform among alternatives, hence the prior is initialized with $p(g_m) = \frac{1}{3}$. The posterior belief is used as the new prior upon observation of the next robot action according to Eq. (6.2). The likelihoods $p(a_{i,j}|g_m)$ are computed based on internal weights $w_{i,j,m}^{t-1}$ of the previous trial $t-1$ using the softmax function according to Eq. (6.3). These weights associate the set of observable actions $a_{i,j}$ with the goal alternatives g_m and as such describe the current knowledge of the model. Upon start of the experiment, the weights are initialized uniformly among all possible actions and goals with $w_{i,j,m}^0 = 0$ for all i, j, m.

$$p(g_m|a_{i,j}^1) = \frac{p(a_{i,j}^1|g_m)p(g_m)}{\sum_{m=1}^M p(a_{i,j}^1|g_m)p(g_m)} \tag{6.1}$$

$$p(g_m|a_{i,j}^{1:k}) = \frac{p(a_{i,j}^k|g_m)p(g_m|a_{i,j}^{1:k-1})}{\sum_{m=1}^M p(a_{i,j}^k|g_m)p(g_m|a_{i,j}^{1:k-1})} \tag{6.2}$$

$$p(a_{i,j}|g_m) = \frac{e^{w_{i,j,m}^{t-1}}}{\sum_{m=1}^M e^{w_{i,j,m}^{t-1}}} \tag{6.3}$$

6.2.2.2 Decision-making

The decision-making module is responsible for turning the current belief $p(g_m|a_{i,j}^k)$ into an explicit action stating the predicted agent's goal $\hat{g}$. Here, we draw from the formulation of the PRO-model by Alexander & Brown in 2011 in which decisions are initiated when the belief exceeds a certain threshold (commonly called *decision bound* and referred to thereafter for the remainder of this chapter) [101]. However, as mentioned initially, the PRO-model was instantiated mainly based on CRTs, in which participants were asked to respond as quickly as possible. In our case, participants could freely decide about when to respond; decision time did not play an explicit role (e.g. was not rewarded or penalized) in the experiment. This means that on the one hand participants could continue observing the agent even though they were already certain about its goal in the hope of collecting more evidence supporting the current belief (delayed response). On the other hand, participants could respond earlier even though they were not certain about the agent's goal, possibly because they lost confidence that future observations would further consolidate the current belief (advanced uninformed response). The latter case is expected in particular in the beginning of the co-adaptation run when the agent behavior follows a random policy which by nature does not allow participants to make informed decisions. Therefore, it is assumed that participant's decisions did not only depend on the belief, but also on the elapsed time spent on observing the agent. Therefore,

the decision variable $\beta(g_m)$ was implemented as a function of the posterior belief and the time step k. Both variables were linearly combined according to Eq. (6.4) and the influence of k was scaled with a factor ϵ (further referred to as *timeout* parameter). This way the decision variable captures the participant's *timeout*-behavior, e.g. participants taking a decision despite limited certainty (posterior belief is similar among options) after some time of observation. Furthermore, decision time decreases relative to an increasing difference of the posterior belief among options. A decision $\hat{g} = \mathrm{argmax}_i \beta(g_m)$ is initiated whenever the maximum of $\beta(g_m)$ is equal or greater than the decision bound Γ, otherwise, the next observation is awaited and the decision postponed, according to Eq. (6.5). The decision bound is recomputed in every trial t based on a Gaussian random process with two free model parameters μ_β and σ_β^2 according to Eq. (6.6).

$$\beta(g_m) = p(g_m | a_{i,j}^{1:k}) + \epsilon k \tag{6.4}$$

$$\hat{g} = \begin{cases} \mathrm{argmax}_m \beta(g_m) & \text{if } \max_m \beta(g_m) \geq \Gamma \\ 0 \text{ "await next action"} & \text{otherwise} \end{cases} \tag{6.5}$$

$$\Gamma \sim \mathcal{N}(\mu_\beta, \, \sigma_\beta^2) \tag{6.6}$$

6.2.2.3 Outcome evaluation

The evaluation module is responsible for validating the outcome of the trial with regard to the decision about the predicted agent's goal $\hat{g}$ and the feedback about the true agent's goal g. First, a binary outcome O is computed based on Eq. (6.7). Second, the prediction error δ is computed according to Eq. (6.8) based on the expected outcome and the true outcome (agent feedback). The expected outcome is the posterior belief of the chosen object at the moment of which the decision was taken.

$$O = \begin{cases} 1 & \text{if } \hat{g} = g \text{ (non-error)} \\ 0 & \text{if } \hat{g} \neq g \text{ (error)} \end{cases} \tag{6.7}$$

$$\delta = O - p(\hat{g} | a_{i,j}^{1:k}) = \begin{cases} 0 \leq \delta \leq +1 & \text{if } \hat{g} = g \\ -1 \leq \delta \leq 0 & \text{if } \hat{g} \neq g \end{cases} \tag{6.8}$$

6.2.2.4 Learning

The learning module is responsible for updating the internal weights after the outcome of the trial has been evaluated. The learning module was based on previous works describing behavioral and neurophysiological responses in CRT and gambling tasks as a reinforcement learning [210] process [101, 82], specifically on the formulation of Cohen and colleagues in 2007 which proposed a model describing human behavioral data in a gambling task [230]. In Eq. (6.9) the weights used in the current trial are updated for the next trial by adding the probabilities of observed actions, weighted by the prediction error δ and scaled by the learning rate λ. As participants may have reacted differently to erroneous versus correct guesses, we introduced separate learning rates for success ($\lambda^{(+)}$) and failure ($\lambda^{(-)}$) outcome. The gating parameter $\eta(\hat{g})$ is 1 for the chosen target and 0 for all non-chosen targets according to Eq. (6.10). This way, the weights associated with the chosen target are either positively or negatively reinforced depending on the outcome of the trial. The discount parameter γ decreases the impact of past trials on the current update if chosen < 1. After updating, the weights are finally turned into updated likelihoods according to Eq. (6.3) used by the perception and inference module in the next trial.

$$w_{i,j,m}^{t} = \begin{cases} \gamma w_{i,j,m}^{t-1} + \eta(\hat{g})\delta\lambda^{(+)}\frac{1}{n}\sum_{k=1}^{n} a_{i,j}^{1:k} & \text{if } O = 1 \\ \gamma w_{i,j,m}^{t-1} + \eta(\hat{g})\delta\lambda^{(-)}\frac{1}{n}\sum_{k=1}^{n} a_{i,j}^{1:k} & \text{if } O = 0 \end{cases} \qquad (6.9)$$

$$\eta(g_m) = \begin{cases} 1 & \text{for } g_m = \hat{g} \\ 0 & \text{otherwise} \end{cases} \qquad (6.10)$$

6.2.3. Parameter fitting

The computational model consists of 8 parameters (γ, $\lambda^{(+)}$, $\lambda^{(-)}$, μ_β, σ_β^2, ϵ, TNR, TPR; see overview of parameters in Table 7). ErrP-decoder true-negative and true-positive rates (TNR, TPR) were determined based on the online decoding performance during the experimental study[2]. The remaining four model parameters ($\lambda^{(+)}$, $\lambda^{(-)}$, μ_β, σ_β^2) were fit to the behavioral data of each participant individually. The original experiment data consisted of 16 participants out of which the data of one participant (s06) had to be removed due to incomplete marker information. Model fitting was performed in a sequential two step procedure: (1) Success and failure learning rates were first fitted using an optimization procedure, and (2) the decision bound parameters were retrieved by fitting a Gaussian process to the distribution of $\beta(\hat{g})$ across trials. Finally, the timeout parameter ϵ turned out to only affect how well the model describes the experimental data in terms of decision times (number of observed agent actions until participant decision). Since decision times were not in the focus of our investigation, the

---

[2] In accordance to the definition in the Chapters 4 and 5, non-error (match) events were defined as the negative class, whereas error (mismatch) events were defined as the positive class

Table 7 Overview of model parameters: participant-specific learning rates and decision bound are estimated via the fitting procedure (fit); participant-specific ErrP-decoder rates are taken directly from the data obtained during the experimental study (exp). Timeout and discount factor were empirically set and fixed across participants.

Parameter	Value	Description	Equation
ϵ	0.02	timeout scaling factor	Eq. (6.4)
γ	1.0	discount factor	Eq. (6.9)
TNR	exp	non-error decoding rate	Eq. (6.16)
TPR	exp	error decoding rate	Eq. (6.16)
$\lambda^{(+)}$	fit	Success learning rate	Eq. (6.9)
$\lambda^{(-)}$	fit	Failure learning rate	Eq. (6.9)
μ_β	fit	mean decision bound	Eq. (6.6)
σ_β^2	fit	variance decision bound	Eq. (6.6)

timeout parameter was empirically set to $\epsilon = 0.02$ during model validation described in Section 6.2.5. This resulted in a good match of simulated and real decision times for all participants.

6.2.3.1 Fitting success-failure learning rates

The computational model was presented the actions performed by the robot in the actual experimental sessions, starting with the calibration session and continuing with the co-adaptation runs CORL-I, CORL-II, and CORL-IV[3]. The model started with zero knowledge and continued learning across the entire experiment. The model's simulated decisions were compared to the participant's real decisions on a single-trial level. The cost function to be minimized is depicted in Eq. (6.11) and computes the RMSE (root mean squared error) between the simulated $O_{SIM}^{1:t}$ and the actual participant's trial outcomes $O_{DATA}^{1:t}$ across all trials T (see Eq. (6.7)). Outcomes of individual trials were first smoothed with a 10-trial kernel before being administered to Eq. (6.11), a procedure often used to examine correspondence between model predictions and behavioral data [231, 232, 233, 230]. The cost function was computed separately for the calibration (150 trials) and the three co-adaptation sessions (each 50 trials) and subsequently averaged to obtain a single measure of goodness of fit F_{total} across the entire experimental session according to Eq. (6.12). Optimization of success and failure learning rates were performed in a sequential two-step procedure based on Eq. (6.13): (1) by localization of the global minimum using a 2D grid search across the values $\lambda^{(+)} = \{0, 0.2, ..., 3\}$ and $\lambda^{(-)} = \{0, 0.2, ..., 1\}$, and (2) by fine-grained optimization using the nonlinear, uncon-

---

[3] In CORL-III a different protocol was employed which did not capture ground truth information about participant's decisions; therefore it was excluded from this research.

strained Nelder-Mead simplex method [234] starting with initial values based on the previous grid search[4].

$$F = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (\overline{O}_{SIM}^{1:t} - \overline{O}_{DATA}^{1:t})^2}; \ T = \text{number of trials} \tag{6.11}$$

$$F_{total} = \frac{1}{4}(F_{CALIB} + F_{CORL-I} + F_{CORL-II} + F_{CORL-IV}) \tag{6.12}$$

$$min_{\lambda(+),\lambda(-)}(F_{total}) \tag{6.13}$$

6.2.3.2 Fitting decision bound

The discrete probability distributions of the decision bound across trials were approximately normal distributed in all participants[5]. Therefore, the variations of the decision bound were assumed to result from a Gaussian random process reflecting the participant's uncertainty about the underlying belief. Based on this, the decision bound of each participant was modelled individually via a Gaussian random process according to Eq. (6.14), whereby x is the discrete probability distribution of the decision bound (described by μ_β and σ_β^2) based on trials in which the model's decision matched the decision of the participant ($\hat{g}_{SIM} = \hat{g}_{DATA}$) according to Eq. (6.15).

$$f(x|\mu_\beta, \sigma_\beta^2) = \frac{1}{\sigma_\beta \sqrt{2\pi}} e^{-(x-\mu_\beta)^2/2\sigma_\beta^2} \tag{6.14}$$

$$x = \beta(\hat{g})^{1:t} \text{ for all } \hat{g}_{SIM}^t = \hat{g}_{DATA}^t \tag{6.15}$$

6.2.3.3 Fitting results

Exemplary results of the fitting procedure for six participants can be visually inspected in Figure 30 (for all participants, see Supplementary Figure 45, Appendix A); numerical results are provided in Table 8. Figure 30 shows that the simulated learning curves match well those of

---

[4] Note, that this procedure assumes stationary learning rates for each participant, since optimization is performed across the entire experiment, including the calibration and co-adaptation sessions. Although, it is likely that participants changed/adapted their learning rates throughout the experiment, in particular, during the transition from calibration to co-adaptation, this assumption was made because of the overall limited amount of data and to constrain the complexity of the model.

[5] This was observed by visual inspection of the decision bound histograms for each participant individually. Since no obvious deviations from normal distributions were observed, no further quantitative tests were performed.

the experimental data with almost perfect fit in some participants (e.g. s03, s10) and moderate fit in other participants (e.g. s04, s11). Table 8 provides details about the overall goodness of fit (F_{total}), the goodness of fit for each experimental session (F_{CALIB}, $F_{CORL-I-IV}$), and the percentage of correctly predicted participant responses. Furthermore, the optimized values for the free model parameters ($\lambda^{(+)}$, $\lambda^{(-)}$, μ_β, σ_β^2) are provided. Overall, simulations match well the experimental data with an average across participants goodness of fit of $F = 0.19 \pm 0.04$ (AVG±SD). This translates to an average percentage of correct predictions of participant decisions of $PRED = 70.4 \pm 11.1\%$ (AVG±SD), with 12 out of 15 participants resulting in $PRED \geq 65\%$. The results suggest that the model captured relevant behavioral effects observed in the human data.
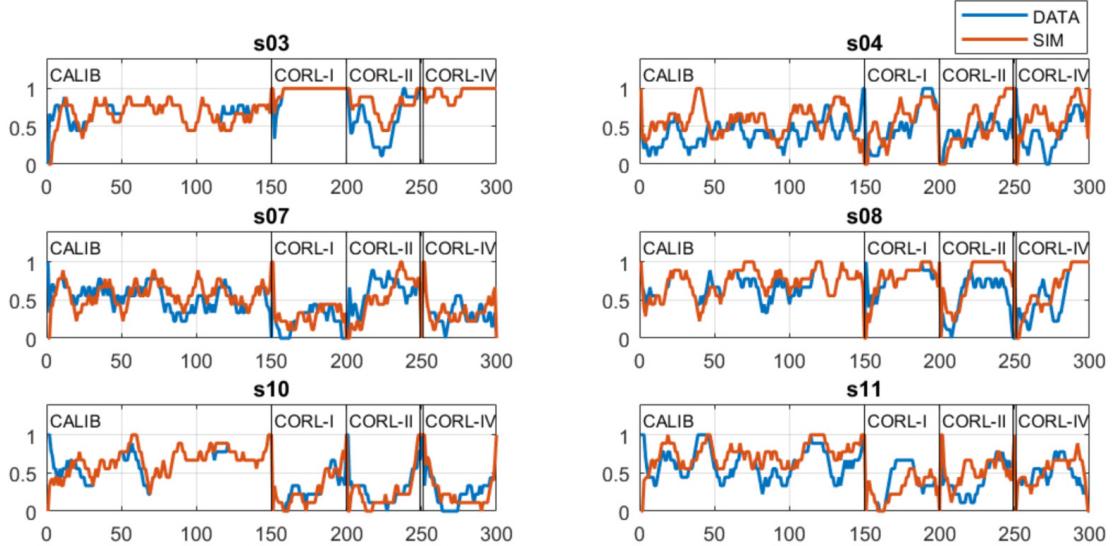
Figure 30 Exemplary qualitative results of goodness of fit for six participants: Each panel depicts individual results of simulated (orange) and real (blue) guessing performance per participant for CALIB, CORL-I, CORL-II, and CORL-IV (smoothed with a 10-trials kernel for each part of the experiment separately). Overall, the results show a good match between simulation and ground truth, indicating that the model captured relevant behavioral effects observed in the human data.

6.2.4. Integration of the human computational model into the co-learning environment

After fitting the free parameters of the computational model based on the experimental data, the model is ready to be integrated into the co-learning environment (see Figure 29) including the adaptive robotic agent, and the BCI (ErrP-decoder) providing implicit human feedback to the agent. Algorithm 1 below describes the perception and inference, decision-making, outcome evaluation, and learning procedure of the human participant. Algorithm 2 describes the action generation and learning procedure of the robotic agent. The simulation of the ErrP-decoder interfacing human and agent is described in Eq. (6.16). The function computes a reward estimate $\tilde{R}$ for agent policy updating (see Algorithm 2) based on the true outcome of the trial and a comparison of a uniform random number $\mathcal{U}(0, 100)$ with the participant's individual ErrP-decoder rates. As such, the function emulates the derivation of a reward from an imprecise ErrP-decoder decision scaled with the experimental parameters TNR and TPR.

Table 8 Optimized free model parameters ($\lambda^{(+)}, \lambda^{(-)}, \mu_\beta, \sigma_\beta^2$), experimentally determined online ErrP-decoder performance (TNR, TPR [%]) and resulting goodness of fit F, expressed as RMSE according to Eq. (6.11) and as percentage of correct predictions of participant decisions ($PRED$ [%]).

ID	$\lambda^{(+)}$	$\lambda^{(-)}$	μ_β	σ_β^2	TNR	TPR	F_{CALIB}	F_{CORL-I}	$F_{CORL-II}$	$F_{CORL-IV}$	F_{total}	PRED
			Model Parameters						Fitting Results			
s03	1.46	0.27	1.27	0.24	92.10	69.03	0.12	0.06	0.26	0.00	0.11	87.00
s04	0.00	0.20	0.53	0.08	96.80	86.40	0.29	0.20	0.29	0.33	0.28	44.00
s05	2.40	0.80	1.04	0.21	62.83	42.17	0.13	0.35	0.24	0.25	0.24	66.00
s07	2.20	1.00	0.87	0.24	83.17	81.87	0.19	0.16	0.24	0.18	0.19	56.33
s08	2.59	0.83	1.07	0.19	95.53	71.07	0.12	0.17	0.31	0.24	0.21	77.33
s09	2.60	0.40	1.18	0.19	90.10	77.03	0.16	0.09	0.08	0.17	0.13	84.67
s10	0.20	0.16	0.68	0.14	100.00	0.97	0.14	0.12	0.19	0.18	0.16	72.00
s11	2.40	0.40	1.11	0.23	76.53	56.57	0.23	0.19	0.24	0.17	0.21	60.67
s12	2.81	0.33	1.32	0.33	95.00	71.57	0.16	0.29	0.23	0.18	0.22	74.67
s13	2.63	0.43	1.16	0.19	71.73	63.90	0.15	0.20	0.28	0.22	0.21	69.33
s14	0.69	0.33	1.03	0.22	90.63	66.00	0.16	0.10	0.21	0.27	0.19	73.67
s15	1.00	0.40	1.01	0.23	90.53	68.17	0.17	0.00	0.32	0.05	0.14	82.00
s16	2.00	0.20	1.16	0.34	58.37	81.47	0.24	0.15	0.15	0.06	0.15	73.00
s17	2.60	0.40	1.12	0.19	84.30	89.43	0.12	0.22	0.22	0.14	0.18	68.67
s18	1.40	0.80	0.98	0.22	89.00	65.17	0.12	0.17	0.39	0.08	0.19	67.00
AVG	**1.80**	**0.46**	**1.04**	**0.20**	**85.1**	**66.1**	**0.17**	**0.17**	**0.24**	**0.17**	**0.19**	**70.4**
±SD	0.94	0.26	0.22	0.05	12.5	21.6	0.05	0.09	0.07	0.09	0.04	11.1

$$\tilde{R} = \begin{cases} +1 & \text{if } (O = 1)\&(\mathcal{U}(0, 100) \le TNR); \text{else } -1 \\ -1 & \text{if } (O = 0)\&(\mathcal{U}(0, 100) \le TPR); \text{else } +1 \end{cases} \tag{6.16}$$

6.2.5. Model validation

The fitting results demonstrated that the model is capable of sufficiently capturing behavioral effects observed in the human data. However, whether and to what extent the model can be employed for large-scale simulations of the co-learning system require additional validation of the integrated model including the simulated agent and the simulated BCI interface. As such, the following sections assess the validity of the computational model in the context of three questions: (1) do large-scale simulations result in outcomes which are consistent with the empirical data observed in the experimental study? (2) does the model account also for / explain neurophysiological effects despite not having been explicitly used for model fitting? (3) Is the model complexity justified or can the model be simplified, e.g. by reducing the number of free parameters?

Algorithm 1: HUMAN perception and inference, decision-making, outcome evaluation, and learning algorithm

---

Initialize weights $w_{i,j,m}^0 = 0$ for all i, j, m

for $t \leftarrow 1$ **to** T **do**

    Initialize prior belief: $p(g_m) = \frac{1}{M}$

    Initialize observed action count: $k = 0$

    Initialize observed action history: $A_{i,j} = 0$ for all i, j

    Compute likelihoods based on weights: $w_{i,j,m}^{t-1}$ [Eq. (6.3)]

    Compute decision bound: $\Gamma \sim \mathcal{N}(\mu_\beta, \sigma_\beta^2)$ [Eq. 6.6]

    while $\beta(g_m) < \Gamma$ *("await next action")* **do**

        $k \leftarrow 1$

        Observe agent action a^k

        Update posterior belief based on prior and likelihood of action a^k [Eq. (6.2)]

        Update decision variable $\beta(g_m)$ [Eq. (6.4)]

        Update prior belief: $p(g_m|a_{i,j}^{1:k}) \leftarrow p(g_m|a_{i,j}^{1:k-1})$ [Eq. (6.1) and (6.2)]

        Update observed action history: $a_{i,j}^k \leftarrow 1$

    end

    Take decision $\hat{g}$ [Eq. (6.5)]

    Observe agent feedback g

    Evaluate outcome O [Eq. (6.7)]

    Compute prediction error δ [Eq. (6.8)]

    Update weights $w_{i,j,m}^t \leftarrow w_{i,j,m}^{t-1}$ [Eq. (6.9)]

end

Algorithm 2: AGENT action generation and learning algorithm

Initialize policy: $\pi = \frac{1}{4}$ for all s, a

for $t \leftarrow 1$ **to** T **do**

    Initialize action count: $k = 0$

    Choose initial gaze state s_i with $s = \mathcal{U}\{s_{O1}, s_{O2}, s_{O3}, s_H\}$

    Choose goal g with $g = \mathcal{U}\{g_{O1}, g_{O2}, g_{O3}\}$

    Initialize gaze transition history: $a_{i,j} = 0$ for all i, j

    while $\hat{g} = 0$ *("participant wait")* **do**

        $k \leftarrow 1$

        Choose action a_j from state s_i using $\pi(g)$

        Execute action a_j (perform gaze shift), observe next gaze state s_j

        Update gaze transition history: $a_{i,j}^k \leftarrow 1$

        $s_i \leftarrow s_j$

    end

    Present feedback g to participant

    Observe reward $\tilde{R}$ from ErrP-decoder

    Update policy: $\pi \leftarrow \pi + \tilde{R}\alpha \sum_k a_{i,j}^{1:k}$

    Truncate policy: $\pi \leftarrow clamp(\pi)_0^1$ for all s, a

    Normalize policy: $\pi \leftarrow \frac{\pi}{\sum_s \pi}$

end

6.2.5.1 Comparison of large-scale simulation with experimental data

This section validates the model by comparing simulated data with the experimental data. To have equal conditions for the comparison, the original experimental protocol was simulated with a large number of individual experiments. The simulation was focused on the calibration-session (CALIB) and subsequent first co-adaptation run (CORL-I) only, as these made up the most critical parts of the original experiment. The comparison between simulated and real experimental data was performed on the "study-level", e.g. a single simulated study comprised all 15 participants and reported measures of co-adaptation performance averaged across participants. Four measures of co-adaptation performance were compared:

- *guessing performance*: mean percentage of guessing performance in 50 trials of CORL averaged across participants.

- *success-rate*: percentage of successful[6] co-adaptation runs across participants.

- *abs. decision time*: mean absolute number of gaze transitions until participant decision in 50 trials of CORL averaged across participants.

- *rel. decision time*: mean percentage change of number of gaze transitions until participant decision relative to the first 10 trials in 50 trials of CORL averaged across participants.

In total 500 studies were simulated; distributions of the above measures were then compared to the empirical measures from the original study. The underlying proposition of this comparison is as follows: the model reflects well the experimental data (and can be trusted at the level of large-scale simulations) if the empirical measures lie within the single standard deviation of the simulated measures.

Results are depicted in Figure 31(a) and show the empirical performance measures (dashed green line) superimposed on the distributions of simulated measures. Sufficient consistency between simulated and real data can be observed for success-rate, abs. and rel. decision time. Simulated average guessing performance turned out to be consistently higher than the empirical average guessing performance (see Figure 31(a), upper-left panel). This discrepancy could be due to the fact that the model could not be fitted equally well to all participants. Fitting results turned out rather low in a few participants (see Table 8). A comparison was therefore performed based only on participants for which the model could be well fitted, specifically, for participants whose decisions the model could predict with $PRED > 65\%$ (this excluded three out of 15 participants: s04, s07, and s11). Results, depicted in Figure 31(b) show that now empirical measures (dashed green lines) are well within the single standard deviation of the simulated data (dashed black lines) for all four measures. This consistency between the empirical and simulated data corroborates the above hypothesis that the observed discrepancy was essentially due to the cases for which no good fit could be found. However,

---
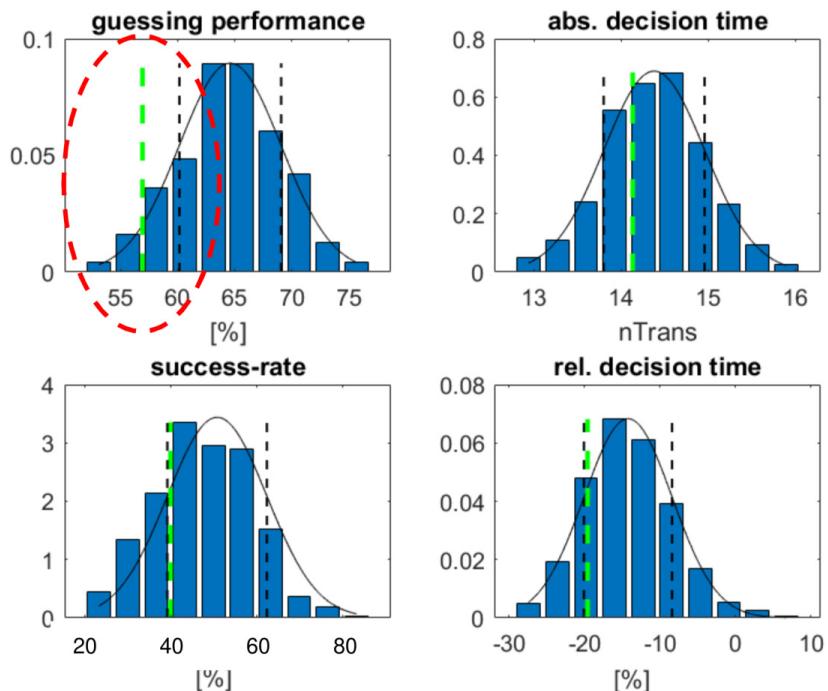
[6] Definition according to the analysis of experimental data (see Chapter 5, Section 5.3.2): A co-adaptation run consisting of 50 trials was denoted successful if guessing performance was $\geq 70\%$ in three subsequent segments of 10 trials, $p < 7.6 * 10^{-6}$, one-sided binomial test with $p_{chance} = 1/3$.

they also show the limitations of the model, since the latter can only account for a subset (approximately 80%) of the participants. Nonetheless, the consistency observed validates using the model for large-scale simulations, as long as these are based and constrained to this subset of participants. All further investigations and simulations therefore excluded participants s04, s07, and s11 from the analyses.

6.2.5.2 Relationship between model prediction error and neurophysiological characteristics of error-related potentials

So far, the computational model has been considered only in the context of behavioral data, as reflected in per trial participant decisions, corresponding outcomes, and decision time. This section will investigate the question of whether the computational model predicts the neurophysiological effects of ErrPs, as reflected in the participant's EEG signals during the experiment. The purpose of this analysis is to provide further evidence supporting the validity of the proposed computational model in the given experimental paradigm. Prior research [93, 101, 230] suggested that the ErrP reflects the prediction error in reinforcement learning-based models of the PFC. Following this proposition, the present analysis is based on the hypothesis that the model prediction error δ (see Eq. (6.8)) correlates with specific components of the ErrP on a single trial basis. The analysis was performed on the data of each participant individually by computing Spearman's rank correlation coefficients between the single-trial ERP and δ over all trials. δ can take on positive and negative values for correct and incorrect outcomes, respectively (see Eq. (6.8)). The absolute prediction error ($abs(\delta)$) can be understood as expressing the level of discrepancy between expectation and outcome irrespective of correct or incorrect outcome and was therefore used for this analysis. Correlation coefficients were computed for all EEG channels ch and sample time-points ts individually (see Eq. (6.17)) for each participant individually and across all trials including CALIB, CORL-I, CORL-II, and CORL-IV (300 trials per participant). Computation of the grand average across participants ($n = 12$) was performed by averaging correlation coefficients of individual participants for each ch and ts. Grand average results of spatio-temporal distribution of correlation coefficients are depicted in topographic plots in Figure 32. Overall, the results show significant negative correlations during times in which the N200 component is expected (Figure 32(a-i)) and significant positive correlations during times in which the P300 component is expected (Figure 32(a-ii)) (one-sample Wilcoxon signed rank test, $p < 0.05$, FDR (false discovery rate)-corrected, $n = 12$). In both cases, the spatial patterns are fronto-centrally focused. This means, larger absolute prediction errors are reflected in more negative N200 and more positive P300 deflections. In addition, a late posterior negative correlation was observed around 800ms (Figure 32(a-iii)). The origin of this effect is currently unexplained. The mid and bottom row of Figure 32(b-c) show the condition-wise correlation coefficients for non-error and error trials, respectively. Around N200 latencies, negative (but insignificant) correlation is again observed for error trials (Figure 32(c-iv)), whereas non-error trials show a weaker positive correlation around the latencies of N200. The P300 effect does only become present in the condition-wise analysis with a weak positive correlation in the non-error condition. The

(a) all subjects

(b) 12 subjects (PRED > 65%)
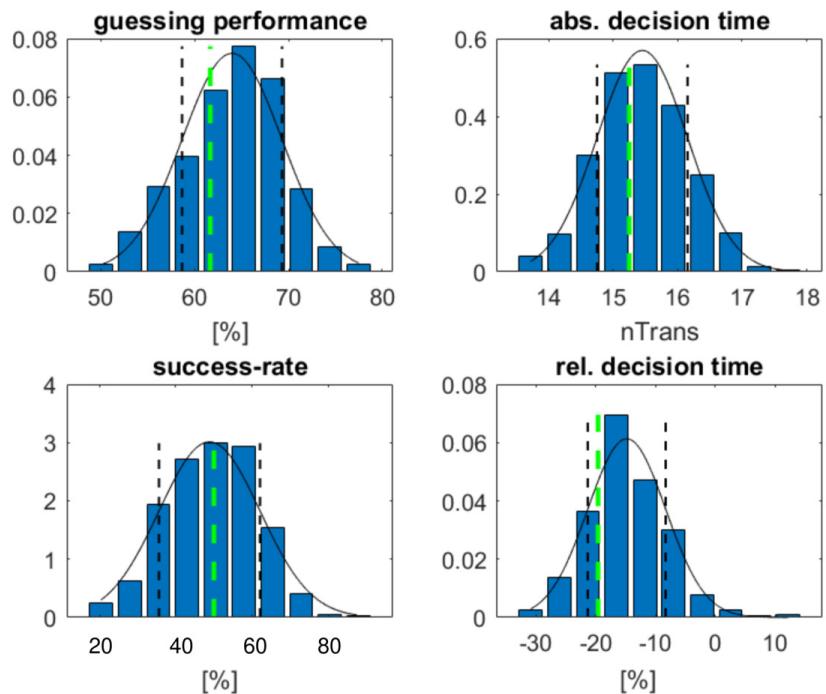
Figure 31 Large-scale simulation of original experimental protocol: Distributions of performance measures (guessing performance, success-rate, abs. and rel. decision time) across 500 simulated studies including all participants (a) and a subset of participants whose decisions the model could predict with $PRED > 65\%$ (b). Results show that empirical measures of the original study (green dashed lines) are well within the single standard deviation of the simulated data (black dashed lines) for all four measures when considering the subset of participants compared to considering all participants.

results demonstrate that, across the subset of tested participants ($n = 12$), the model prediction error reflects the neurophysiological characteristics of ErrPs. Given that this relationship was not explicitly introduced into the proposed computational model, this finding provides not only strong evidence for its validity, but also for its biological plausibility (participant-specific spatio-temporal distribution of correlation coefficients are depicted as topographic plots in Supplementary Figure 46, Appendix A).

$$\rho = corr(ERP^{1:t}_{ch,ts}, abs(\delta^{1:t})) \tag{6.17}$$
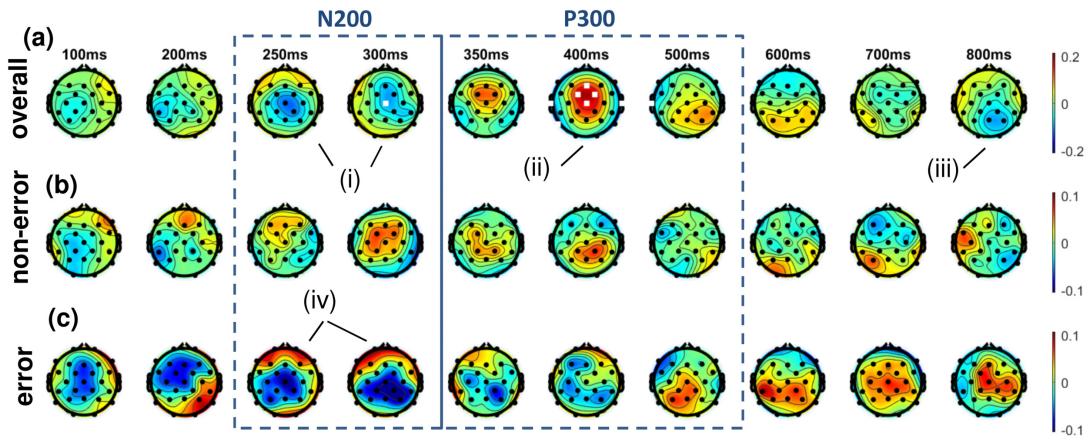
Figure 32 Single-trial spatiotemporal correlation coefficients between absolute model prediction error and ERP: (a) Upper row shows correlation coefficients across all trials. Significant negative correlations are observed during N200 latencies (i), significant positive correlations are observed during P300 latencies (ii) and a late negative correlation are observed around 800ms (iii). The mid (b) and bottom (c) rows show condition-wise correlation coefficients for non-error and error-trials, respectively. Also here, error-trials show negative correlations during N200 latencies (iv). White markers show significant channels at corresponding time-points (one-sample Wilcoxon signed rank test, $p < 0.05$, FDR-corrected, $n = 12$).

6.2.5.3 Assessment of model complexity

An important consideration in computational modeling is the assessment of the potential to simplify the model. A slim model with few parameters not just facilitates the fitting procedure, but also benefits its flexible use in future works. One way to identify if the model can be simplified is to assess dependencies between free model parameters. Strong dependencies would indicate that parameters share information and may be combined into a single parameter. Table 9 shows Spearman's cross-correlation coefficients between model parameters across the subset of participants ($n = 12$). Two pairs of parameters show significant positive correlations: success learning rate and mean decision bound, as well as mean decision bound and error decoding rate during online operation (TPR). A tentative explanation for the former dependency could be that participants with higher mean decision bounds were more certain when making decisions because they collected on average more evidence. In turn, successful outcomes matching highly certain beliefs could have resulted in quicker learning compared to successful outcomes matching uncertain beliefs, hence modulating success learning rates to be higher. A tentative explanation for the latter relationship could be as follows: If we assume

that the later a decision is made (higher mean decision bound) the more certain participants were of their decision, then, in case expectations are not met, the mismatch will be larger and will correspondingly elicit a more distinct and easier to decode ErrPs. Vice-versa, earlier decisions (with participants having had less certainty about the expected outcome), should elicit less distinct, and therefore more challenging to decode, ErrPs. Interestingly, success and failure learning rates are not significantly correlated. This supports the initial assumption of the necessity of including separate learning rates for both, success and failure outcomes. Overall, the results indicate that the model can be simplified since free model parameters share information.

Table 9 Spearman's cross-correlation between free model parameters computed for the subset of participants ($n = 12$). Asterisks (*) denote significance level at and $p < 0.05$.

$n = 12$	$\lambda^{(+)}$	$\lambda^{(-)}$	μ_β	σ_β^2	TNR	TPR
$\lambda^{(+)}$	x	.30	**.73***	-.05	-.26	.46
$\lambda^{(-)}$		x	-.21	-.31	-.24	-.14
μ_β			x	.41	-.11	**.62***
σ_β^2				x	-.13	.31
TNR					x	-.14
TPR						x

6.2.5.4 Overview of model validation results

This section investigated the validity of the proposed human computational model covering three separate aspects. Part one assessed as to which extent model-based large-scale simulations are consistent with the experimental data. Results revealed a discrepancy between simulation and real data. This discrepancy could be resolved by removing three out of 15 participants whose experimental data could not be well described by the model during the fitting procedure (see Section 6.2.3). While this result pointed out that the model is suitable for large-scale simulations, it also showed that it is limited to trustworthy description of approximately 80% of the participants recruited for the study. This finding will be further discussed in Section 6.4. Part two supported further the model by showing significant correlations between the model prediction error and the amplitude of relevant components of the ErrP. This finding underscores the biological plausibility of the model, in that it reflects neurophysiological effects, even though these data not having been explicitly included into the model fitting procedure. Part three assessed dependencies between model parameters and revealed significant relationships between some parameters. This yielded suggestions on how the model can be simplified which is further discussed in Section 6.4.

6.3. Model-based simulations and results

The following sections describe simulation-based experiments investigating the impact of various factors of the integrated human-agent system on the overall performance of co-adaptation. All simulations were conducted based on the premise that model parameters are stationary (do not change throughout experiments) and independent of varying experimental conditions (are not modulated by and can be transferred across varying experimental conditions)[7]. Among factors potentially contributing to and influencing co-adaptation, the following were investigated: (1) experimental protocol, (2) ErrP-decoder performance, (3) human factors as described by free model parameters, (4) the agent learning paradigm, and (5) state-space dimensionality in which human and agent operate.

The following measures were defined to quantify co-adaptation performance and to compare different simulated conditions.

- *guessing performance*: mean percentage of guessing performance in 50 trials of CORL averaged across participants.

- *segment-wise guessing performance*: percentage of correct trial outcomes in a segment of 10 trials.

- *success-rate*: percentage of successful[8] co-adaptation runs across participants.

- *convergence horizon*: number of iterations at which a policy is considered to have converged. The convergence is quantified based on the policy change rate $\Delta\pi = \max(|\pi^{t-1} - \pi^t|)$; the horizon is reached once the policy change rate falls below a pre-defined threshold ($\Delta\pi < 1e-4$).

All simulations were performed by running large sets of individual experiments on single participant level, e.g. in each experiment the model parameters ($\lambda^{(+)}, \lambda^{(-)}, \mu_\beta, \sigma_\beta^2, TNR, TPR$) were fixed to one specific participant. To capture robust statistics, all simulations comprised 50 experiments per participant and condition, e.g. 50 simulated studies. This resulted in 600 experiments (12 participants x 50) per condition, if not stated otherwise.

6.3.1. Simulation I: Variations of the experimental protocol

The original experimental study started with 150 trials of calibration with a fixed and non-adaptive agent policy, followed by four separate closed-loop co-adaptation runs in which the agent policy was adapted based on online decoded ErrPs (each comprising 50 trials). These

---

[7] Please note, that some parameters are likely dependent on varying experimental conditions, such as learning rates, which can be assumed to be influenced by ErrP-decoder performances. This consideration is left for future work and involves further empirical studies in conjunction with revisions of the model (see Section 6.4 for details). At the current stage of research, the presented simulation-based approach is the most practical means to obtain insight into how to optimize the overall system.

[8] Definition according to the analysis of experimental data (see Chapter 5, Section 5.3.2): A co-adaptation run consisting of 50 trials was denoted successful if guessing performance was $\geq 70\%$ in three subsequent segments of 10 trials, $p < 7.6 * 10^{-6}$, one-sided binomial test with $p_{chance} = 1/3$.

design choices were a reasonable trade-off between the total experiment duration (< 1h) and a sufficient number of trials (see Chapter 5, Section 5.2.2). At the time of experiment design and conductance, it was unclear if the design choices were in favor of successful co-adaptation. In order to elucidate the role of these design choices, this section investigates co-adaptation performance in the context of varying experimental protocols and addresses two separate questions: (1) The amount of iterations per CORL were fixed to 50 trials in the original study and experimental results suggested that the number of iterations were insufficient for policy convergence (see Chapter 5, Section 5.3.2). Model-based simulations allow prolonging the CORLs and as such to obtain insight into long-term dynamics of co-adaptation and the expected number of iterations for policy convergence. (2) Furthermore, the original experiment required a calibration session which was nearly identical to the subsequent co-adaptation runs. While this was required to obtain participant-independent ErrP-decoders, it remained unclear if and to what extent the calibration session influenced subsequent co-adaptation in terms of participant priming. The second question thus explores the extent to which the calibration session influenced performance in subsequent co-adaptation. Therefore, simulations were performed by comparing guessing performance and success-rate under two conditions, both in the context of prolonged co-adaptation (500 trials per experiment): (a) with preceding calibration session (as in the experimental study), and (b) without preceding calibration session. The latter condition simulates the situation of participants starting to co-adapt with the agent without having been exposed, and therefore primed by the preceding calibration session involving a non-adaptive agent. In both conditions, participant-individual measures of online ErrP-decoder performance (TNR and TPR) from the original experiment were used.

Results are depicted in Figure 33 and show the development of average guessing performance (Figure 33(a)), absolute decision time (Figure 33(b)), and policy change rate $\Delta \pi$ (Figure 33(c)), over the course of 500 trials. Regarding the first question, the results show that the system does not converge within the first 50 trials. The average learning curve suggests convergence to be rather expected around trial 150 which translates to an experiment time of approximately 18-24 min[9]. The results of this simulation suggest that convergence would have been more likely if participants had been given more time during the co-adaptation runs in the original study (e.g. ≥ 150 trials). An important observation from this simulation is that guessing performance does not approach optimal values in the long run, but on average around 90%. This observation is consistent with the average of long-term policy change rate which approaches approximately $\Delta \pi \approx 0.1$. This indicates that the system did not converge within 500 trials in a certain amount of simulated experiments. Regarding the second question, the results indicate that the preceding calibration-session has a positive effect on co-adaptation performance in terms of the steepness of the learning curve, particularly during the first 150 trials (see gray-shaded marked significantly different segments in Figure 33, Wilcoxon paired signed rank test, $p < 0.05$, FDR-corrected). This observation is in line with the original hypothesis in that co-adaptation performance is expected to be better if participants start with a prior

---

[9] Approximation based on the average duration of a single CORL in the experimental study, see Chapter 5, Table 5.

acquired during the calibration session. Nevertheless, in the long run (after approximately trial 150), both average learning curves become indistinguishable, indicating that the system converges successfully and within the same timescale also when participants start without any prior. This finding suggests that the calibration-session is not required for successful co-adaptation.
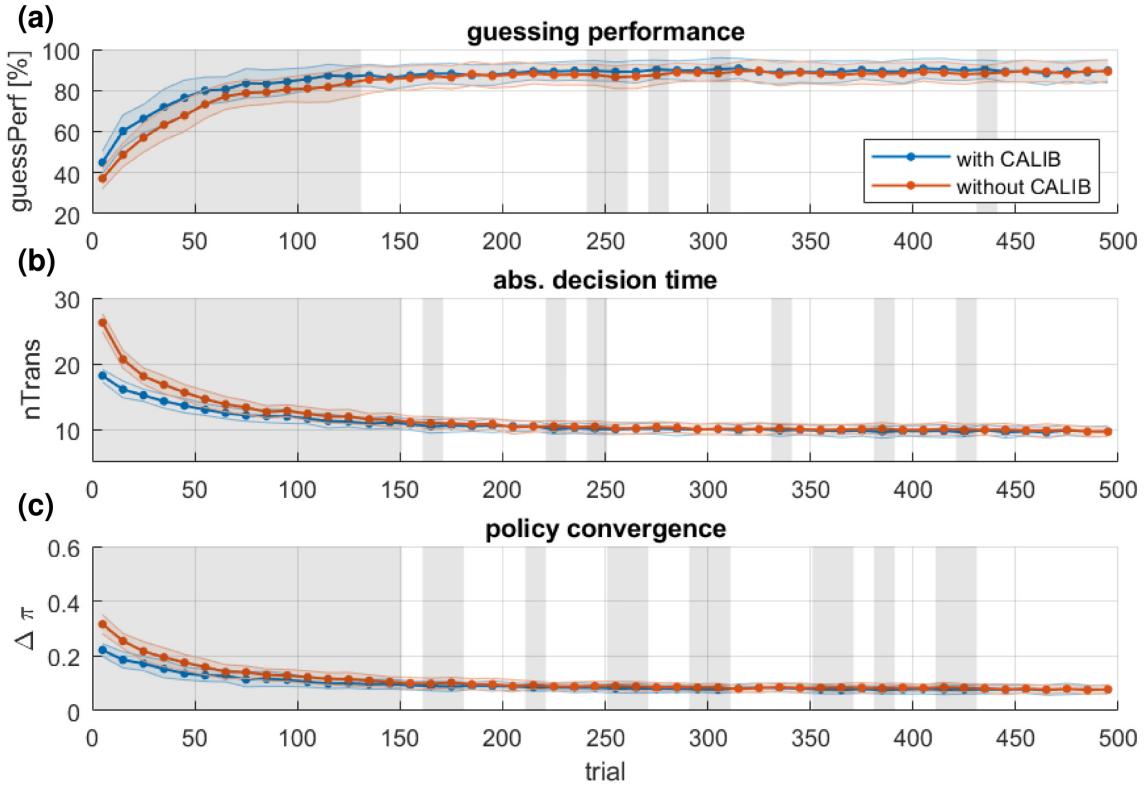
Figure 33 Long-term dynamics of co-adaptation and influence of CALIB on CORL: (a) average guessing performance, (b) absolute decision time, (c) policy change-rate $\Delta\pi$ over the course of 500 trials. Shaded areas represent the single standard deviation of the mean across 50 simulated studies. Results indicate the system converges around trial 150 (after approximately 20-30 min). Gray shaded areas indicate significant differences between co-adaptation runs which were preceded by calibration sessions and those which were not (Wilcoxon paired signed rank test, $p < 0.05$, FDR-corrected). Significant differences are evident during the first 150 trials in which preceding CALIB has a positive effect on co-adaptation performance.

6.3.2. Simulation II: Impact of ErrP-decoder performance

This section addresses the question of how ErrP-decoder performance influences co-adaptation performance. Higher ErrP-decoder performances are naturally assumed to yield better co-adaptation performance as the amount of false feedback provided to the robotic agent is reduced. This relationship was also observed in the results of the empirical study (see Chapter 5, Section 5.3.2). What remained unclear is to what extent hypothetical improvements of ErrP-decoder performance would benefit co-adaptation performance. Furthermore, the lower boundary of ErrP-decoder performance that would still lead to acceptable co-adaptation performance is currently unknown. As in the previous simulation (see Section 6.3.1), the analysis was performed by simulating 600 individual experiments (12 participants x 50 studies) per condition, each starting with 150 trials of CALIB, followed by 50 trials of CORL. Each condition was determined by predefined ErrP-decoder true-negative and true-positive rates (TNR, TPR) which were fixed across participants (no other free model parameters were modified).

Guessing performance and success-rate were computed across all experiments per condition and compared against baseline measures. Baseline measures were obtained during the large-scale simulation (see Section 6.2.5.1) and are based on participant's individual online ErrP decoding performance (see Table 8).
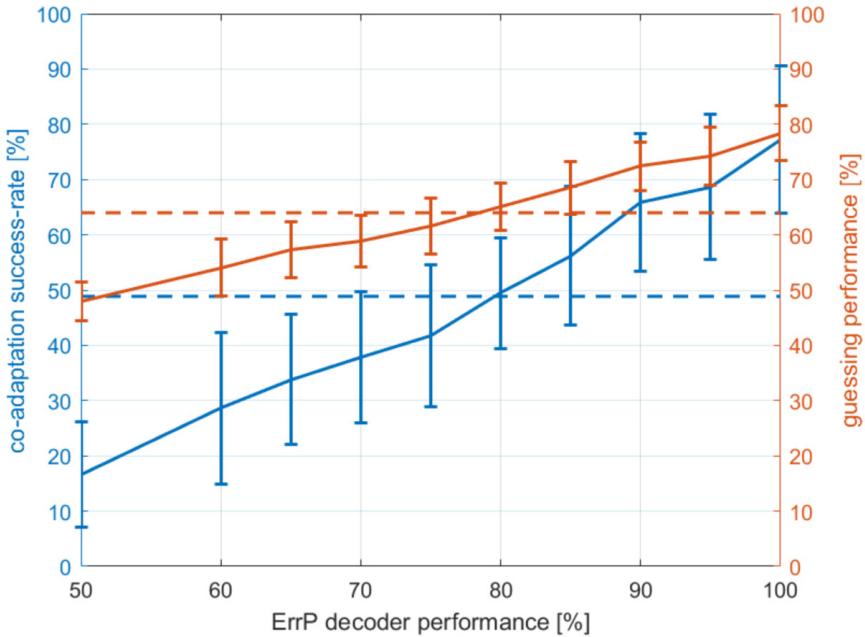
Results are depicted in Figure 34 and indicate an approximately linear relationship between ErrP-decoder performance and co-adaptation performance measures. The results suggest that even a non-functioning ErrP-decoder ($TNR = TPR = 50\%$) results in an average guessing performance of $\sim 48\%$ and an average success-rate of $\sim 16\%$. Optimal ErrP-decoder performance ($TNR = TPR = 100\%$) on the other hand, results in an average guessing performance of $\sim 78\%$ and a success-rate of $\sim 77\%$. According to the simulation, an overall improvement of approximately 10% in ErrP-decoder performance results in an approximate $\sim 5\%$ improvement of guessing performance and an approximate of $\sim 10\%$ improvement of success-rate. As hypothesized, the results reflect that ErrP-decoder performance plays a critical role, but also suggest that it is not the only factor affecting co-adaptation. Some level of co-adaptation can even be achieved with a non-functioning ErrP-decoder (possibly compensated by the human counterpart) and, more importantly, optimal ErrP decoding performance did not straightforwardly result in optimal co-adaptation performance. Further results (see Figure 34(b)) indicate that variations of co-adaptation performance are mainly driven by variations of TNR (non-error decoding rate). Variations of TPR (error decoding rate) on the other hand, have a negligible effect on co-adaptation performance. In this regard, misclassifications of error trials seem to have a weaker effect on the overall co-learning system (human and agent) than misclassifications of non-error trials. This finding is further discussed in Section 6.4. In summary this implies that efforts to improve ErrP-decoder performance, in particular, non-error decoding rates, are generally worthwhile but not the only technical factor influencing co-adaptation performance in the given scenario.

6.3.3. Simulation III: Impact of human factors

Besides ErrP-decoder performance human factors are likely to also influence co-adaptation performance. We hypothesize that specific decision-making and adaptation styles of participants, described by success- and failure learning rates and decision bound parameters, may encourage or hinder fast co-adaptation. In order to disentangle these factors, large-scale simulations were performed under two conditions: (1) with participant-specific online ErrP-decoder performance based on results from the original experiment, and (2) with online ErrP-decoder performance fixed to the average across participants of $TNR = 85.1\%$ and $TPR = 63.8\%$. By fixing ErrP-decoder performance across participants in the second condition, variations of co-adaptation performance due to human factors only, can be isolated.
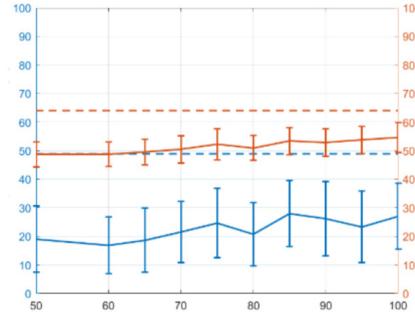
Results are depicted in Figure 35 and show the distribution of average guessing performance across different success- and failure learning rates (left panels), means and variances of decision bound (middle panels), and ErrP-decoder performances (right panel). The results of the first condition (Figure 35(a)) show an approximately linear relationship between ErrP-decoder
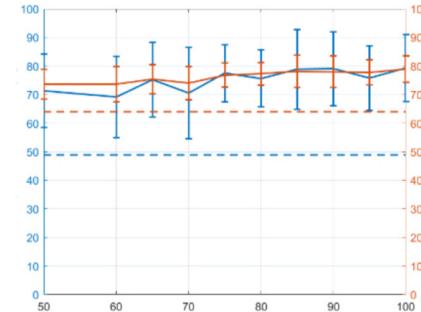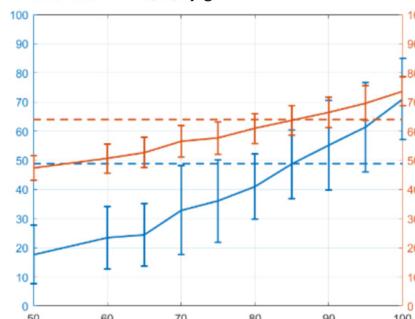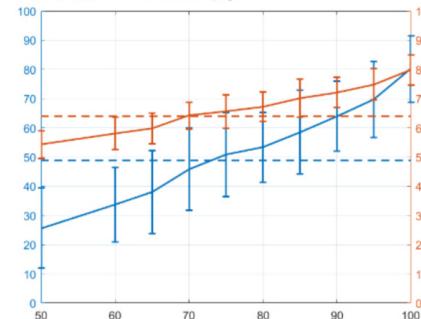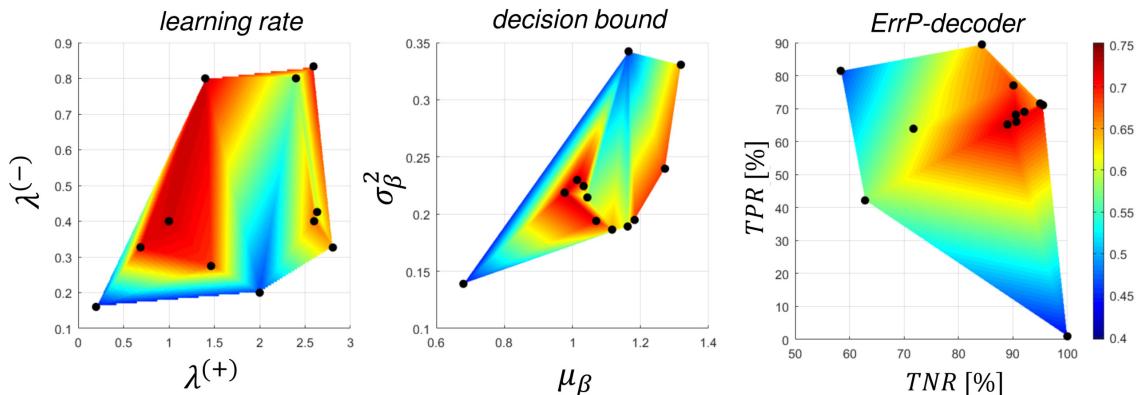
Figure 34 Simulated impact of ErrP-decoder performance: overall guessing performance (orange lines) and success-rate (blue lines) for fixed ErrP-decoder performance across participants with (a) balanced non-error and error decoding performance ($TNR = TPR$), (b) fixed TNR and varying TPR (upper panels), and fixed TPR and varying TNR (lower panels). Error bars represent the single standard deviation of the mean across 50 studies. Results indicate an approximately linear relationship between ErrP-decoder performance and co-adaptation performance measures. Variations of co-adaptation performance seem to be mainly driven by TNR; variations of TPR have a negligible effect on co-adaptation performance. Average baseline measures (computed based on individual participant's TNR and TPR) are depicted as dashed lines.

performance and average guessing performance in line with the observations of the previous simulation (see Section 6.3.2). Human factors on the other hand, do not seem to be in an obvious functional relationship with guessing performance. The results of the second condition (Figure 35(b)) show that with fixed ErrP-decoder performance, variations of average guessing performance across participants diminish drastically. Nevertheless, some dependency between variations of model parameters and co-adaptation performance can be observed: In particular, lower performance (approximately 60% versus 75%) can be observed for high values of success-learning rate, low values of failure-learning rates, together with high values for mean and variance of decision bound result. While this indicates that human factors do have an influence on co-adaptation performance as initially hypothesized, this influence seems to play a minor role in comparison to the impact observed from variations of ErrP-decoder performance. This finding suggests feasibility of generalizing the ErrP-based co-adaptation approach across participants with varying decision-making and learning behavior.

(a) subject-specific ErrP-decoder rates

(b) ErrP-decoder rates fixed across subjects
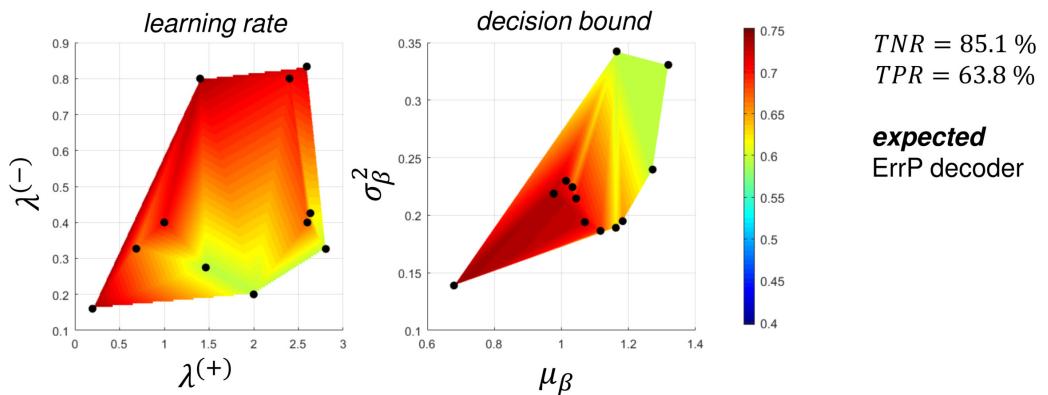
Figure 35 Human factors: Distributions of average guessing performance across variations of free model parameters with dots representing individual participants ($n = 12$). (a) Average guessing performance including participant-specific online ErrP-decoder performance and (b) average guessing performance with average (or expected) online ErrP-decoder performance fixed across participants. Results indicate a major role of ErrP-decoder performance and a minor role of human factors in co-adaptation performance in the given experimental context.

6.3.4. Simulation IV: Impact of agent learning paradigm

Besides human factors and ErrP decoding performance, the agent learning paradigm and corresponding parametrization has likely also an effect on co-adaptation performance. In

the original experiment, a policy gradient based update strategy with a fixed learning rate of $\alpha = 0.1$ was chosen (see Chapter 5, Section 5.2.3.4). The method was employed because it was reported to be capable of dealing with complex learning tasks involving many degrees of freedom [213, 214]; the learning rate was empirically adjusted in pilot experiments prior to the conductance of the experimental study. It remained unclear if the learning rate was chosen appropriately given the variability across participants. Large variations in results across participants in the experimental study suggested that an adaptation of the learning rate to individual participants could have a beneficial effect on co-adaptation performance. Furthermore, it remained unclear if policy gradient based learning constitutes the best suitable method given the particularities of this co-learning task. These are, in particular: a noisy reward signal (due to imprecise ErrP decoding) and a non-stationary reward function (reward function, implicitly provided by the participants, may undergo changes when participants revise their learning and/or decision strategy). The most competitive alternative to the policy gradient based method is possibly Q-learning, which was introduced by Watkins in 1989 [235]. Q-learning is one of the most widely used and studied methods in reinforcement learning [210] and therefore a suitable choice for systematic comparison. Instead of on-policy adaptation as in policy gradient based methods, in Q-learning, the so-called Q-function (or Q-table) is updated which represents the "quality" of an action taken in a specific state. The policy can be obtained at any time from the Q-table by using the softmax function. For realizing the Q-learning based agent, the original agent action generation and learning algorithm (see Algorithm 2) was replaced with Algorithm 3. The major differences between the two algorithms are: (1) in Q-learning, a Q-table is updated based on obtained reward, instead of the policy π (as in the policy gradient based method); the policy is computed from the Q-table using the softmax function with a temperature parameter of $\tau_{QL} = 0.1$. (2) The Q-table is updated after each action executed by the agent; as long as the participant does not respond with a decision, the Q-table is updated with a reward of $\tilde{R} = 0$. In contrast, in the policy gradient based method, π is updated only after each trial. (3) Future rewards are marginally discounted with the parameter $\gamma_{QL} = 0.99$. In the policy gradient based method, updates were not discounted. The discount parameter γ_{QL} was empirically adjusted during pilot simulations by qualitatively comparing different discount parameters ($\gamma_{QL} = \{0.7, 0.8, 0.9, 0.99, 1.0\}$). The chosen marginal discount of $\gamma_{QL} = 0.99$ resulted in the best performance in terms of quick and stable policy convergence. The softmax temperature τ_{QL} was chosen according to recommendations given by [210]. The learning rate α_{QL} was not fixed and variations investigated during the simulation-based experiments: Performance measures of simulated co-adaptation experiments were analyzed for different learning rates, for both methods. As in the previous experiments (Sections 6.3.1, 6.3.2, and 6.3.3), 50 individual experimental runs including CALIB (150 trials) and the first CORL (50 trials) were simulated per participant and condition. To realize identical starting conditions for both learning methods, the computational model was first trained during CALIB and then administered to separate CORLs, one employing the policy gradient based (Algorithm 2) and the other the Q-learning based agent (Algorithm 3).

Algorithm 3: AGENT action generation and learning algorithm (Q-learning)

Initialize Q-table with $Q(s,a) = 0$ for all s, a

for $t \leftarrow 1$ **to** T **do**

    Choose initial gaze state s_i with $s = \mathcal{U}\{s_{O1}, s_{O2}, s_{O3}, s_H\}$

    Choose goal g with $g = \mathcal{U}\{g_{O1}, g_{O2}, g_{O3}\}$

    Compute policy π from Q-table: $\pi = \dfrac{e^{Q/\tau_{QL}}}{\sum_a e^{Q/\tau_{QL}}}$

    while $\hat{g} = 0$ *("participant wait")* **do**

        Choose action a_j from state s_i using $\pi(g)$

        Execute action a_j (perform gaze shift), observe next gaze state s_j

        Update Q-table with reward $\tilde{R} = 0$:

          $Q(s_i, a_j) \leftarrow Q(s_i, a_j) + \alpha_{QL}(\tilde{R} + \gamma_{QL} \max_a Q(s_j) - Q(s_i, a_j))$

        $s_i \leftarrow s_j$

    end

    Present feedback g to participant

    Observe reward $\tilde{R}$ from ErrP-decoder

    Update Q-table: $Q(s_i, a_j) \leftarrow Q(s_i, a_j) + \alpha_{QL}(\tilde{R} + \gamma_{QL} \max_a Q(s_j) - Q(s_i, a_j))$

end

Results are depicted in Figure 36, with left panels for the policy gradient based method (Algorithm 2) and right panels for the Q-learning method (Algorithm 3). Subject average results for the policy gradient based method (Figure 36(a), upper panel) indicate a clear peak for average guessing performance and success-rate at a learning rate of $\alpha_{PG} = 0.05$. This suggests that the original choice of learning rate was close to the optimum, however, a slightly lower learning rate would have led to overall better co-adaptation performance in the experimental study. Variations of the learning rate in the Q-learning method (Figure 36(b), upper panel) on the other hand, do not exhibit a clear optimum, but rather a range ($0.3 \geq \alpha_{QL} \geq 0.7$) in which co-adaptation performance turned out superior to the policy gradient based method. The lower panels of Figure 36 show participant-individual results of average guessing performance across different learning rates. Results indicate higher performance and less variability across participants for Q-learning as compared to policy gradient based learning. Subjects exhibiting weak performance with the policy gradient based method exhibit at least decent performance with the Q-learning method (e.g. participants s06, s10, s16). This observation is further supported by Spearman's rank correlation coefficients between participant-individual optimal learning rates and computational model parameters for each learning method and co-adaptation performance measure (see Table 10). Overall, results show lower correlation coefficients for the Q-learning method compared to the policy gradient based method. These findings indicate that Q-learning does not only outperform the policy-gradient based method, but is also less dependent on varying parameters of the computational model which suggest greater independency of variations across participants. Importantly, this seems to also account for variations of ErrP-decoder performance, where Q-learning appears to be the more robust method. To complement these findings, a long-term simulation was performed for prolonged co-adaptation with 500 trials. For a fair comparison of both methods, each method's optimal learning rate across participants was chosen ($\alpha_{PG} = 0.05$ and $\alpha_{QL} = 0.5$). Results are depicted in Figure 37 and show that Q-learning significantly outperforms policy gradi-

Table 10 Spearman's rank correlation coefficients between participant individual optimal agent learning rate α and free model parameters ($\lambda^{(+)}, \lambda^{(-)}, \mu_\beta, \sigma_\beta^2, TNR, TPR$) for policy-gradient based and Q-learning separately. Asterisks (*) denote significance level at $p < 0.05$.

PG-learning	$\lambda^{(+)}$	$\lambda^{(-)}$	μ_β	σ_β^2	TNR	TPR
guess.-perf.	-.27	-.30	-.26	-.15	**.68***	.05
success-rate	-.40	-.46	-.35	.22	.22	-.18
Q-learning	$\lambda^{(+)}$	$\lambda^{(-)}$	μ_β	σ_β^2	TNR	TPR
guess.-perf.	.06	.02	.29	-.02	.00	.11
success-rate	-.33	-.37	-.07	.14	-.10	-.19

ent based learning from the first segment (trials 1-10) onwards, both in converging faster to higher average guessing performance as well as lower absolute decision times (Figure 37(a)). Further support for Q-learning resulting in more robust convergence is provided by the comparison of the median convergence horizon $t_{conv} \approx 250$ for policy gradient based learning and $t_{conv} \approx 20$ for Q-learning (Figure 37(b)).
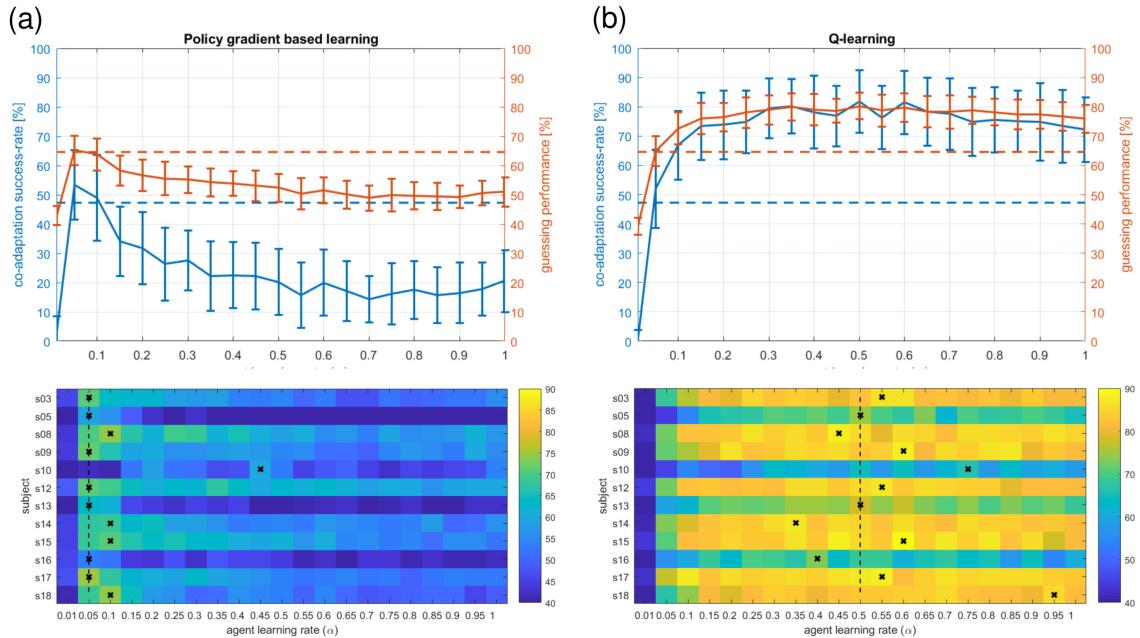
Figure 36 Comparison of alternative agent learning paradigms for different learning rates: Average guessing performance and success-rates for different learning rates comparing policy gradient based learning (a) and Q-learning (b). Across participants average results (upper panels) show superior co-adaptation performance for Q-learning, as compared to policy gradient based method across a range of learning rates. Subject-individual results (lower panels) show superior and less variant co-adaptation performance across participant. Dots represent participant-specific optimal learning rates; dashed lines show across participants optimal learning rate for each method.

6.3.5. Simulation V: Assessment of generalizability and scalability

Finally, we investigated if the ErrP-based co-adaptation ErrP-based co-adaptation approach can be generalized and scaled to higher dimensional state-spaces. This has practical implications as it gives insight on if and how the approach can be transferred to different, more
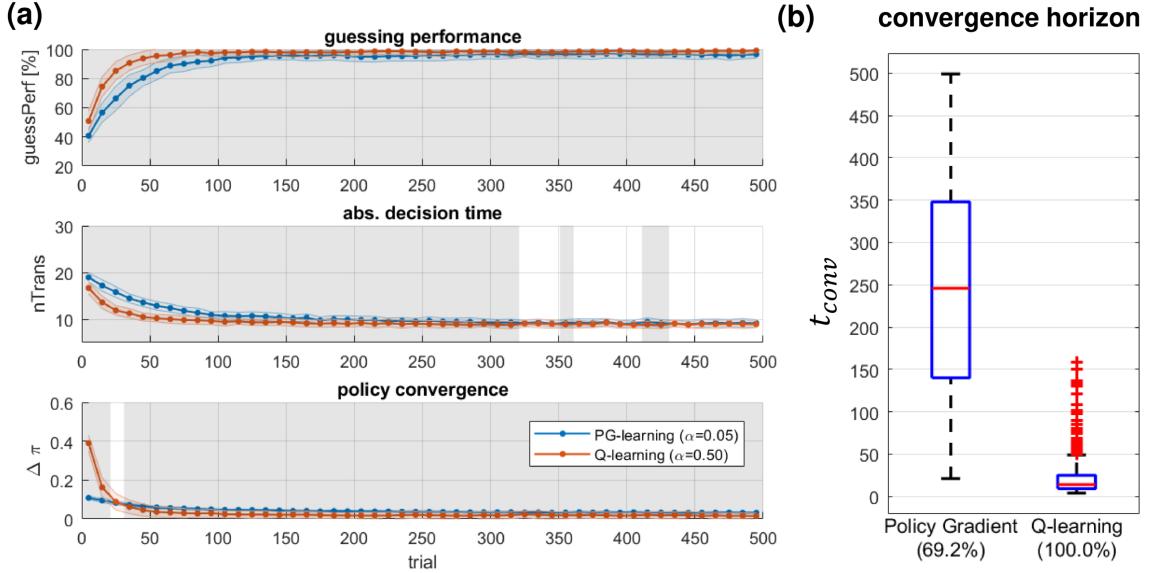
Figure 37 Comparison of long-term dynamics of alternative agent learning paradigms: (a) guessing performance, decision time (number of agent actions until decision), and policy change-rate $\Delta\pi$ of subsequent policy updates over the course of 500 trials. Shaded areas represent the single standard deviation of the mean across 50 simulated studies. Results indicate clear superiority of Q-learning over policy gradient based learning in terms of faster convergence to higher guessing performance and lower decision time. Gray shaded areas indicate significant differences between methods (Wilcoxon paired signed rank test, $p < 0.05$, FDR-corrected). Significant differences in favor for Q-learning are evident from the first segment onwards (trial 1-10). (b) Distribution of convergence horizon of individual experimental runs ($n = 600$, 12 participants x 50 studies).

complex HRI scenarios. In the experimental study, the state-space was deliberately kept relatively small to demonstrate principal feasibility of the approach. The agent behavioral policy comprised four gaze states, four possible gaze transitions (actions) per state, and three possible goals/intentions. For simplification, the state-action space was furthermore remapped depending on varying goals/intentions, which resulted in a policy with 16 degrees-of-freedom (4 states x 4 actions). This section addresses simulated extensions of the state-action-goal-space and investigates resulting performance measures of co-adaptation, in particular, policy convergence and guessing performance. As the previous section pointed out better performance of Q-learning compared to policy gradient based learning (see Section 6.3.4), all subsequent simulations were exclusively performed with the Q-learning approach. For generalization, the agent's Q-learning function was extended with the goal-dimension[10] (see Eq. (6.18)) and replaced the Q-learning function in Algorithm 3. To account for the most conservative test cases, the number of actions were always identical to the number of states (the set of possible actions in a given state always comprised transitions to all other states including staying in the current state). The dimensionality of the human computational model weights $w_{i,j,m}$ was initialized corresponding to the dimensionality of the agent's state-action-goal space (see Algorithm 1), as if "knowing" the agent's behavioral degrees-of-freedom from the beginning[11]. Subsequently, a varying number of states, actions, and goals were simulated. Starting with a

---

[10]Although, goals may be treated as just an extension of the state-space, it was introduced as another dimension in the Q-table for the sake of transparency and comprehensibility.

[11]Flexible adaptations of $w_{i,j,m}$ upon observation of new states, actions, and goals during co-adaptation are programmatically feasible and would more realistically simulate a human participant without prior knowledge about the agent's degrees-of-freedom. However, such an alternative implementation would not affect the simulation results presented here.

tangible example, the original guessing game from the experimental study was extended from initially 3 goals with 4 gaze states to 5 goals with 6 gaze states, and finally, to 10 goals with 11 gaze states. For the sake of generalization, these simulations only comprised co-adaptation experiments without preceding calibration. This ensured that both human and agent start without prior knowledge into co-adaptation, which was found to be the more conservative test case (see Section 6.3.1). As in all previous analyses, average performance measures across 600 experiments per condition are reported (12 participants, 50 studies).

Results are depicted in Figure 38(a-b): As expected, with increasing number of goals, the learning curves become flatter and policies converge on average later. However, as can be seen in Figure 38(a), average guessing performance exceeds the 95% confidence interval of chance-level performance[12] in all three test cases at some point, e.g. around trial 25 for 3 goals, around trial 40 for 5 goals, and around trial 140 for 10 goals. Figure 38(c) shows the results for an extended simulation with jointly varying state-action- (number of states = $2, 4, 6, 8, 10, 15, 20, 50$) and goal-dimensions ($M = 2, 3, 5, 7, 10, 15, 20$). This spans a range of state-space dimensionality from 8 (2 states, 2 actions, 2 goals) to 50.000 (50 states, 50 actions, 20 goals)[13]. The simulation results show that the convergence horizon t_{conv} increases logarithmically in relation to increases of the state-space dimensionality (Figure 38(c-left)) from < 50 trials for small state-spaces up to around 400 trials for the highest tested state-space dimensionality. Similarly, the guessing performance necessary to exceed on average the 95% confidence interval of chance-level (Figure 38(c-left)) ranges from < 50 trials for small up to > 500 trials for large state-spaces. The reason for guessing performance to reach low performance in the case of small state-spaces in conjunction with high number of goals is that a certain number of goals require a certain state-space dimensionality to be unambiguously de-/encodable. It is therefore logical that the computational model did not perform well in these test cases. The analysis further illustrates that state-space extensions only affect the convergence horizon in that successful co-adaptation takes up an impractically long time. The basic principle of the ErrP-based co-adaptation approach, appears to remain in turn unaffected by increases of the agent's behavioral complexity. In order to identify state-space extensions which can be considered practically acceptable (within practically acceptable timescales), we assume guessing performance should exceed the 95% confidence interval of chance-level within the first 150 trials, which corresponds to approximately 18-24 minutes of co-adaptation[14]. State-action-goal combinations which conform to this constraint

---

[12]inverse binomial test for 10 subsequent trials given the corresponding number of options/goals

[13]Because of the high number of conditions, computation time was here increased excessively. Therefore, this analysis only comprised 120 experiments per condition (12 participants, 10 studies), instead of 500 as in previous simulations. Repeated computations confirmed that results were replicable despite the lower number of experiments per condition; the observed variations across repeated computations do not affect the main conclusions drawn from this analysis.

[14]Approximation based on the average duration of a single CORL in the experimental study, see Chapter 5, Table 5.

are depicted within the red dashed frames in Figure 38(c). Based on this constraint, simulation results suggest that extensions of up to 12500 degrees-of-freedom (e.g. 50 states, 50 actions, and 5 goals) are practically acceptable.

$$Q(s_i, g_m, a_j) \leftarrow Q(s_i, g_m, a_j) + \alpha_{qL}(\tilde{R} + \gamma_{qL} \max_a Q(s_j, g_m) - Q(s_i, g_m, a_j)) \qquad (6.18)$$
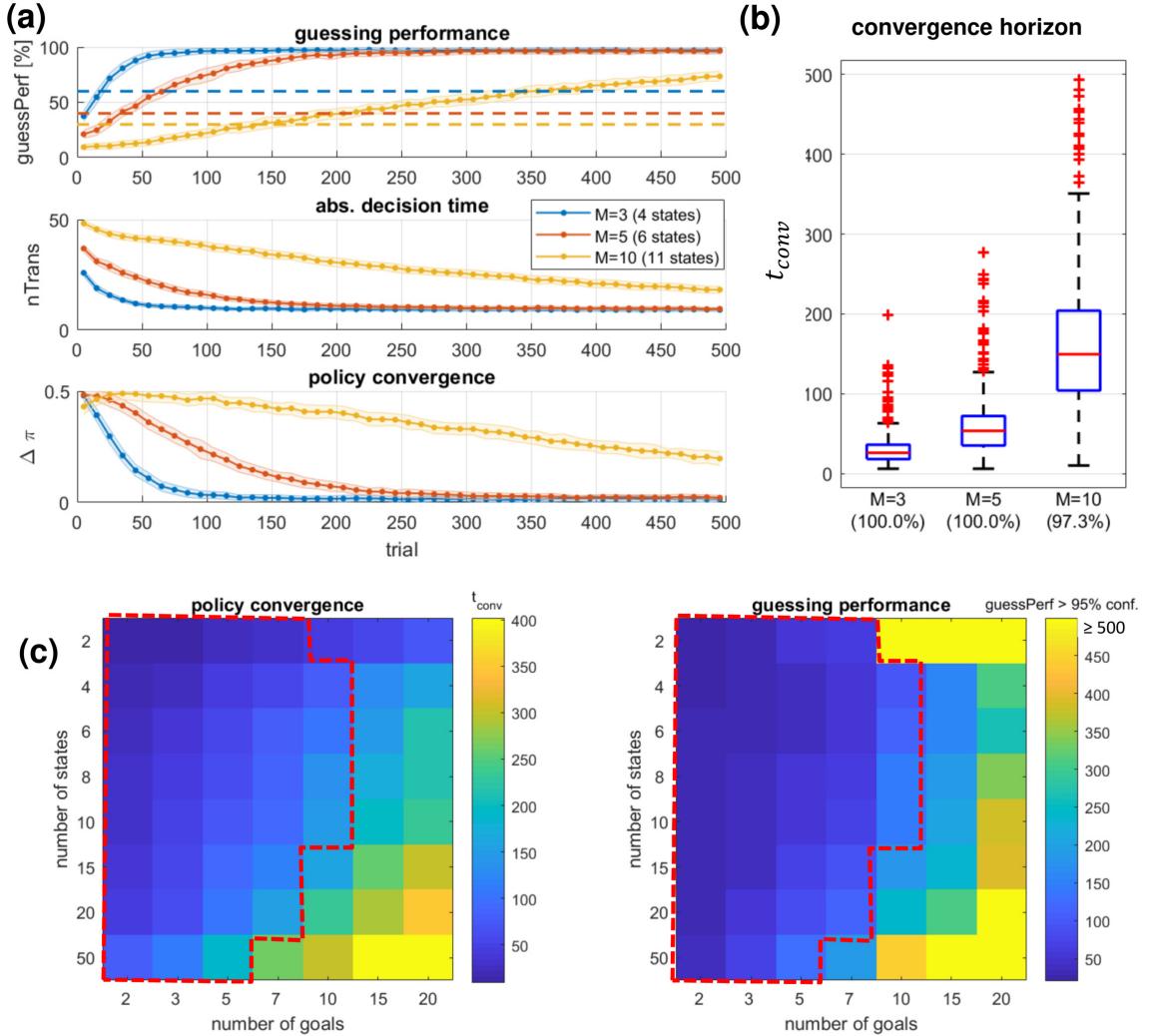
Figure 38 Scalability for extended state-action-goal space dimensionality: (a) Guessing performance, decision time and policy change rate and (b) distribution of convergence horizon over the course of 500 trials for varying number of goals M with fixed state-action-space (4 states, 4 actions per state). (c) Convergence horizon (left) and average guessing performance exceeding 95% confidence interval of chance-level for joint variations of state-space and number of goals. Red dashed frames encompass state-action-goal state extensions for which average guessing performance exceeds 95% confidence interval of chance-level within the first 150 trials.

6.3.6. Overview of simulation results

The set of simulations of co-adaptation experiments, performed using the human computational model previously fitted based on the experimental data provided a clearer picture of the impact different factors can have on the performance of the co-adaptation system. As a result, recommendations for overall improvement of performance could furthermore be derived from this knowledge: (1) Simulations showed that the convergence horizon is about threefold

higher than initially assumed. They suggested that robust convergence under the given circumstances requires about 150 policy iterations (trials). In addition to this, results indicated that, although calibration has a positive effect on the performance in subsequent co-adaptation sessions, it is not required for successful co-adaptation in the long run (see Section 6.3.1). (2) Simulations showed that from all participant-specific parameters of the human computational model, co-adaptation performance mainly depended on ErrP decoding rates (see Section 6.3.2). More specifically, the precision of non-error decoding (TNR) turned out to be more critical than the precision of error decoding (TPR). Variations of the other model parameters (learning rates and decision bound) turned out to have a negligible effect on co-adaptation performance (see Section 6.3.3). (3) Simulations suggested that the initially chosen agent learning method (policy gradient based method) is rather sensitive to variations across participants including ErrP-decoder performance. The alternative Q-learning method, tested here for comparison, unequivocally outperformed policy gradient based learning, both in terms of faster convergence to higher co-adaptation performance, as well as less dependence on participant variations (see Section 6.3.4). (4) Simulations with more complex state-action-goal spaces suggested that extensions to more than 12.5K degrees-of-freedom are practically acceptable (see Section 6.3.5). These findings are further discussed in detail in the following sections.

6.4. Discussion on the proposed computational model and simulation results

The overarching goal of the research presented in this chapter was to provide a better understanding of ErrP-based co-adaptation in human-agent interaction with the purpose to derive guiding principles for further work. Whilst the outcomes of the experimental study presented in Chapter 5 support the functionality and utility of the approach in HRI, a better fundamental understanding of the dyadic interacting human-agent system remained unexplored. The approach chosen to address this gap was (1) to develop a computational model of human decision-making behavior suitable for the given experimental setup, (2) to fit the model using the empirical data of the experimental study, (3) to integrate the model into the co-learning environment with a simulated agent, and (4) to perform large-scale simulations of co-adaptation experiments with different simulated experimental conditions. As such, this model-based approach allowed to investigate, in an efficient manner, the impact different experimental factors can have on the performance of the system studied. The simulation results obtained have important implications for the research on ErrP-based human-agent interactions and provide useful recommendations for future work.

Guidelines for the design of future studies:

An important first observation concerned the design chosen for the experimental study (see Chapter 5). Simulation-based results indicated that the timescale for successful convergence

of co-adaptation is approximately threefold higher than in the original study. This may partly explain why a significant amount of co-adaptation runs turned out unsuccessful and suggested that overall, the experimental results would have benefited from allowing longer co-adaptation. Follow up studies in line with this research should consider this finding in the experimental design. In particular, it is recommended to carefully estimate the approximate convergence horizon based on simulations prior to defining the experiment design. Although informed simulations prior to a new study may not be possible due to lack of empirical data, employing a simplified human computational model, e.g based on the one presented in this chapter, may be sufficient for an approximation of the co-adaptation convergence horizon.

Co-adaptation is mainly determined by ErrP-decoder performance and less by varying human factors:

The second finding concerned possible factors influencing co-adaptation, both human-related as well as technical factors. The simulation-based results suggested the main factor for successful co-adaptation to be the ErrP-decoder performance. This observation is consistent with the significant correlation between the overall guessing performance and the online ErrP-decoder performance observed empirically and reported in Chapter 5 (see Section 5.3.2). The likewise negligible influence of human factors (learning rates and decision bound) on successful co-adaptation was rather unexpected. Human factors were assumed to account for and help explain cases in which participants performed poorly in co-adaptation despite high ErrP-decoder performance (see Chapter 5, Section 5.4). Since the simulations did not provide evidence for such a relationship, a likely remaining explanation is that co-learning may have taken unfortunate courses in some participants due to the probabilistic nature of the agent's action generation and learning algorithm (Algorithm 2). Importantly, the negligible effect of human factors on co-adaptation performance suggests the possibility of generalizing the ErrP-based co-adaptation approach across participants, despite the variations in decision-making behavior and learning. As a result, generalization across participants is likely to be fostered by finding ways to reduce variations in ErrP decoding performance across participants, as this would contribute to improve overall predictability of the co-learning system. This could be achieved, for instance, by simplifying the calibration protocol, e.g. by employing the *cursor task* presented in Chapter 4 (see Section 4.2.2). Simplifications may reduce variations of task interpretation and as such result in more coherent ErrP responses across participants. The decoding method itself may also contribute to reduce variations across participants. One promising approach which resulted in high decoding performances around 90% across seven participants was presented recently by Kim and colleagues [43]. Simulation-based results further suggested that investing effort into improving ErrP-decoder performance, in particular, the precision of non-error decoding, is worthwhile. In the given interaction scenario, non-error decoding rates play the more crucial role than error decoding rates. This rather surprising observation could be related to participants' learning style: participants seem to have put more emphasis on success than failure trials in the process of learning the agent's behavior (suggested by higher success- than failure learning rates, see Table 8). In practice this

suggests that co-adaptation performance can be improved by biasing ErrP-decoders towards emphasizing TNR at the cost reducing TPR. A similar observation was reported by Llera and colleagues in 2011 [153], although in a complementary perspective. In their work, they explored the usability of ErrPs for online adaptation of a BCI classifier for a binary choice task. Their results indicated a more negative effect on the adaptation process resulting from incorrect decoding of ErrPs in response to the BCI output matching participants' intentions and practically no influence resulting from incorrect decoding of ErrPs in response to the BCI output mismatching participants' intentions. This supports both our results and post-hoc those reported by Llera and colleagues. Whether and to what extent this observation can be generalized across further interaction scenarios is subject to future research.

Alternative agent learning methods can significantly improve co-adaptation performance:

The most promising finding concerning possible technical modifications in favor of improved co-adaptation performance resulted from the comparison between the original policy gradient based learning with an alternative Q-learning based agent algorithms. Simulation results demonstrated Q-learning to outperform policy gradient based learning, both in terms of faster convergence to higher co-adaptation performance, as well as improved robustness against variations across participants (including ErrP-decoder variations). The reason for this may be twofold: (1) Q-learning enforces optimality by maximizing reward in the long run; the policy gradient based method on the other hand, does not take into consideration expected future rewards and updates the policy only based on information from the current iteration (trial). (2) The Q-learning algorithm (Algorithm 3) was implemented such that Q-table updates were carried out after each agent action (Section 6.3.4). Although this update was performed with a reward of $\tilde{R} = 0$ as long as no decision was taken by the participant, this nevertheless had an effect because of the reward discounting property of Q-learning. This resulted in the algorithm to not just optimize for rewards based on trial outcomes, but also to minimize the number of actions needed to complete the trial (e.g. minimizing participant decision time). What remained unexplored is the capability of both algorithms to follow possible sudden changes in participant's decision-making and learning strategy. For instance, it appears conceivable that participants may at some point decide to reject assumptions built during previous trials and switch to new or alternative assumptions. While policy gradient based learning may be capable of following such changes flexibly as it is not explicitly concerned with future reward, Q-learning on the other hand, may not be as flexible as it enforces optimality by maximizing future reward. Although the latter leads to better co-adaptation performance in the long run, the former method may more flexibly align with participants' preferences. The investigation of this question is left open for future research.

ErrP-based co-adaptation can be transferred to more complex HRI scenarios:

The finding obtained in the last simulation provided important insight into the transferability of the ErrP-based co-adaptation approach to other, more complex, HRI scenarios. Here,

co-adaptation performance was investigated in the context of simulated extensions of the state-action-goal-space. Results suggested that extensions up to 12.5K degrees-of-freedom, subsuming behavioral states, actions, and goals are practically acceptable within reasonable timescales of co-adaptation convergence. It is worth noting that acceptable timescales required for convergence are strongly dependent on the interaction scenario: the simulation results here presented (see Figure 38) shall inform future research about estimates of convergence horizon given a predefined state-action-goal-space, or vice-versa, about how to modify the state-action-goal-space to conform to a practically acceptable convergence horizon. The findings from this simulation are limited by the fact that the general framework in which these state-space extensions were investigated were based on the repeated guessing game used in the experimental study (see Chapter 5, Section 5.2.1). This constrains the transferability of results in the following ways: The model-based approach including all simulations here presented assume (1) the agent's behavior consisting of sequences of discrete actions which (2) encode the agent's underlying goals or intentions. (3) These goals / intentions are unambiguously disclosed to the human interaction partner at some point, e.g. explicitly through feedback or implicitly through a confirmative action concluding the action sequence. (4) The human interaction partner is intrinsically motivated to infer the agent's goals/intentions to adapt his/her future actions or behavioral strategy, e.g. to optimize interaction performance or subjective interaction preferences. To conform to the above constraints, it is recommended to structure future alternative interaction scenarios as game-like protocols. While this may at first seem a limiting factor for generalization to a wider range of HRI scenarios, it is worth noting that many situations arising in natural human-human interaction can be simplified to and formalized as games (the formal framework is provided by a large body of research on (non-)cooperative game theory [236, 237]). Such game-like formalization could, for instance, be applied to human-robot collaborative assembly or sorting tasks in which both interaction partners have to find agreements about assignment of subtasks to succeed in achieving a common goal. This approach has already been formally proposed, for instance in [238] and practically applied, for instance in [239].

Limitations and future work:

While the human computational model here proposed successfully captured behavioral effects in the given experimental scenario, a number of aspects may limit its applicability to alternative interaction scenarios: (1) The proposed computational model turned out to be limited to describing approximately 80% of the participants recruited for the study. The model fit turned out rather poor for three participants which let us to exclude them from the simulation. The reasons for this are likely related to factors specific to these participants which were not sufficiently captured by the model. As these factors could be related to temporary aspects, such as the participants' attention level or task interpretation, suggestions for model improvements require further research. (2) The perception and inference module of the proposed computational model determines beliefs based on event statistics (occurrence probabilities) and does not explicitly take into consideration the order of event occurrences. While this seemed to be

no limitation in the context of the given experimental paradigm, it may have an effect in the context of interaction scenarios where the order of events plasy a more critical role. In such cases, the Bayesian approach in the inference module could be replaced by an approach based on Hidden Markov Models. (3) The current realization of the computational model assumes stationary model parameters, e.g. learning rates and decision bound do not change over the course of the experiment. While this was necessary to keep the model simple, it is conceivable that these parameters actually varied throughout the experiment, e.g. depending on the participant's attention level. This could be a reason why the model parameters could not be well fitted to three out of 15 participants. A future revision of the model should therefore include a careful analysis of the stationarity of model parameters. This investigation should also account for stationarity in the context of varying experimental conditions. (4) The assessment of model complexity suggested that simplifications are possible, such as by integrating success learning rate and mean decision bound as these parameters turned out to be significantly correlated across participants. This could be implemented by deriving (or adapting) the current success learning rate $\lambda^{(+),t}$ from the decision bound of the current (μ_β^t) and/or previous trials ($\mu_\beta^{1:t-1}$). Further simplifications could be applied to the perception and inference module, e.g. computing the belief directly from the model weights $w_{i,j,m}$ (without employing Bayesian updating). This would reduce the number of necessary computations and slim down the model to be more easily employed in new scenarios.

6.5. Summary

This chapter proposed a computational model for human decision-making and learning in the context of ErrP-based mediation of human-agent co-adaptation. The model was fitted to empirical data obtained in our previous study (Chapter 5) on single-participant level and validated by comparing large-scale model-based simulations with empirical results. The proposed computational model enables the prediction of human decision-making and learning in the context of ErrP-based human-agent co-adaptation in the given HRI scenario. As such, it allows for simulation of future empirical studies, and thereby provides a means for accelerating progress along this line of research in a resource-saving manner. This was here demonstrated by employing the model in the systematic study of factors influencing co-adaptation performance by means of the simulation of experimental sessions for different conditions. The main results suggested that co-adaptation can be significantly improved and stabilized across participants by (1) employing an alternative agent learning paradigm (Q-learning) and (2) by improving ErrP-decoder performance. Our findings suggested a more critical role of non-error decoding rate than error decoding rate in the co-adaptation process; an observation which was found consistent with previous research by others. This means that, in practice, interaction performance will benefit, in particular, from improving the non-error decoding rates. Moreover, model-based simulations further indicated that extensions to more complex agent behavior are feasible. Critically, simulation results suggested that extensions to larger state-spaces in

which human and agent operate affect only convergence timescales, but not the principal functionality of ErrP-based co-adaptation approach. Within the constraints of practical timescales ($<$ 30 min), extensions of up to 12.5K degrees-of-freedom were shown feasible.

7. Conclusions and Outlook

Summary and conclusions

The development of human-centered robotic systems whose behavior is comfortable and acceptable to humans is a multidisciplinary endeavor, demanding knowledge from artificial intelligence systems, robotics, design, experimental psychology, cognitive and social sciences. The development of robotic systems for interaction with human partners has so far been driven by two approaches: offline *assessment* (experimental psychology approach) and online *adaptation* (robotics engineering approach). The two approaches have limitations, that are addressed in this thesis by drawing upon the concept of passive BCI, specifically, upon the passive decoding of ErrPs in response to the human partner's observation of robot behavior during HRI. Following this approach, we demonstrated that ErrPs can be reliably decoded from non-invasive EEG recordings of human participants during interaction with different robotic systems in a variety of HRI scenarios, and be employed to offline assess or online adapt robot's behavior according to the expectation of the human interaction partner, this way facilitating and enhancing interaction performance.

Human-robot interaction involves neuronal mechanisms of performance monitoring

The ErrP, and specifically the N200 component [140, 39], has been described to originate from a "generic mismatch detector" in the brain [98]. More specifically, previous works converged towards describing ErrPs in the context of performance monitoring as a response generically informative about mismatches of expectations formed by internal models for predicting external events and outcomes [101]. In this thesis, ErrPs were observed for all experimental studies performed, studies featuring different interaction tasks and different stimuli, from screen-based stimuli to real humanoid robotic systems. This suggests that in all cases performing the task involved the generation of a belief towards the robot, and of associated ErrPs when those beliefs are not confirmed. Previous research proposed a link between human performance monitoring and higher cognition involved in social interaction, mentalizing, and self- and other-referencing [39]. More specifically, there is increasing evidence supporting human performance monitoring as a central process involved in a variety of tasks from rudimentary sensorimotor learning (monitoring of own behavior) up to social cognition (monitoring and evaluating the internal states of others) [40]. The results here presented support this notion, in that neuronal correlates of performance monitoring mechanisms were observed for a range of tasks, from a task requiring the monitoring of simple stimulus-outcome congruency (Chapter 4) up to a scenario requiring inference of covert intentions/goals from overt behavior (Chapter 5), as well as a scenario which employed manipulation of participants' belief about the robot's internal workings (Chapter 3). In particular, the latter study showed ERP modulations in the EEG showing a decrease at a latency of 200-ms and a peak at 350-ms, both over

fronto-central recording sites. The resemblance between these spatio-temporal patterns and that of ErrPs suggests that performance monitoring mechanisms are involved in the engagement task (*contribution 1b*, Chapter 1, Section 1.4).

In addition, model-based simulations confirmed a relationship between the electrophysiological characteristics of the ErrP collected during the empirical study (Chapter 5) and the prediction error employed by the computational model during simulation of the human participant (Chapter 6). Thus, in conjunction with previous works, our empirical (Chapters 3-5) and simulation-based findings (Chapter 6) support the ErrP as an informative, reliable, and well decodable (in single-trial) neuro-cognitive measure of expectation mismatch across a wide spectrum of scenarios. In conclusion, these findings point toward the need of more prominent exploitation of the ErrP in future empirical HRI studies involving neuro-cognitive measures.

ErrPs for offline assessment of robot behavior and HRI

The first two studies (Chapter 3 and 4) demonstrated the feasibility to decode from the EEG of the human interaction partner useful information for the assessment of robot behavior and HRI. In the first study (Chapter 3), a person's intention to engage or to get engaged with the robot via gaze contact was shown to be decodable offline from the EEG of the human interaction partner with an offline accuracy of average 80.4% across 5 participants (*contribution 1a*, see Section 1.4). The results further demonstrated how differential human beliefs about the robot and the social role participants believed to have in the interaction are reflected in significant modulations of the ERP upon human-robot gaze-contact, modulations that qualitatively resembled those of an ErrP (*contribution 1b*, see Section 1.4). This finding let us to focus on ErrPs and performance monitoring as the central neurophysiological mechanism in the subsequent study (Chapter 4). Here, we demonstrated that ErrPs in response to the human observation of robot behavior are informative with regard to the correctness of performed robot actions and can be decoded in a binary fashion with an offline accuracy of average 69% across 11 participants (*contribution 2a*, Chapter 1, Section 1.4). The discrepancy observed in cross-comparing (*contribution 2b*, Chapter 1, Section 1.4) the decoding performance of ErrPs in response to robot actions as compared to procedurally equivalent screen-based cursor actions (offline accuracies of 91% across 11 participants) were a surprising result, but consistent with the observations reported by others, such as in [174]. In conclusion, the experimental results of this thesis demonstrate the feasibility of assessing robot behavior on a single-trial level (assessment of individual robot actions) by means of EEG-based measures, in particular, ErrPs. The findings on between stimuli differences of ErrP observability point towards the demand for further thorough investigation of the factors influencing/modulating ErrPs, to further consolidate implications towards their usability for assessment of robot behavior and HRI (see Outlook below).

ErrPs for online adaptation of robot behavior and HRI

Beyond the observability and decodability of ErrPs in the context of HRI, the thesis showed that such decoded ErrPs can be used to online adapt robot behavior towards human expecta-

tions, which themselves could be volatile, e.g. human adapting also to the robot (human-robot co-adaptation). Specifically, we addressed this in a closed-loop experimental setup in which participants were asked to guess from the gazing behavior of a humanoid robot which of three objects it had selected (Chapter 5). The robot's behavioral policy started with zero prior knowledge and was updated only based on online decoded ErrPs from the human interaction partner. This had an online accuracy of 81.8% across 13 participants and, crucially, while the robot's policy was updated, the human partner could as well change/adapt expectations towards the robot. Even though human-robot co-adaptation varied across participants, co-adaptation was successful for the majority of them (10 out of 13 participants), with most robot behavioral policies converging towards plausible policies (*contribution 3a*, Chapter 1, Section 1.4). The experimental study demonstrated furthermore that relatively complex robot behavior consisting of sequences of actions can be updated based on delayed feedback derived from online decoded ErrPs (after human observation of the action sequence), an important aspect in view of the development of more naturalistic HRI systems (*contribution 3b*, Chapter 1, Section 1.4). The results of this study provided first empirical evidence for the usefulness of ErrPs as a feedback signal to mediate human-robot co-adaptation.

Technical, rather than human factors, influence ErrP-based co-adaptation in HRI: guidelines for future studies and systems development

Beyond the empirical findings of the human participant study (Chapter 5), we proposed a computational model of human decision making and learning (Chapter 6). This model (*contribution 4a*, Chapter 1, Section 1.4) was subsequently employed in model-based simulations to explore the influence of potential human and technical factors on success of human-robot co-adaptation (*contribution 4b*, Chapter 1, Section 1.4): Overall, the simulation results attributed a more important role to technical factors and a minor role to human factors. This finding supports the generalization of the experimental setup across participants as long as accompanied by technical improvements. In particular, simulation results attributed different importance to ErrP-decoder true-positive and true-negative rates, with successful co-adaptation being mainly driven by non-error rather than error decoding rates. This finding suggests that future similar experimental studies may benefit from the use of biased ErrP-decoders.

The simulations further enabled the cross-comparison of different robot learning paradigms and demonstrated Q-learning to significantly outperform the algorithm used in the empirical study. Hence, while the empirical study (Chapter 5) showed convergence to be possible with a policy gradient based algorithm, according to simulations, the results of our study would have significantly benefited from using Q-learning algorithm instead (Chapter 6). Simulations also confirmed that extensions to larger state-spaces are feasible and mainly accompanied by longer convergence times. This proved, on a theoretical level, that ErrP-based human-robot co-adaptation can be extended to more complex HRI. Overall the findings of the model-based simulations demonstrated the ErrP-based co-adaptation to be a promising approach in HRI research and may enable not only the development of a new type of user-adaptive

artificial systems, but likewise provide an alternative means to study the neurophysiological mechanisms of human social cognition in relation to performance monitoring (see Outlook below).

Outlook

Considering the work developed in this thesis, general topics for future research concerning both, the utilization of ErrPs for assessment and adaptation of robotic systems, include: (1) novel decoding methods for overall improvement of ErrP decoding rates and reduction of variability across participants, (2) novel methods for reduction of calibration time towards zero-calibration, universally applicable "plug-and-play" ErrP-decoders (across participants and tasks), (3) novel methods for asynchronous decoding of ErrPs, making ErrP decoding useful in scenarios where no precise information about the timing of events is available (4) novel approaches to practical and wide-spread accessible EEG technology allowing for easy handling and reliable measurement and ErrP decoding outside the laboratory.

Over the last years, a relevant side-topic of the author's research included the development of functional prototypes of ease-of-use, inexpensive, and unobtrusive EEG systems. The prototypes and their relevance for enhanced practicality in measurement and decoding of ErrPs are reported in Appendix B.

Assessment: Identification of factors influencing observability and decodability of ErrPs in HRI

One important aspect requiring further clarification concerns the investigation of possible factors influencing ErrPs obtained in response to robot behavior. We propose a series of human participant studies targeting the following dimensions:

- ErrPs seem to be modulated by the ***degree of event characteristic of observed stimuli***, e.g. more distinct in response to event-based stimuli which are temporally well isolated versus gradually unfolding stimuli embedded in continuous trajectories [159]. The first dimension, thus, concerns the investigation of ErrPs in response to stimuli with varying levels of event-characteristic. On the experimental level, we plan to realize this by varying the speed (or torque) of certain robot actions, e.g. robot head movements in a joint attention task, robot manipulator reaching actions in a collaborative assembly task, or changes in robot facial expressions in a social interactive task. We hypothesize the strength of the ErrP to be correlated negatively with the duration of events (duration of execution of incremental robot actions).

- While it has been shown feasible to detect ErrPs in response to unexpected robot actions, it remained so far unexplored if ErrPs can be observed also in response to ***failed occurrence of expected stimuli***. Being able to decode ErrPs in the latter case is particularly important in the domain of collaborative HRI, e.g. to cover situations where the human

partner expects the robot to take an action it fails to execute. The second dimension, thus, concerns the investigation of ErrPs in response to non-occurring expected stimuli. On the experimental level, we plan to realize this in a human-robot collaborative sorting task in which the robot fails in some situations to sort specific objects it has been assigned to or does so only after an interval of time unknown to the participant. We hypothesize that ErrPs can be observed in response to non-occurrence of expected events, however with significantly lower effect strength compared to ErrPs in response to occurrence of unexpected events. The effect strength likely depends on the participant's precision of time-estimation as to when to expect the event, a variable which certainly varies within and across participants due to fluctuating levels of attention.

- A challenging dimension to investigate concerns the extent as to which ErrPs are influenced by ***human's prior expectations towards the robotic system*** [240]. Robots with human-like appearance, such as androids, can raise expectations of human-like behavior and capabilities; violations of these expectations due to slight deviations from human-like characteristics can cause odd feelings in human observers (known as the uncanny valley effect [241]). The degree of expectation violation naturally depends on the level of expectation prior to the observed violation. Therefore, a robotic system which does not raise strong human expectations may result in less distinct ErrPs. On the experimental level, we plan to evaluate this by varying levels of human likeness of a simulated robotic agent executing the same or similar actions (as in [18] or [19]). Alternatively, one could use *belief-manipulations* (as in Chapter 3), and this way manipulate expectations about the robot's goals / intentions underlying a particular behavior. Modulations of ErrPs could be derived from analyses of correlations with subjective ratings of participants' individual beliefs and expectations towards the robotic system.

- Besides the appearance and characteristics of the robotic system, the ***human's perceived importance of robot actions*** is likely to be another factor influencing ErrPs. The degree of importance of robot actions to the human interaction partner's individual and/or team success likely modulates how closely he/she would monitor the robot's behavior. This in turn likely affects the observability and decodeablity of ErrPs. On the experimental level, we plan to address this by manipulating the significance of robot actions for the human partner by varying rewards given for joint or individual task success. This requires an experimental HRI scenario, e.g. a human-robot assembly task, in which task success does not rely on collaboration, but collaboration is gradually encouraged by providing varying levels of joint reward. We hypothesize the strength of the ErrP to vary strongly on the single-trial level for robot actions with low importance for the human partner (e.g. high during phases in which the participant evaluates the importance of robot actions). Vice versa, we expect ErrPs to be more consistent on single-trial level if robot actions are consistently important for the participants' individual or team success.

Adaptation: Towards formalization of ErrP-based co-adaptation in collaborative human-agent interaction

In the co-adaptation study presented in Chapter 5 adaptation of agent behavior was based on (mis-)matching beliefs of the human participant about the agent's goals. Essentially, this resulted in the agent's behavior converging towards participants' expectation about its behavior and goals, while these expectations may as well dynamically change (adapt to the agent) throughout the experiment. The experimental data demonstrated that this approach resulted in successful co-adaptation in the majority of cases. However, the entanglement of both the agent policy and participant's expectations do not allow further post-hoc analyses about how this co-adaptation was achieved (e.g. who adapted in which phases to whom and when did the implicit agreement occur).

We propose therefore a theoretical extension of our approach to cover above considerations. Here, we draw upon the concept of "ToM (theory of mind)" according to Premack and Woodruff (1978) [242] which describes a mental process going beyond the ability to mentalize about the internal states, intentions, and goals of others. In particular, ToM proposes that in the context of cooperative or competitive interaction, the ability of mentalizing must be extended to include recursive representations of reciprocal beliefs (others' belief about our intentions; their model about our belief about their intentions, etc; see Figure 39(b)). The proposed framework is illustrated in 39(a) and describes a reciprocal interaction between two partners, e.g. a human (H) and an agent (R). Both partners operate based their own policy (π_H and π_R) describing their behavior in the respective interaction scenario (first order: *"my belief of my policy"*). Furthermore, both partners maintain an estimate of the other's policy (π'_H and π'_A) which is inferred by observing the other's behavior (second order: *"my belief of your policy"*). Lastly, both partners maintain an estimate (π''_H and π''_R) of the other's estimate of their own policy (third order: *"my belief of your belief of my policy"*). Crucially, the third order inference of the agent's estimate π''_R of the humans estimate of the agent's policy is inferred from online decoded ErrPs during human inference π'_R of the agent's policy. Adaptations of own policies involve policy estimates of second and third order inference: $\pi_H^{new} = f(\pi_H^{old}, \pi'_R, \pi''_H)$ and $\pi_R^{new} = f(\pi_R^{old}, \pi'_H, \pi''_R)$. To address this on the experimental level, we propose to start with a simplified testbed with a constrained state-space and ground truth access to all actions performed by both human and agent. This allows for reciprocal policy inference within reasonable time constraints and ensures policies are of limited complexity to facilitate data analysis and interpretation. In 2008 and 2010, Yoshida and colleagues proposed an experimental setup which constitutes a suitable testbed for our case [243, 244]. They used a variant of *Stag Hunt* (see Figure 39(c)), a game originally described by Jean-Jacques Rousseau (see also [245]). Here, two players go on a hunt and can choose between hunting a stag or a hare. Crucially, hunting the stag (high payoff) requires cooperation between the players, whereas hunting the hare (low payoff) can be performed individually. This creates a "trust dilemma" (conflict between safety and social cooperation) forcing both partners to infer the other's policy (intention to cooperate) to obtain high payoff in the long run. Yoshida and colleagues used this setup to model ToM based on behavioral data of human participants [243] as well as for the study of neural processes in-

volved in mentalizing using fMRI [244]. As a preliminary step towards validating the testbed for our research goals, we performed a pilot EEG study with five participants. Here, participants played the game with a computer agent which followed a simple deterministic decision policy. The goal of the pilot was to validate if defective agent behavior (actions deviating from the policy) would result in observable ErrPs. Indeed, the ErrP-characteristic N2-P3 complex was visible in the majority of participants (see examples of two participants in Figure 39(d)). Based on these results, we consider the testbed as a suitable experimental setup to validate the proposed theoretical extension of ErrP-based human-agent co-adaptation. Upon validation, we plan to address the following aspects in follow-up empirical studies:

- Investigation on how to perform policy updates based on functional combination of first, second and third order policies. To approach this, we will start by collecting human-human interaction data in the same testbed and modelling the behavioral data based on the approach of [243]. This will provide insight on how this functional combination is performed in human participants allowing for the derivation of a computational account implementable in the agent framework.

- Investigation of co-adaptation by joint study of human and agent policy adaptations. Here, we plan to manipulate the agent's level of adaptiveness from low (e.g. enforcing optimal team behavior on the human) to high adaptiveness (e.g. following human subjective preferences). The analysis of joint task performance in relation to low versus high agent adaptiveness allows then the quantification of how much the human participant adapted to the agent.

- Validation of the extended ErrP-based human-agent co-adaptation approach in the context of agents with limited perceptual abilities, e.g. unable to infer reliable estimates π'_H of the human interaction partner's policy. This simulates the case of a major limitation in real-world HRI, where often reliable modeling of the human partner is infeasible due to limited or noisy robot perception. The goal of this investigation is to validate if comparable co-adaptation performance can be realized when agents update their policy without a policy estimate of the human partner: $\pi_R^{new} = f(\pi_R^{old}, \pi''_R)$. This would significantly facilitate co-adaptation in real-world HRI, by circumventing the need for robot inference of human policies.
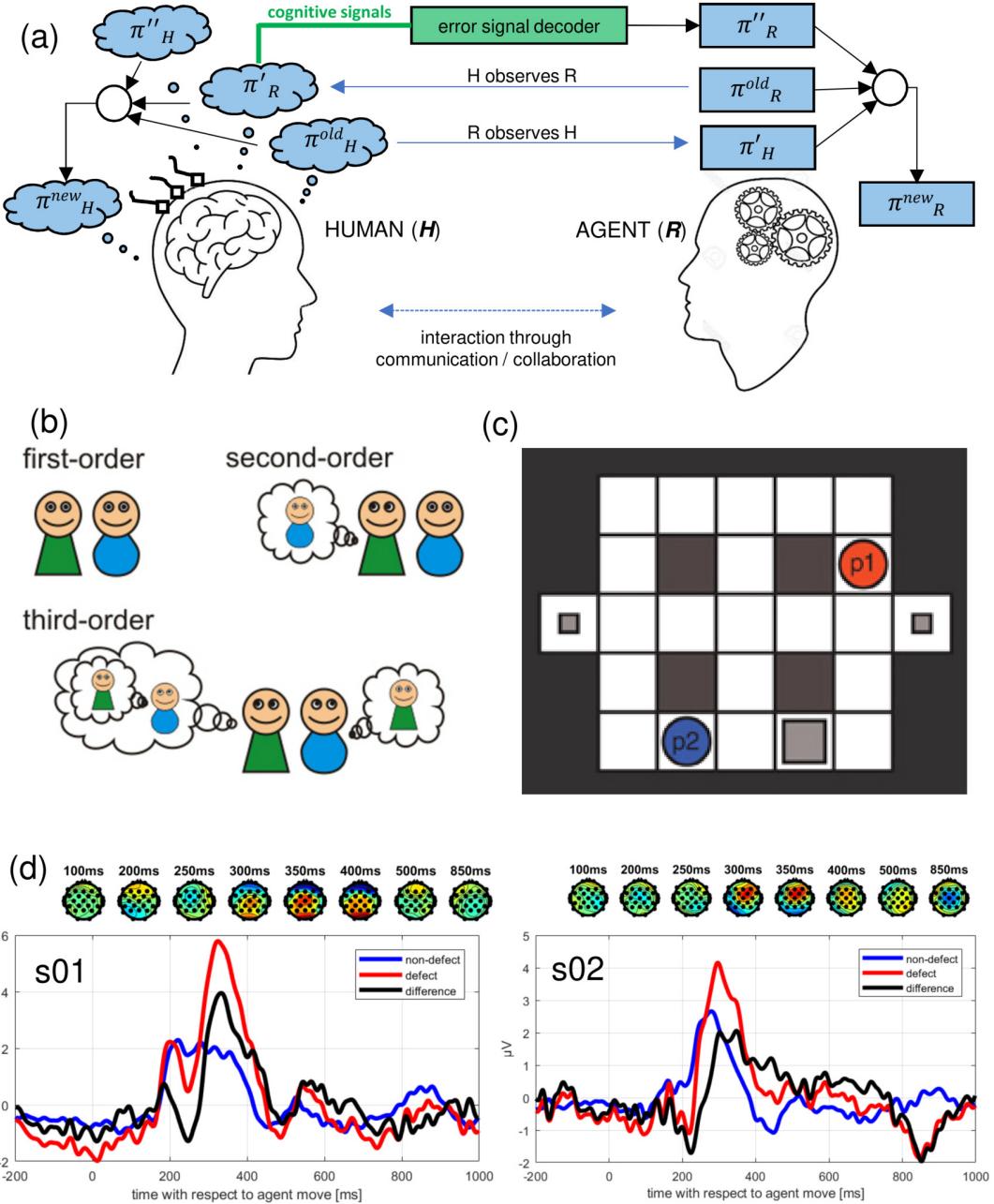
Figure 39 Formalization of ErrP-based co-adaptation: (a) Extended framework for human-agent co-adaptation using error-related potentials. Both interaction partners operate based their own policies (first order: π_H and π_R), estimates of the other's policy (second order: π'_H and π'_A), as well an estimate of other's estimate of their own policy (third order: π''_H and π''_A). Agent third order inference is performed by inferring from online decoded ErrPs. (b) First-, second-, and third-order inference of mentalizing about the internal states, intentions, and goals of others. (c) Variant of *Stag Hunt*, a famous game originally described by Jean-Jacques Rousseau as a suitable testbed for the extended framework for human-agent co-adaptation using error-related potentials. Players p1 and p2 can choose between collaboratively hunting a moving stag (big grey square) or individually hunt static hares (small grey squares) [B and C, reprinted from Yoshida et al. 2008 and 2010 [243, 244]]. (d) Average time-courses of event-related potentials in response to non-defective (blue) and defective (red) agent moves in two participants. Black line shows the difference wave (defect minus non-defect) with the characteristic N2-P3 complex, also visible in the topographic representation at specific time-points.

Bibliography

[1] Chris D Frith and Uta Frith. How we predict what other people are going to do. *Brain Research*, 1079(1):36–46, 2006.

[2] Michael Tomasello, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5):675–691, 2005.

[3] Michael Tomasello and Malinda Carpenter. Shared intentionality. *Developmental Science*, 10(1):121–125, 2007.

[4] Bradley Hayes and Brian Scassellati. Challenges in shared-environment human-robot collaboration. *Learning*, 8(9), 2013.

[5] Agnieszka Wykowska, Thierry Chaminade, and Gordon Cheng. Embodied artificial agents for understanding human social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693):20150375, 2016.

[6] Eva Wiese, Giorgio Metta, and Agnieszka Wykowska. Robots as intentional agents: using neuroscientific methods to make robots appear more social. *Frontiers in Psychology*, 8:1663, 2017.

[7] Thomas B Sheridan. Human–robot interaction: status and challenges. *Human Factors*, 58(4):525–532, 2016.

[8] Aaron Steinfeld, Terrence Fong, David Kaber, Michael Lewis, Jean Scholtz, Alan Schultz, and Michael Goodrich. Common metrics for human-robot interaction. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*, pages 33–40. ACM, 2006.

[9] Astrid Weiss, Judith Igelsböck, Manfred Tscheligi, Andrea Bauer, Kolja Kühnlenz, Dirk Wollherr, and Martin Buss. Robots asking for directions: the willingness of passers-by to support robots. In *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction*, pages 23–30. IEEE Press, 2010.

[10] Agnieszka Wykowska, Jasmin Kajopoulos, Miguel Obando-Leitón, Sushil Singh Chauhan, John-John Cabibihan, and Gordon Cheng. Humans are well tuned to detecting agents among non-agents: examining the sensitivity of human perception to behavioral characteristics of intentional systems. *International Journal of Social Robotics*, 7(5):767–781, 2015.

[11] Cindy L Bethel and Robin R Murphy. Review of human studies methods in HRI and recommendations. *International Journal of Social Robotics*, 2(4):347–359, 2010.

[12] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1(1):71–81, 2009.

[13] Colleen M Carpinella, Alisa B Wyman, Michael A Perez, and Steven J Stroessner. The robotic social attributes scale (rosas): Development and validation. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pages 254–262. ACM, 2017.

[14] Astrid Weiss, Regina Bernhaupt, Michael Lankes, and Manfred Tscheligi. The USUS evaluation framework for human-robot interaction. In *AISB2009: Proceedings of the Symposium on new Frontiers in Human-Robot Interaction*, volume 4, pages 11–26, 2009.

[15] Sandra G Hart and Lowell E Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology*, volume 52, pages 139–183. Elsevier, 1988.

[16] Pablo Lanillos, João Filipe Ferreira, and Jorge Dias. Evaluating the influence of automatic attentional mechanisms in human-robot interaction. In *Workshop: A bridge between Robotics and Neuroscience Workshop in Human-Robot Interaction, 9th ACM/IEEE International Conference on, Bielefeld, Germany, March 2014.*, 2014.

[17] Agnieszka Wykowska and Anna Schubö. Action intentions modulate allocation of visual attention: electrophysiological evidence. *Frontiers in Psychology*, 3:379, 2012.

[18] Ayse Pinar Saygin, Thierry Chaminade, Hiroshi Ishiguro, Jon Driver, and Chris Frith. The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social Cognitive and Affective Neuroscience*, 7(4):413–422, 2011.

[19] Thierry Chaminade, Delphine Rosset, David Da Fonseca, Bruno Nazarian, Ewald Lutscher, Gordon Cheng, and Christine Deruelle. How do we think machines think? An fMRI study of alleged competition with an artificial intelligence. *Frontiers in Human Neuroscience*, 6:103, 2012.

[20] Hiroshi Ishiguro. The future life supported by interactive humanoid. In *2015 Transducers-2015 18th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS)*, pages 7–10. IEEE, 2015.

[21] Gonçalo S Martins, Luís Santos, and Jorge Dias. User-adaptive interaction in social robots: A survey focusing on non-physical interaction. *International Journal of Social Robotics*, 11(1):185–205, 2019.

[22] Anthony F Norcio and Jaki Stanley. Adaptive human-computer interfaces: A literature survey and perspective. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(2):399–408, 1989.

[23] Julie Shah, James Wiken, Brian Williams, and Cynthia Breazeal. Improved human-robot team performance using chaski, a human-inspired plan execution system. In *Proceedings of the 6th International Conference on Human-Robot Interaction*, pages 29–36. ACM, 2011.

[24] Cynthia Breazeal, Jesse Gray, and Matt Berlin. An embodied cognition approach to mindreading skills for socially intelligent robots. *The International Journal of Robotics Research*, 28(5):656–680, 2009.

[25] Stefanos Nikolaidis and Julie Shah. Human-robot cross-training: computational formulation, modeling and evaluation of a human team training strategy. In *Proceedings of the 8th ACM/IEEE International Conference on Human Robot Interaction*, pages 33–40. IEEE, 2013.

[26] Thomas Cederborg. Artificial learners adopting normative conventions from human teachers. *Paladyn, Journal of Behavioral Robotics*, 8(1):70–99, 2017.

[27] Dorsa Sadigh, S Shankar Sastry, Sanjit A Seshia, and Anca Dragan. Information gathering actions over human internal state. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 66–73. IEEE, 2016.

[28] Stefanos Nikolaidis, Swaprava Nath, Ariel D Procaccia, and Siddhartha Srinivasa. Game-theoretic modeling of human adaptation in human-robot collaboration. In *Proceedings of the 12th ACM/IEEE International Conference on Human-Robot Interaction*, pages 323–331. IEEE, 2017.

[29] Stefanos Nikolaidis, Anton Kuznetsov, David Hsu, and Siddharta Srinivasa. Formalizing human-robot mutual adaptation: A bounded memory model. In *Proceedings of the 11th ACM/IEEE International Conference on Human Robot Interaction*, pages 75–82. IEEE, 2016.

[30] Stefanos Nikolaidis, David Hsu, and Siddhartha Srinivasa. Human-robot mutual adaptation in collaborative tasks: Models and experiments. *The International Journal of Robotics Research*, 36(5-7):618–634, 2017.

[31] Barbara Gonsior, Stefan Sosnowski, Christoph Mayer, Jürgen Blume, Bernd Radig, Dirk Wollherr, and Kolja Kühnlenz. Improving aspects of empathy and subjective performance for HRI through mirroring facial expressions. In *2011 RO-MAN*, pages 350–356. IEEE, 2011.

[32] Barbara Kühnlenz, Stefan Sosnowski, Malte Buß, Dirk Wollherr, Kolja Kühnlenz, and Martin Buss. Increasing helpfulness towards a robot by emotional adaption to the user. *International Journal of Social Robotics*, 5(4):457–476, 2013.

[33] Maha Salem, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. Would you trust a (faulty) robot?: Effects of error, task type and personality on human-robot cooperation and trust. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 141–148. ACM, 2015.

[34] Stefanos Nikolaidis, Jodi Forlizzi, David Hsu, Julie Shah, and Siddhartha Srinivasa. Mathematical models of adaptation in human-robot collaboration. *arXiv preprint arXiv:1707.02586*, 2017.

[35] Thorsten O Zander, Christian Kothe, Sabine Jatzev, and Matti Gaertner. Enhancing human-computer interaction with input from active and passive brain-computer interfaces. In *Brain-Computer Interfaces*, pages 181–199. Springer, 2010.

[36] Thorsten O Zander and Christian Kothe. Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of Neural Engineering*, 8(2):025005, 2011.

[37] Jan BF van Erp, Anne-Marie Brouwer, and Thorsten O Zander. Using neurophysiological signals that reflect cognitive or affective state. *Frontiers in Neuroscience*, 9:193, 2015.

[38] Benjamin Blankertz, Laura Acqualagna, Sven Dähne, Stefan Haufe, Matthias Schultze-Kraft, Irene Sturm, Marija Ušćumlic, Markus A Wenzel, Gabriel Curio, and Klaus-Robert Müller. The berlin brain-computer interface: progress beyond communication and control. *Frontiers in Neuroscience*, 10:530, 2016.

[39] Markus Ullsperger, Claudia Danielmeier, and Gerhard Jocham. Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, 94(1):35–79, 2014.

[40] David M Amodio and Chris D Frith. Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7(4):268, 2006.

[41] Iñaki Iturrate, Luis Montesano, and Javier Minguez. Robot reinforcement learning using EEG-based reward signals. In *2010 IEEE International Conference on Robotics and Automation*, pages 4822–4829. IEEE, 2010.

[42] Stefan Ehrlich and Gordon Cheng. A neuro-based method for detecting context-dependent erroneous robot action. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots*, pages 477–482. IEEE, 2016.

[43] Su Kyoung Kim, Elsa Andrea Kirchner, Arne Stefes, and Frank Kirchner. Intrinsic interactive reinforcement learning–using error-related potentials for real world human-robot interaction. *Scientific Reports*, 7(1):17562, 2017.

[44] Stefan Ehrlich, Agnieszka Wykowska, Karinne Ramirez-Amaro, and Gordon Cheng. When to engage in interaction—And how? EEG-based enhancement of robot's ability to sense social signals in HRI. In *2014 IEEE-RAS International Conference on Humanoid Robots*, pages 1104–1109. IEEE, 2014.

[45] Stefan K Ehrlich and Gordon Cheng. A Feasibility Study for Validating Robot Actions Using EEG-Based Error-Related Potentials. *International Journal of Social Robotics*, 11(2):271–283, 2019.

[46] Stefan Ehrlich and Gordon Cheng. Human-agent co-adaptation using error-related potentials. *Journal of Neural Engineering*, 2018.

[47] Stefan K Ehrlich and Gordon Cheng. A computational model of human decision making and learning for assessment of co-adaptation in neuro-adaptive human-robot interaction. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 264–271. IEEE, 2019.

[48] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767–791, 2002.

[49] Clemens Brunner, Niels Birbaumer, Benjamin Blankertz, Christoph Guger, Andrea Kübler, Donatella Mattia, José del R Millán, Felip Miralles, Anton Nijholt, Eloy Opisso, et al. BNCI Horizon 2020: towards a roadmap for the BCI community. *Brain-Computer Interfaces*, 2(1):1–10, 2015.

[50] Hans Berger. Über das elektrenkephalogramm des menschen. *DMW-Deutsche Medizinische Wochenschrift*, 60(51):1947–1949, 1934.

[51] Jacques J Vidal. Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering*, 2(1):157–180, 1973.

[52] Mikhail A Lebedev and Miguel AL Nicolelis. Brain–machine interfaces: past, present and future. *TRENDS in Neurosciences*, 29(9):536–546, 2006.

[53] Janis J Daly and Jonathan R Wolpaw. Brain–computer interfaces in neurological rehabilitation. *The Lancet Neurology*, 7(11):1032–1043, 2008.

[54] Jonathan Wolpaw and Elizabeth Winter Wolpaw. *Brain-computer interfaces: principles and practice.* OUP USA, 2012.

[55] Han-Jeong Hwang, Soyoun Kim, Soobeom Choi, and Chang-Hwan Im. EEG-based brain-computer interfaces: a thorough literature survey. *International Journal of Human-Computer Interaction*, 29(12):814–826, 2013.

[56] Reza Abiri, Soheil Borhani, Eric W Sellers, Yang Jiang, and Xiaopeng Zhao. A comprehensive review of EEG-based brain-computer interface paradigms. *Journal of Neural Engineering*, 2018.

[57] Johan Wessberg, Christopher R Stambaugh, Jerald D Kralik, Pamela D Beck, Mark Laubach, John K Chapin, Jung Kim, S James Biggs, Mandayam A Srinivasan, and Miguel AL Nicolelis. Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature*, 408(6810):361, 2000.

[58] Jose M Carmena, Mikhail A Lebedev, Roy E Crist, Joseph E O'Doherty, David M Santucci, Dragan F Dimitrov, Parag G Patil, Craig S Henriquez, and Miguel AL Nicolelis. Learning to control a brain–machine interface for reaching and grasping by primates. *PLoS Biology*, 1(2):e42, 2003.

[59] Mijail D Serruya, Nicholas G Hatsopoulos, Liam Paninski, Matthew R Fellows, and John P Donoghue. Brain-machine interface: Instant neural control of a movement signal. *Nature*, 416(6877):141, 2002.

[60] Dawn M Taylor, Stephen I Helms Tillery, and Andrew B Schwartz. Direct cortical control of 3D neuroprosthetic devices. *Science*, 296(5574):1829–1832, 2002.

[61] Ranganatha Sitaram, Tomas Ros, Luke Stoeckel, Sven Haller, Frank Scharnowski, Jarrod Lewis-Peacock, Nikolaus Weiskopf, Maria Laura Blefari, Mohit Rana, Ethan Oblak, et al. Closed-loop brain training: the science of neurofeedback. *Nature Reviews Neuroscience*, 18(2):86, 2017.

[62] Kai Keng Ang, Cuntai Guan, Karen Sui Geok Chua, Beng Ti Ang, Christopher Kuah, Chuanchu Wang, Kok Soon Phua, Zheng Yang Chin, and Haihong Zhang. Clinical study of neurorehabilitation in stroke using eeg-based motor imagery brain-computer interface with robotic feedback. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 5549–5552. IEEE, 2010.

[63] Janis J Daly, Roger Cheng, Jean Rogers, Krisanne Litinas, Kenneth Hrovat, and Mark Dohring. Feasibility of a new application of noninvasive brain computer interface (BCI): a case study of training for recovery of volitional motor control after stroke. *Journal of Neurologic Physical Therapy*, 33(4):203–211, 2009.

[64] Jan Van Erp, Fabien Lotte, and Michael Tangermann. Brain-computer interfaces: beyond medical applications. *Computer*, 45(4):26–34, 2012.

[65] Kiel Gilleade, Alan Dix, and Jen Allanson. Affective videogames and modes of affective gaming: assist me, challenge me, emote me. *DiGRA 2005: Changing Views–Worlds in Play.*, 2005.

[66] Stephen H Fairclough. BCI and physiological computing for computer games: Differences, similarities & intuitive control. In *Proceedings of CHI'08*, 2008.

[67] Daniel Szafir and Bilge Mutlu. Pay attention!: Designing adaptive agents that monitor and improve user engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 11–20. ACM, 2012.

[68] Alice F Jackson and Donald J Bolger. The neurophysiological bases of EEG and EEG measurement: A review for the rest of us. *Psychophysiology*, 51(11):1061–1071, 2014.

[69] Ali Mazaheri and Ole Jensen. Rhythmic pulsing: linking ongoing brain activity with evoked responses. *Frontiers in Human Neuroscience*, 4:177, 2010.

[70] Hans H Kornhuber and Lüder Deecke. Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, 284(1):1–17, 1965.

[71] Niels Birbaumer. Slow cortical potentials: plasticity, operant control, and behavioral effects. *The Neuroscientist*, 5(2):74–78, 1999.

[72] Lawrence M Ward. Synchronous neural oscillations and cognitive processes. *Trends in Cognitive Sciences*, 7(12):553–559, 2003.

[73] Gert Pfurtscheller, Clemens Brunner, Alois Schlögl, and FH Lopes Da Silva. Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks. *NeuroImage*, 31(1):153–159, 2006.

[74] Gert Pfurtscheller. EEG event-related desynchronization (ERD) and synchronization (ERS). *Electroencephalography and Clinical Neurophysiology*, 1(103):26, 1997.

[75] Andrea Kübler, Femke Nijboer, Jürgen Mellinger, Theresa M Vaughan, Hannelore Pawelzik, Gerwin Schalk, Dennis J McFarland, Niels Birbaumer, and Jonathan R Wolpaw. Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology*, 64(10):1775–1777, 2005.

[76] Chris Berka, Daniel J Levendowski, Michelle N Lumicao, Alan Yau, Gene Davis, Vladimir T Zivkovic, Richard E Olmstead, Patrice D Tremoulet, and Patrick L Craven. EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine*, 78(5):B231–B244, 2007.

[77] Erich E Sutter. The visual evoked response as a communication channel. In *Proceedings of the IEEE Symposium on Biosensors*, pages 95–100, 1984.

[78] Steven J Luck, Geoffrey F Woodman, and Edward K Vogel. Event-related potential studies of attention. *Trends in Cognitive Sciences*, 4(11):432–440, 2000.

[79] PM Rabbitt. Errors and error correction in choice-response tasks. *Journal of Experimental Psychology*, 71(2):264, 1966.

[80] NP Bechtereva and VB Gretchin. Physiological foundations of mental activity. In *International Review of Neurobiology*, volume 11, pages 329–352. Elsevier, 1969.

[81] William H Alexander and Joshua W Brown. Computational models of performance monitoring and cognitive control. *Topics in Cognitive Science*, 2(4):658–677, 2010.

[82] Eliana Vassena, Clay B Holroyd, and William H Alexander. Computational models of anterior cingulate cortex: At the crossroads between prediction and effort. *Frontiers in Neuroscience*, 11:316, 2017.

[83] Massimo Silvetti, William Alexander, Tom Verguts, and Joshua W Brown. From conflict management to reward-based decision making: actors and critics in primate medial frontal cortex. *Neuroscience & Biobehavioral Reviews*, 46:44–57, 2014.

[84] Arthur R Jensen. *Clocking the mind: Mental chronometry and individual differences*. Elsevier, 2006.

[85] Barbara A Eriksen and Charles W Eriksen. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1):143–149, 1974.

[86] J Ridley Stroop. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6):643, 1935.

[87] Antoine Bechara, Antonio R Damasio, Hanna Damasio, and Steven W Anderson. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50(1-3):7–15, 1994.

[88] Leonie Koban and Gilles Pourtois. Brain systems underlying the affective and social monitoring of actions: an integrative review. *Neuroscience & Biobehavioral Reviews*, 46:71–84, 2014.

[89] Giorgio Coricelli, Hugo D Critchley, Mateus Joffily, John P O'Doherty, Angela Sirigu, and Raymond J Dolan. Regret and its avoidance: a neuroimaging study of choice behavior. *Nature Neuroscience*, 8(9):1255, 2005.

[90] William J Gehring, Brian Goss, Michael GH Coles, David E Meyer, and Emanuel Donchin. A neural system for error detection and compensation. *Psychological Science*, 4(6):385–390, 1993.

[91] Cameron S Carter, Todd S Braver, Deanna M Barch, Matthew M Botvinick, Douglas Noll, and Jonathan D Cohen. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, 280(5364):747–749, 1998.

[92] Michael Falkenstein, Jörg Hoormann, Stefan Christ, and Joachim Hohnsbein. ERP components on reaction errors and their functional significance: a tutorial. *Biological Psychology*, 51(2-3):87–107, 2000.

[93] Clay B Holroyd and Michael GH Coles. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109(4):679, 2002.

[94] Deanna M Barch, Todd S Braver, Erbil Akbudak, Tom Conturo, John Ollinger, and Avraham Snyder. Anterior cingulate cortex and response conflict: effects of response modality and processing domain. *Cerebral Cortex*, 11(9):837–848, 2001.

[95] Jonathan R Folstein and Cyma Van Petten. Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, 45(1):152–170, 2008.

[96] Nick Yeung, Matthew M Botvinick, and Jonathan D Cohen. The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological Review*, 111(4):931, 2004.

[97] Nick Yeung, Rafal Bogacz, Clay B Holroyd, Sander Nieuwenhuis, and Jonathan D Cohen. Theta phase resetting and the error-related negativity. *Psychophysiology*, 44(1):39–49, 2007.

[98] Wolfgang HR Miltner, Christoph H Braun, and Michael GH Coles. Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, 9(6):788–798, 1997.

[99] Claudia Danielmeier, Jan R Wessel, Marco Steinhauser, and Markus Ullsperger. Modulation of the error-related negativity by response conflict. *Psychophysiology*, 46(6):1288–1298, 2009.

[100] Wolfram Schultz. Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.*, 30:259–288, 2007.

[101] William H Alexander and Joshua W Brown. Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, 14(10):1338, 2011.

[102] Pierre W Ferrez and José del R Millán. Error-related EEG potentials generated during simulated brain–computer interaction. *IEEE Transactions on Biomedical Engineering*, 55(3):923–929, 2008.

[103] Ricardo Chavarriaga, Aleksander Sobolewski, and José del R Millán. Errare machinale est: the use of error-related potentials in brain-machine interfaces. *Frontiers in Neuroscience*, 8:208, 2014.

[104] Martin Spüler and Christian Niethammer. Error-related potentials during continuous feedback: using EEG to detect errors of different type and severity. *Frontiers in Human Neuroscience*, 9:155, 2015.

[105] Iñaki Iturrate, Ricardo Chavarriaga, Luis Montesano, Javier Minguez, and José del R Millán. Teaching brain-machine interfaces as an alternative paradigm to neuroprosthetics control. *Scientific Reports*, 5:13893, 2015.

[106] Michael Falkenstein, Joachim Hohnsbein, Jörg Hoormann, and Ludger Blanke. Effects of crossmodal divided attention on late ERP components. II. Error processing in choice

reaction tasks. *Electroencephalography and Clinical Neurophysiology*, 78(6):447–455, 1991.

[107] Edith Kaan, Anthony Harris, Edward Gibson, and Phillip Holcomb. The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15(2):159–201, 2000.

[108] Marieke Van Herten, Herman HJ Kolk, and Dorothee J Chwilla. An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, 22(2):241–255, 2005.

[109] Theo OJ Gruendler, Markus Ullsperger, and René J Huster. Event-related potential correlates of performance-monitoring in a lateralized time-estimation task. *PloS one*, 6(10):e25591, 2011.

[110] Roy F Baumeister and Mark R Leary. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3):497, 1995.

[111] Michael Tomasello et al. *Why we cooperate*, volume 206. MIT Press Cambridge, MA, 2009.

[112] Simon Baron-Cohen. *Mindblindness: An essay on autism and theory of mind*. MIT press, 1997.

[113] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3):143–166, 2003.

[114] Chrystopher L Nehaniv. *Imitation and Social Learning in Robots, Humans and Animals*. Cambridge Univ. Press, 2007.

[115] Kate C Ewing, Stephen H Fairclough, and Kiel Gilleade. Evaluation of an adaptive game that uses EEG measures validated during the design process as inputs to a biocybernetic loop. *Frontiers in Human Neuroscience*, 10:223, 2016.

[116] Anne-Marie Brouwer, Maarten A Hogervorst, Jan BF Van Erp, Tobias Heffelaar, Patrick H Zimmerman, and Robert Oostenveld. Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of Neural Engineering*, 9(4):045008, 2012.

[117] Domen Novak, Benjamin Beyeler, Ximena Omlin, and Robert Riener. Passive brain-computer interfaces for robot-assisted rehabilitation. In *Brain-Computer Interface Research*, pages 73–95. Springer, 2014.

[118] Thibault Gateau, Hasan Ayaz, and Frédéric Dehais. In silico versus 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*, 12:187, 2018.

[119] Thorsten O Zander and Sabine Jatzev. Detecting affective covert user states with passive brain-computer interfaces. In *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pages 1–9. IEEE, 2009.

[120] Megan Strait and Matthias Scheutz. Using near infrared spectroscopy to index temporal changes in affect in realistic human-robot interactions. In *PhyCS*, pages 385–392, 2014.

[121] Christian Mühl, Brendan Allison, Anton Nijholt, and Guillaume Chanel. A survey of affective brain computer interfaces: principles, state-of-the-art, and challenges. *Brain-Computer Interfaces*, 1(2):66–84, 2014.

[122] Riccardo Berta, Francesco Bellotti, Alessandro De Gloria, Danu Pranantha, and Carlotta Schatten. Electroencephalogram and physiological signal analysis for assessing flow in games. *IEEE Transactions on Computational Intelligence and AI in Games*, 5(2):164–175, 2013.

[123] Daniel Szafir and Bilge Mutlu. Artful: adaptive review technology for flipped learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1001–1010. ACM, 2013.

[124] Leonhard Schilbach, Marcus Wilms, Simon B Eickhoff, Sandro Romanzetti, Ralf Tepest, Gary Bente, N Jon Shah, Gereon R Fink, and Kai Vogeley. Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *Journal of Cognitive Neuroscience*, 22(12):2702–2715, 2010.

[125] Giorgio Metta, Lorenzo Natale, Francesco Nori, Giulio Sandini, David Vernon, Luciano Fadiga, Claes Von Hofsten, Kerstin Rosander, Manuel Lopes, José Santos-Victor, et al. The iCub humanoid robot: An open-systems platform for research in cognitive development. *Neural Networks*, 23(8-9):1125–1134, 2010.

[126] Giorgio Metta, Paul Fitzpatrick, and Lorenzo Natale. YARP: yet another robot platform. *International Journal on Advanced Robotics Systems*, 3(1):43–48, 2006.

[127] Ugo Pattacini. Modular cartesian controllers for humanoid robots: Design and implementation on the icub. *Dizertacná práca, Istituto Italiano di Tecnologia*, 2011.

[128] Richard W Homan, John Herman, and Phillip Purdy. Cerebral location of international

10–20 system electrode placement. *Electroencephalography and Clinical Neurophysiology*, 66(4):376–382, 1987.

[129] Li Hu and Zhiguo Zhang. *EEG Signal Processing and Feature Extraction*. Springer, 2019.

[130] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[131] Xie Song-yun, Wang Peng-wei, Zhang Hai-jun, and Zhao Hai-tao. Research on the classification of brain function based on SVM. In *Bioinformatics and Biomedical Engineering, 2008. ICBBE 2008. The 2nd International Conference on*, pages 1931–1934. IEEE, 2008.

[132] Wolfgang Klimesch, Michael Doppelmayr, Harald Russegger, Thomas Pachinger, and J Schwaiger. Induced alpha band power changes in the human EEG and attention. *Neuroscience Letters*, 244(2):73–76, 1998.

[133] Ardesheer Talati and Joy Hirsch. Functional specialization within the medial frontal gyrus for perceptual go/no-go decisions based on what, when, and where related information: an fMRI study. *Journal of Cognitive Neuroscience*, 17(7):981–993, 2005.

[134] Christa Neuper, Reinhold Scherer, Miriam Reiner, and Gert Pfurtscheller. Imagery of motor actions: Differential effects of kinesthetic and visual–motor mode of imagery in single-trial EEG. *Cognitive Brain Research*, 25(3):668–677, 2005.

[135] Laura Acqualagna, Loic Botrel, Carmen Vidaurre, Andrea Kübler, and Benjamin Blankertz. Large-scale assessment of a fully automatic co-adaptive motor imagery-based brain computer interface. *PloS one*, 11(2):e0148886, 2016.

[136] Fabien Lotte and Cuntai Guan. Learning from other subjects helps reducing brain-computer interface calibration time. In *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 614–617. IEEE, 2010.

[137] Arnaud Delorme and Scott Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, 2004.

[138] Andrea Mognon, Jorge Jovicich, Lorenzo Bruzzone, and Marco Buiatti. ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, 48(2):229–240, 2011.

[139] Steven J Luck. *An introduction to the event-related potential technique.* MIT press, 2014.

[140] Salil H Patel and Pierre N Azzam. Characterization of N200 and P300: selected studies of the event-related potential. *International Journal of Medical Sciences*, 2(4):147, 2005.

[141] Inaki Iturrate, Luis Montesano, and Javier Minguez. Task-dependent signal variations in EEG error-related potentials for brain–computer interfaces. *Journal of Neural Engineering*, 10(2):026024, 2013.

[142] Olivia K Carrick, James C Thompson, James A Epling, and Aina Puce. It's all in the eyes: neural responses to socially significant gaze shifts. *Neuroreport*, 18(8):763, 2007.

[143] Michael A Goodrich and Alan C Schultz. Human-robot interaction: a survey. *Foundations and Trends in Human-Computer Interaction*, 1(3):203–275, 2007.

[144] Kerstin Dautenhahn, Sarah Woods, Christina Kaouri, Michael L Walters, Kheng Lee Koay, and Iain Werry. What is a robot companion-friend, assistant or butler? In *Intelligent Robots and Systems, 2005. 2005 IEEE/RSJ International Conference on*, pages 1192–1197. IEEE, 2005.

[145] Chien-Ming Huang, Maya Cakmak, and Bilge Mutlu. Adaptive coordination strategies for human-robot handovers. In *Robotics: Science and Systems*, 2015.

[146] Matthew C Gombolay, Reymundo A Gutierrez, Shanelle G Clarke, Giancarlo F Sturla, and Julie A Shah. Decision-making authority, team efficiency and human worker satisfaction in mixed human–robot teams. *Autonomous Robots*, 39(3):293–312, 2015.

[147] Chien-Ming Huang and Bilge Mutlu. Robot behavior toolkit: generating effective social behaviors for robots. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 25–32. ACM, 2012.

[148] Christoph Bartneck, Elizabeth Croft, and Dana Kulic. Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots. In *Metrics for HRI workshop, Technical Report*, volume 471, pages 37–44. Citeseer, 2008.

[149] Michiel Joosse, Aziez Sardar, Manja Lohse, and Vanessa Evers. BEHAVE-II: The revised set of measures to assess users? attitudinal and behavioral responses to a social robot. *International Journal of Social Robotics*, 5(3):379–388, 2013.

[150] Bilge Mutlu and Jodi Forlizzi. Robots in organizations: the role of workflow, social, and

environmental factors in human-robot interaction. In *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*, pages 287–294. ACM, 2008.

[151] Ricardo Chavarriaga and José del R Millán. Learning from EEG error-related potentials in noninvasive brain-computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(4):381–388, 2010.

[152] Jason Omedes, Inaki Iturrate, Luis Montesano, and Javier Minguez. Using frequency-domain features for the generalization of EEG error-related potentials among different tasks. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5263–5266. IEEE, 2013.

[153] Alberto Llera, Marcel AJ van Gerven, Vicenç Gómez, Ole Jensen, and Hilbert J Kappen. On the use of interaction error potentials for adaptive brain computer interfaces. *Neural Networks*, 24(10):1120–1127, 2011.

[154] Larbi Boubchir, Youcef Touati, Boubaker Daachi, and Arab Ali Chérif. EEG error potentials detection and classification using time-frequency features for robot reinforcement learning. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1761–1764. IEEE, 2015.

[155] Jason Omedes, Inaki Iturrate, and Luis Montesano. Brain connectivity in continuous error tasks. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3997–4000. IEEE, 2014.

[156] Joos Behncke, Robin T Schirrmeister, Wolfram Burgard, and Tonio Ball. The signature of robot action success in EEG signals of a human observer: Decoding and visualization using deep convolutional neural networks. In *2018 6th International Conference on Brain-Computer Interface (BCI)*, pages 1–6. IEEE, 2018.

[157] Jason Omedes, Inaki Iturrate, and Luis Montesano. Asynchronous detection of error potentials. In *Proceedings of the 6th Brain-Computer Interface Conference 2014*, 2014.

[158] Martin Spüler, Christian Niethammer, Wolfgang Rosenstiel, and Martin Bogdan. Classification of error-related potentials in EEG during continuous feedback. In *Proceedings of the 6th International Brain-Computer Interface Conference (BCI2014)(Graz)*, pages 24–27, 2014.

[159] Jason Omedes, Iñaki Iturrate, Javier Minguez, and Luis Montesano. Analysis and asynchronous detection of gradually unfolding errors during monitoring tasks. *Journal of Neural Engineering*, 12(5):056001, 2015.

[160] Iñaki Iturrate, Luis Montesano, Ricardo Chavarriaga, José del R Millán, and Javier Minguez. Minimizing calibration time using inter-subject information of single-trial recognition of error potentials in brain-computer interfaces. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, 2011*, pages pp. 6369–6372, 2011.

[161] Jonathan Grizou, Inaki Iturrate, Luis Montesano, Manuel Lopes, and Pierre-Yves Oudeyer. Zero-calibration BMIs for sequential tasks using error-related potentials. In *IROS 2013 Workshop on Neuroscience and Robotics*, 2013.

[162] Jonathan Grizou, Iñaki Iturrate, Luis Montesano, Pierre-Yves Oudeyer, and Manuel Lopes. Calibration-free BCI based control. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014.

[163] Inaki Iturrate, Ricardo Chavarriaga, Luis Montesano, Javier Minguez, and José del R Millán. Latency correction of error-related potentials reduces BCI calibration time. In *6th Brain-Computer Interface Conference 2014*, 2014.

[164] Florian M Schönleitner, Lukas Otter, Stefan K Ehrlich, and Gordon Cheng. A comparative study on adaptive subject-independent classification models for zero-calibration error-potential decoding. In *2019 IEEE International Conference on Cyborg and Bionic Systems (CBS)*, pages 85–90. IEEE, 2019.

[165] Florian Schönleitner, Lukas Otter, Stefan K Ehrlich, and Gordon Cheng. Calibration-free error-related potential decoding with adaptive subject-independent models: A comparative study. Under review in IEEE Transactions on Medical Robotics and Bionics (T-MRB).

[166] Su Kyoung Kim and Elsa Andrea Kirchner. Classifier transferability in the detection of error related potentials from observation to interaction. In *2013 IEEE International Conference on Systems, Man, and Cybernetics*, pages 3360–3365. IEEE, 2013.

[167] Iñaki Iturrate, Jonathan Grizou, Jason Omedes, Pierre-Yves Oudeyer, Manuel Lopes, and Luis Montesano. Exploiting task constraints for self-calibrated brain-machine interface control using error-related potentials. *PloS one*, 10(7):e0131491, 2015.

[168] Su Kyoung Kim and Elsa Andrea Kirchner. Handling few training data: classifier transfer between different types of error-related potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(3):320–332, 2016.

[169] Pierre W Ferrez and José del R Millán. You are wrong!—automatic detection of interac-

tion errors from brain waves. In *Proceedings of the 19th international joint conference on Artificial intelligence*, 2005.

[170] Martin Spüler, Wolfgang Rosenstiel, and Martin Bogdan. Online adaptation of a c-VEP brain-computer interface (BCI) based on error-related potentials and unsupervised learning. *PloS one*, 7(12):e51077, 2012.

[171] Inaki Iturrate, Luis Montesano, and Javier Minguez. Single trial recognition of error-related potentials during observation of robot operation. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 4181–4184. IEEE, 2010.

[172] Alex Kreilinger, Christa Neuper, and Gernot R Müller-Putz. Error potential detection during continuous movement of an artificial arm controlled by brain–computer interface. *Medical & Biological Engineering & Computing*, 50(3):223–230, 2012.

[173] Andres F Salazar-Gomez, Joseph DelPreto, Stephanie Gil, Frank H Guenther, and Daniela Rus. Correcting robot mistakes in real time using EEG signals. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6570–6577. IEEE, 2017.

[174] Dominik Welke, Joos Behncke, Marina Hader, Robin Tibor Schirrmeister, Andreas Schönau, Boris Eßmann, Oliver Müller, Wolfram Burgard, and Tonio Ball. Brain responses during robot-error observation. *arXiv preprint arXiv:1708.01465*, 2017.

[175] Alexandra Frischen, Andrew P Bayliss, and Steven P Tipper. Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin*, 133(4):694, 2007.

[176] Bilge Mutlu, Toshiyuki Shiwa, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. Footing in human-robot conversations: how robots might shape participant roles using gaze cues. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction*, pages 61–68. ACM, 2009.

[177] Jonathan W Peirce. PsychoPy - psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1-2):8–13, 2007.

[178] David Gouaillier, Vincent Hugel, Pierre Blazevic, Chris Kilner, Jerome Monceaux, Pascal Lafourcade, Brice Marnier, Julien Serre, and Bruno Maisonnier. The NAO humanoid: a combination of performance and affordability. *CoRR abs/0807.3223*, 2008.

[179] Alois Schlögl, Claudia Keinrath, Doris Zimmermann, Reinhold Scherer, Robert Leeb,

and Gert Pfurtscheller. A fully automated correction method of EOG artifacts in EEG recordings. *Clinical Neurophysiology*, 118(1):98–104, 2007.

[180] Jerome H Friedman. Regularized discriminant analysis. *Journal of the American Statistical Association*, 84(405):165–175, 1989.

[181] Benjamin Blankertz, Steven Lemm, Matthias Treder, Stefan Haufe, and Klaus-Robert Müller. Single-trial analysis and classification of ERP components - a tutorial. *NeuroImage*, 56(2):814–825, 2011.

[182] Juliane Schäfer and Korbinian Strimmer. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*, 4(1), 2005.

[183] Etienne Combrisson and Karim Jerbi. Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *Journal of Neuroscience Methods*, 250:126–136, 2015.

[184] Clay B Holroyd, Joseph Dien, and Michael GH Coles. Error-related scalp potentials elicited by hand and foot movements: evidence for an output-independent error-processing system in humans. *Neuroscience Letters*, 242(2):65–68, 1998.

[185] W David Hairston, Keith W Whitaker, Anthony J Ries, Jean M Vettel, J Cortney Bradford, Scott E Kerick, and Kaleb McDowell. Usability of four commercially-oriented EEG systems. *Journal of Neural Engineering*, 11(4):046018, 2014.

[186] Ana Alves-Pinto, Stefan Ehrlich, Gordon Cheng, Varvara Turova, Tobias Blumenstein, and Renée Lampe. Effects of short-term piano training on measures of finger tapping, somatosensory perception and motor-related brain activity in patients with cerebral palsy. *Neuropsychiatric Disease and Treatment*, 13:2705, 2017.

[187] Flavio TP Oliveira, John J McDonald, and David Goodman. Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *Journal of Cognitive Neuroscience*, 19(12):1994–2004, 2007.

[188] Jérôme Sallet, René Quilodran, Marie Rothé, Julien Vezoli, Jean-Paul Joseph, and Emmanuel Procyk. Expectations, gains, and losses in the anterior cingulate cortex. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4):327–336, 2007.

[189] Gerwin Schalk, Jonathan R Wolpaw, Dennis J McFarland, and Gert Pfurtscheller.

EEG-based communication: presence of an error potential. *Clinical Neurophysiology*, 111(12):2138–2144, 2000.

[190] Gert Pfurtscheller, Brendan Z Allison, Günther Bauernfeind, Clemens Brunner, Teodoro Solis Escalante, Reinhold Scherer, Thorsten O Zander, Gernot Mueller-Putz, Christa Neuper, and Niels Birbaumer. The hybrid BCI. *Frontiers in Neuroscience*, 4:3, 2010.

[191] Lucas C Parra, Clay D Spence, Adam D Gerson, and Paul Sajda. Response error correction - a demonstration of improved human-machine performance using real-time EEG monitoring. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):173–177, 2003.

[192] Benjamin Blankertz, Guido Dornhege, Christin Schafer, Roman Krepki, Jens Kohlmorgen, K-R Muller, Volker Kunzmann, Florian Losch, and Gabriel Curio. Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial EEG analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):127–131, 2003.

[193] Pierre W Ferrez and José del R Millán. Simultaneous real-time detection of motor imagery and error-related potentials for improved BCI accuracy. In *Proceedings of the 4th International Brain-Computer Interface Workshop and Training Course*, pages 197–202, 2008.

[194] Pierre Ferrez and José Millán. EEG-based brain-computer interaction: Improved accuracy by automatic single-trial error detection. In *Advances in Neural Information Processing Systems*, pages 441–448, 2008.

[195] Nico M Schmidt, Benjamin Blankertz, and Matthias S Treder. Online detection of error-related potentials boosts the performance of mental typewriters. *BMC Neuroscience*, 13(1):19, 2012.

[196] Martin Spüler, Michael Bensch, Sonja Kleih, Wolfgang Rosenstiel, Martin Bogdan, and Andrea Kübler. Online use of error-related potentials in healthy users and people with severe motor impairment increases performance of a P300-BCI. *Clinical Neurophysiology*, 123(7):1328–1337, 2012.

[197] Aniana Cruz, Gabriel Pires, and Urbano J Nunes. Double ErrP detection for automatic error correction in an ERP-based BCI speller. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(1):26–36, 2018.

[198] Saugat Bhattacharyya, Amit Konar, and DN Tibarewala. Motor imagery, P300 and error-

related EEG-based robot arm movement control for rehabilitation purpose. *Medical & Biological Engineering & Computing*, 52(12):1007–1017, 2014.

[199] Mahta Mousavi, Adam S Koerner, Qiong Zhang, Eunho Noh, and Virginia R de Sa. Improving motor imagery BCI with user response to feedback. *Brain-Computer Interfaces*, 4(1-2):74–86, 2017.

[200] Ricardo Chavarriaga, Pierre W Ferrez, and José del R Millán. To err is human: Learning from error potentials in brain-computer interfaces. In *Advances in Cognitive Neurodynamics ICCN 2007*, pages 777–782. Springer, 2008.

[201] Ricardo Chavarriaga, Andrea Biasiucci, Killian Förster, Daniel Roggen, Gerhard Tröster, and José del R Millán. Adaptation of hybrid human-computer interaction systems using EEG error-related potentials. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 4226–4229. IEEE, 2010.

[202] Kilian Förster, Andrea Biasiucci, Ricardo Chavarriaga, José del R Millán, Daniel Roggen, and Gerhard Tröster. On the use of brain decoded signals for online user adaptive gesture recognition systems. In *International Conference on Pervasive Computing*, pages 427–444. Springer, 2010.

[203] Inaki Iturrate, Luis Montesano, and Javier Minguez. Shared-control brain-computer interface for a two dimensional reaching task using EEG error-related potentials. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5258–5262. IEEE, 2013.

[204] Iñaki Iturrate, Jason Omedes, and Luis Montesano. Shared control of a robot using EEG-based feedback signals. In *Proceedings of the 2nd Workshop on Machine Learning for Interactive Systems: Bridging the Gap Between Perception, Action and Communication*, pages 45–50. ACM, 2013.

[205] Huaijian Zhang, Ricardo Chavarriaga, Lucian Gheorghe, and José del R Millán. Inferring driver's turning direction through detection of error related brain activity. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2196–2199. IEEE, 2013.

[206] Huaijian Zhang, Ricardo Chavarriaga, Zahra Khaliliardali, L Gheorghe, Inaki Iturrate, and J d R Millán. EEG-based decoding of error-related brain activity in a real-world driving task. *Journal of Neural Engineering*, 12(6):066028, 2015.

[207] Jin Woo Choi, Taehyean Choi, Shinjeong Kim, and Sungho Jo. Towards utilization of

error-related potentials for brain-to-vehicle communication. In *2019 7th International Winter Conference on Brain-Computer Interface (BCI)*, pages 1–6. IEEE, 2019.

[208] Candace L Sidner, Cory D Kidd, Christopher Lee, and Neal Lesh. Where to look: a study of human-robot engagement. In *Proceedings of the 9th International Conference on Intelligent User Interfaces*, pages 78–84. ACM, 2004.

[209] Christian Kothe. Lab streaming layer (LSL). *https://github. com/sccn/labstreaminglayer. Accessed on October*, 26:2015, 2014.

[210] Richard S Sutton and Andrew G Barto. *Introduction to reinforcement learning*, volume 135. MIT press Cambridge, 1998.

[211] Uri Hadar, Timothy J Steiner, Ewan C Grant, and F Clifford Rose. The timing of shifts of head postures during conservation. *Human Movement Science*, 3(3):237–245, 1984.

[212] Yann Renard, Fabien Lotte, Guillaume Gibert, Marco Congedo, Emmanuel Maby, Vincent Delannoy, Olivier Bertrand, and Anatole Lécuyer. Openvibe: An open-source software platform to design, test, and use brain–computer interfaces in real and virtual environments. *Presence: Teleoperators and Virtual Environments*, 19(1):35–53, 2010.

[213] Vijaykumar Gullapalli, Judy A Franklin, and Hamid Benbrahim. Acquiring robot skills via reinforcement learning. *IEEE Control Systems Magazine*, 14(1):13–24, 1994.

[214] Jan Peters and Stefan Schaal. Reinforcement learning of motor skills with policy gradients. *Neural Networks*, 21(4):682–697, 2008.

[215] John Polich. Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10):2128–2148, 2007.

[216] Bertrand Rivet, Antoine Souloumiac, Virginie Attina, and Guillaume Gibert. xDAWN algorithm to enhance evoked potentials: application to brain–computer interface. *IEEE Transactions on Biomedical Engineering*, 56(8):2035–2043, 2009.

[217] George Townsend, Bernhard Graimann, and Gert Pfurtscheller. Continuous EEG classification during motor imagery-simulation of an asynchronous BCI. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 12(2):258–265, 2004.

[218] Matthew M Botvinick, Jonathan D Cohen, and Cameron S Carter. Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, 8(12):539–546, 2004.

[219] Matthew M Botvinick, Todd S Braver, Deanna M Barch, Cameron S Carter, and Jonathan D Cohen. Conflict monitoring and cognitive control. *Psychological Review*, 108(3):624, 2001.

[220] Joshua W Brown and Todd S Braver. Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, 307(5712):1118–1121, 2005.

[221] Timothy EJ Behrens, Mark W Woolrich, Mark E Walton, and Matthew FS Rushworth. Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9):1214, 2007.

[222] Massimo Silvetti, Ruth Seurinck, and Tom Verguts. Value and prediction error in medial frontal cortex: integrating the single-unit and systems levels of analysis. *Frontiers in Human Neuroscience*, 5:75, 2011.

[223] Clay B Holroyd and Samuel M McClure. Hierarchical control over effortful behavior by rodent medial frontal cortex: A computational model. *Psychological Review*, 122(1):54, 2015.

[224] William H Alexander and Joshua W Brown. Frontal cortex function as derived from hierarchical predictive coding. *Scientific Reports*, 8(1):3843, 2018.

[225] Matthew M Botvinick. Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4):356–366, 2007.

[226] Tom Verguts, Eliana Vassena, and Massimo Silvetti. Adaptive effort investment in cognitive and physical tasks: A neurocomputational model. *Frontiers in Behavioral Neuroscience*, 9:57, 2015.

[227] William H Alexander and Joshua W Brown. Hierarchical error representation: a computational model of anterior cingulate and dorsolateral prefrontal cortex. *Neural Computation*, 27(11):2354–2410, 2015.

[228] Eliana Vassena, James Deraeve, and William H Alexander. Predicting motivation: computational models of PFC can explain neural coding of motivation and effort-based decision-making in health and disease. *Journal of Cognitive Neuroscience*, 29(10):1633–1645, 2017.

[229] M Andrea Pisauro, Elsa Fouragnan, Chris Retzler, and Marios G Philiastides. Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous EEG-fMRI. *Nature Communications*, 8:15808, 2017.

[230] Michael X Cohen and Charan Ranganath. Reinforcement learning signals predict future decisions. *Journal of Neuroscience*, 27(2):371–378, 2007.

[231] Leo P Sugrue, Greg S Corrado, and William T Newsome. Matching behavior and the representation of value in the parietal cortex. *Science*, 304(5678):1782–1787, 2004.

[232] Hannah M Bayer and Paul W Glimcher. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, 47(1):129–141, 2005.

[233] Kazuyuki Samejima, Yasumasa Ueda, Kenji Doya, and Minoru Kimura. Representation of action-specific reward values in the striatum. *Science*, 310(5752):1337–1340, 2005.

[234] Jeffrey C Lagarias, James A Reeds, Margaret H Wright, and Paul E Wright. Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on Optimization*, 9(1):112–147, 1998.

[235] Christopher John Cornish Hellaby Watkins. *Learning from delayed rewards*. PhD thesis, King's College, Cambridge, 1989.

[236] John Nash. Non-cooperative games. *Annals of Mathematics*, pages 286–295, 1951.

[237] Bezalel Peleg and Peter Sudhölter. *Introduction to the theory of cooperative games*, volume 34. Springer Science & Business Media, 2007.

[238] Peter Danielson. *Artificial morality: Virtuous robots for virtual games*. Routledge, 2002.

[239] Andrew G Brooks, Jesse Gray, Guy Hoffman, Andrea Lockerd, Hans Lee, and Cynthia Breazeal. Robot's play: interactive games with sociable machines. *Computers in Entertainment (CIE)*, 2(3):10–10, 2004.

[240] Agnieszka Wykowska, Eva Wiese, Aaron Prosser, and Hermann J Müller. Beliefs about the minds of others influence how we process sensory information. *PLoS One*, 9(4):e94339, 2014.

[241] Masahiro Mori. Bukimi no tani [the uncanny valley]. *Energy*, 7:33–35, 1970.

[242] David Premack and Guy Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4):515–526, 1978.

[243] Wako Yoshida, Ray J Dolan, and Karl J Friston. Game theory of mind. *PLoS Computational Biology*, 4(12):e1000254, 2008.

[244] Wako Yoshida, Ben Seymour, Karl J Friston, and Raymond J Dolan. Neural mechanisms of belief inference during cooperative games. *Journal of Neuroscience*, 30(32):10744–10751, 2010.

[245] Brian Skyrms. *The stag hunt and the evolution of social structure*. Cambridge University Press, 2004.

[246] Andrea Kübler. The history of BCI: From a vision for the future to real support for personhood in people with locked-in syndrome. *Neuroethics*, pages 1–18, 2019.

[247] Vojkan Mihajlović, Bernard Grundlehner, Ruud Vullers, and Julien Penders. Wearable, wireless EEG solutions in daily life applications: what are we missing? *IEEE Journal of Biomedical and Health Informatics*, 19(1):6–21, 2014.

[248] Gert Pfurtscheller and FH Lopes Da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110(11):1842–1857, 1999.

[249] Edwin M Robertson. The serial reaction time task: implicit motor skill learning? *Journal of Neuroscience*, 27(38):10073–10075, 2007.

[250] Stefan Debener, Falk Minow, Reiner Emkes, Katharina Gandras, and Maarten De Vos. How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology*, 49(11):1617–1621, 2012.

[251] Nicholas A Badcock, Kathryn A Preece, Bianca de Wit, Katharine Glenn, Nora Fieder, Johnson Thie, and Genevieve McArthur. Validation of the Emotiv EPOC EEG system for research quality auditory event-related potentials in children. *PeerJ*, 3:e907, 2015.

[252] Scott Makeig, Stefan Debener, Julie Onton, and Arnaud Delorme. Mining event-related brain dynamics. *Trends in Cognitive Sciences*, 8(5):204–210, 2004.

[253] Stefan Ehrlich, Ana Alves-Pinto, Renée Lampe, and Gordon Cheng. A simple and practical sensorimotor EEG device for recording in patients with special needs. In *Neurotechnix2017, CogNeuroEng 2017, Symposium on Cognitive Neural Engineering*, 2017.

[254] Olave E Krigolson, Chad C Williams, Angela Norton, Cameron D Hassall, and Francisco L Colino. Choosing MUSE: Validation of a low-cost, portable EEG system for ERP research. *Frontiers in Neuroscience*, 11:109, 2017.

Appendix A: Supplementary material

Supplementary material for chapter 4

Table 11 CURSOR behavioral data: scenario order (C=cursor; R=robot), mean reaction times ($\overline{RT}$), and number of human committed errors ($nErr$).

	s02	s03	s04	s05	s06	s07	s08	s09	s10	s11	s13	AVG±SD
scen. order	R/C	C/R	R/C	C/R	R/C	C/R	R/C	C/R	R/C	C/R	C/R	
$\overline{RT}$ [ms]	425.9	408.6	340.9	421.5	358.4	516.0	356.7	391.9	599.0	396.6	418.1	421.2±75.4
$nErr$	16	12	15	10	20	4	48	36	3	17	2	16.6±14.1

Table 12 ROBOT behavioral data: scenario order (C=cursor; R=robot), mean reaction times ($\overline{RT}$), and number of human committed errors ($nErr$).

	s02	s03	s04	s05	s06	s07	s08	s09	s10	s11	s13	AVG±SD
scen. order	R/C	C/R	R/C	C/R	R/C	C/R	R/C	C/R	R/C	C/R	C/R	
$\overline{RT}$ [ms]	396.3	390.7	355.4	397.4	359.7	452.3	389.0	379.2	508.5	394.6	411.2	403.1±43.4
$nErr$	20	11	4	14	8	12	18	33	14	28	13	15.9±8.5

Table 13 Spatio-temporal 2D correlation coefficient between difference average ERP of CURSOR and ROBOT scenario.

	s02	s03	s04	s05	s06	s07	s08	s09	s10	s11	s13	AVG±SD
r	0.53	0.48	0.84	0.15	0.54	0.42	0.16	0.24	0.13	0.63	0.55	0.42±0.22

Table 14 CURSOR within-session cross-validation results: Overall accuracy (Acc), true-positive rate / machine-error recognition rate (TPR), true negative rate / non-error recognition rate (TNR), and area under receiver operator curve (AUC).

	s02	s03	s04	s05	s06	s07	s08	s09	s10	s11	s13	AVG±SD
Acc [%]	85.9	88.6	96.3	86.5	88.3	96.1	87.1	94.4	93.6	88.5	90.8	90.6±3.9
TPR [%]	82.0	84.1	91.2	84.4	84.1	94.3	83.9	91.6	92.8	85.1	86.7	87.3±4.3
TNR [%]	87.8	91.1	99.1	87.7	90.4	97.0	88.9	96.0	94.0	90.1	92.5	92.2±3.8
AUC	0.91	0.94	0.98	0.91	0.92	0.99	0.91	0.98	0.98	0.93	0.94	0.95±0.03

Table 15 ROBOT within-session cross-validation results: Overall accuracy (Acc), true-positive rate / machine-error recognition rate (TPR), true negative rate / non-error recognition rate (TNR), and area under receiver operator curve (AUC).

	s02	s03	s04	s05	s06	s07	s08	s09	s10	s11	s13	AVG±SD
Acc [%]	65.9	69.8	80.2	67.0	77.3	76.9	60.2	64.3	54.5	67.1	75.6	69.0±7.9
TPR [%]	61.8	70.3	72.3	64.6	72.6	72.6	61.0	59.9	53.3	67.4	71.5	66.1±6.5
TNR [%]	67.9	69.5	85.0	68.3	80.1	79.0	59.7	66.7	55.0	66.9	78.0	70.6±9.1
AUC	0.68	0.75	0.88	0.72	0.82	0.83	0.64	0.66	0.56	0.70	0.82	0.73±0.10

Table 16 Cross-session validation results. Classification model trained with CURSOR data and tested on ROBOT data: Overall accuracy (Acc), true-positive rate / machine-error recognition rate (TPR), true negative rate / non-error recognition rate (TNR), and area under receiver operator curve (AUC).

	s02	s03	s04	s05	s06	s07	s08	s09	s10	s11	s13	AVG±SD
Acc [%]	72.7	72.8	76.0	60.9	71.7	72.3	56.0	69.0	62.8	67.4	69.4	68.3±6.1
TPR [%]	34.2	45.4	37.8	54.1	43.0	15.2	29.1	42.1	11.6	43.6	28.2	34.9±13.0
TNR [%]	91.1	87.9	99.4	64.6	89.2	99.7	72.3	83.5	88.8	80.1	93.8	86.4±10.8
AUC	0.67	0.71	0.88	0.62	0.76	0.79	0.57	0.65	0.50	0.64	0.71	0.68±0.11

Table 17 Cross-session validation results.Classification model trained with ROBOT data and tested on CURSOR data: Overall accuracy (Acc), true-positive rate / machine-error recognition rate (TPR), true negative rate / non-error recognition rate (TNR), and area under receiver operator curve (AUC).

	s02	s03	s04	s05	s06	s07	s08	s09	s10	s11	s13	AVG±SD
Acc [%]	71.3	76.0	87.6	70.6	84.4	83.7	60.8	84.7	42.9	69.8	72.1	73.1±13.0
TPR [%]	64.7	84.5	97.1	63.5	75.9	95.7	19.5	89.5	36.9	58.4	87.6	70.3±24.8
TNR [%]	74.4	71.3	82.5	74.8	88.7	77.7	84.4	82.1	45.9	75.1	65.2	74.7±11.6
AUC	0.76	0.85	0.98	0.76	0.85	0.95	0.50	0.92	0.39	0.73	0.85	0.78±0.18

Supplementary material for chapter 5

Table 18 Offline (CALIB) cross-validation accuracies [%] including selected regularization parameter λ and online classification accuracies [%] during co-adaptation runs (CORL-I, CORL-II, and CORL-IV). Please note, no results are reported for CO-RL-III, since no ground truth data was recorded during this part of the experiment. Successful co-adaptation runs are marked in blue.

Participant ID	λ	TNR	TPR	Acc	TNR	TPR	Acc	TNR	TPR	Acc	TNR	TPR	Acc
		CALIB			CORL-I			CORL-II			CORL-IV		
s03	0.95	68.8	70.1	69.3	91.7	50.0	90.0	93.1	57.1	78.0	91.5	100.0	92.0
s04	0.95	86.3	86.2	86.3	100.0	83.3	92.0	95.2	82.8	88.0	95.2	93.1	94.0
s06	1.00	83.0	80.0	82.0	85.7	72.2	76.0	80.0	93.3	84.0	80.0	43.3	58.0
s07	0.85	82.8	79.9	81.4	75.0	84.2	82.0	80.7	79.0	80.0	93.8	82.4	86.0
s08	0.95	87.7	81.5	85.7	97.4	58.3	88.0	92.6	73.9	84.0	96.6	81.0	90.0
s09	0.95	93.5	91.4	92.8	89.5	83.3	88.0	88.9	78.6	86.0	91.9	69.2	86.0
s11	0.70	70.4	66.3	68.9	79.0	71.0	74.0	71.4	44.8	56.0	79.2	53.9	66.0
s12	0.95	85.3	83.8	84.7	96.8	79.0	90.0	95.2	50.0	88.0	93.0	85.7	92.0
s14	0.75	89.5	87.5	88.7	95.5	83.3	94.0	80.0	55.6	58.0	96.4	59.1	80.0
s15	0.95	72.6	71.9	72.4	90.5	75.0	88.0	85.7	72.4	78.0	95.4	57.1	90.0
s16	0.95	73.9	72.0	73.3	60.9	81.5	72.0	26.7	82.9	66.0	87.5	80.0	86.0
s17	0.90	81.3	82.7	81.8	71.4	93.1	84.0	94.4	85.7	92.0	87.1	89.5	88.0
s18	0.95	80.5	76.3	78.9	90.9	64.3	76.0	88.9	58.5	64.0	87.2	72.7	84.0
AVG		**80.2**	**79.2**	**80.5**	**86.5**	**75.3**	**84.2**	**82.5**	**70.4**	**77.1**	**90.4**	**74.4**	**84.0**
±SD		**±7.7**	**±7.5**	**±7.5**	**±11.7**	**±12.0**	**±7.4**	**±18.3**	**±15.4**	**±12.1**	**±5.8**	**±17.0**	**±10.6**
s05	1.00	50.0	51.8	50.5	63.0	56.5	60.0	68.4	35.5	48.0	57.1	34.5	44.0
s10	1.00	83.4	30.7	64.4	100.0	2.9	32.0	100.0	0.0	28.0	100.0	0.0	26.0
s13	0.95	51.1	48.9	50.2	85.7	55.2	68.0	61.1	68.8	66.0	68.4	67.7	68.0

Table 19 Average guessing performance [%] within five segments of 10 trials each for each co-adaptation run (CORL-I, CORL-II, and CORL-IV). Please note, no results are reported for CO-RL-III, since no ground truth information was recorded during this part of the experiment. Successful co-adaptation runs are marked in blue.

Participant ID	Trial: 1-10	Trial: 11-20	Trial: 21-30	Trial: 31-40	Trial: 41-50	AVG trial 1-50	Trial: 1-10	Trial: 11-20	Trial: 21-30	Trial: 31-40	Trial: 41-50	AVG trial 1-50	Trial: 1-10	Trial: 11-20	Trial: 21-30	Trial: 31-40	Trial: 41-50	AVG trial 1-50
			CORL-I						CORL-II						CORL-IV			
s03	80	100	100	100	100	**96**	60	50	20	70	90	**58**	90	90	90	100	100	**94**
s04	10	40	50	70	90	**52**	40	20	50	60	40	**42**	30	40	20	60	60	**42**
s06	20	40	10	30	40	**28**	40	60	80	80	90	**70**	60	50	40	50	0	**40**
s07	10	20	40	30	20	**24**	40	70	70	70	60	**62**	40	10	60	20	30	**32**
s08	50	80	80	90	80	**76**	10	60	80	70	50	**54**	30	50	30	80	100	**58**
s09	30	70	90	90	100	**76**	50	40	70	70	100	**72**	30	80	100	80	80	**74**
s11	20	30	70	40	30	**38**	40	10	30	70	60	**42**	40	50	60	60	30	**48**
s12	60	70	70	90	20	**62**	60	60	100	100	100	**84**	70	80	80	100	100	**86**
s14	60	100	100	100	80	**88**	40	0	0	0	10	**10**	30	40	100	80	30	**56**
s15	40	100	90	100	90	**84**	30	0	30	50	100	**42**	70	70	90	100	100	**86**
s16	10	30	50	90	50	**46**	20	10	30	30	60	**30**	50	80	80	90	100	**80**
s17	10	20	40	50	90	**42**	60	90	100	40	70	**72**	30	80	60	60	80	**62**
s18	10	50	40	60	60	**44**	20	0	30	0	40	**18**	40	80	80	90	100	**78**
MEDIAN	**20**	**50**	**70**	**90**	**80**	**52**	**40**	**40**	**50**	**70**	**60**	**54**	**40**	**70**	**80**	**80**	**80**	**62**
s05	60	40	70	30	70	**54**	40	20	60	30	40	**38**	40	50	50	40	30	**42**
s10	10	30	20	50	40	**30**	40	10	10	20	60	**28**	50	0	20	20	40	**26**
s13	70	40	20	50	30	**42**	50	50	20	50	10	**36**	20	40	90	10	30	**38**

Table 20 Average number of robot gaze transitions until participant decision (decision time). Successful co-adaptation runs are marked in blue.

Participant ID	Trial: 1-10	Trial: 11-20	Trial: 21-30	Trial: 31-40	Trial: 41-50	AVG trial 1-50	Trial: 1-10	Trial: 11-20	Trial: 21-30	Trial: 31-40	Trial: 41-50	AVG trial 1-50	Trial: 1-10	Trial: 11-20	Trial: 21-30	Trial: 31-40	Trial: 41-50	AVG trial 1-50
	CORL-I						CORL-II						CORL-IV					
s03	25	25	26	23	14	**23**	20	21	26	21	21	**22**	20	15	10	6	6	**12**
s04	9	13	11	10	10	**11**	11	9	8	8	8	**9**	11	11	12	12	16	**12**
s06	22	21	40	37	41	**32**	29	51	26	28	25	**31**	46	38	56	59	39	**48**
s07	9	8	6	4	7	**7**	11	7	4	4	3	**6**	4	3	4	4	4	**4**
s08	15	11	11	19	18	**15**	17	11	9	8	9	**11**	12	12	12	11	10	**11**
s09	19	17	14	13	12	**15**	16	11	17	8	8	**12**	16	9	8	6	7	**9**
s11	17	8	10	12	11	**12**	10	9	11	10	11	**10**	10	9	10	14	9	**11**
s12	22	16	17	18	16	**18**	23	14	11	8	8	**13**	12	13	12	9	10	**11**
s14	16	14	10	9	11	**12**	22	15	19	16	17	**18**	21	21	14	18	19	**19**
s15	23	13	12	9	6	**12**	20	13	27	17	7	**17**	8	8	7	6	6	**7**
s16	41	20	16	21	26	**25**	28	33	16	19	17	**23**	35	26	8	8	13	**18**
s17	17	11	13	13	12	**13**	12	11	15	15	11	**13**	15	17	18	13	15	**15**
s18	13	11	9	9	13	**11**	12	15	15	11	10	**13**	17	19	18	13	12	**16**
MEDIAN	**17**	**13**	**12**	**13**	**12**	**13**	**17**	**13**	**15**	**11**	**10**	**13**	**15**	**13**	**12**	**11**	**10**	**12**
s05	11	10	12	11	10	**11**	13	12	15	18	15	**14**	12	14	19	14	16	**15**
s10	8	6	6	7	6	**7**	8	6	7	6	9	**7**	11	10	11	9	11	**10**
s13	18	22	23	23	23	**22**	19	13	13	17	21	**16**	22	20	18	20	22	**20**
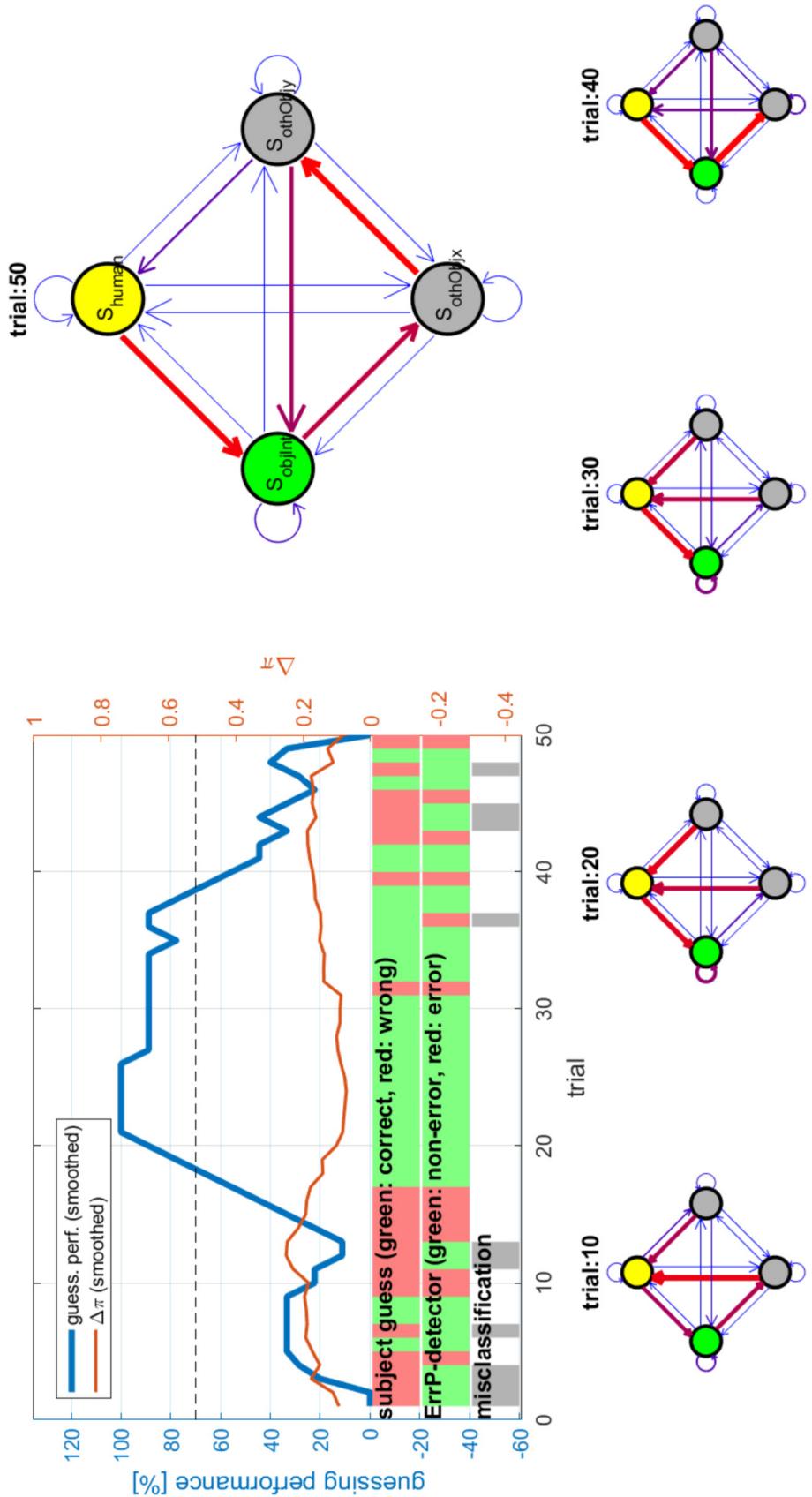
Figure 40 Exemplary case 1 (s14/CORL-IV) of temporary increase of guessing performance followed by decreases (unlearning).
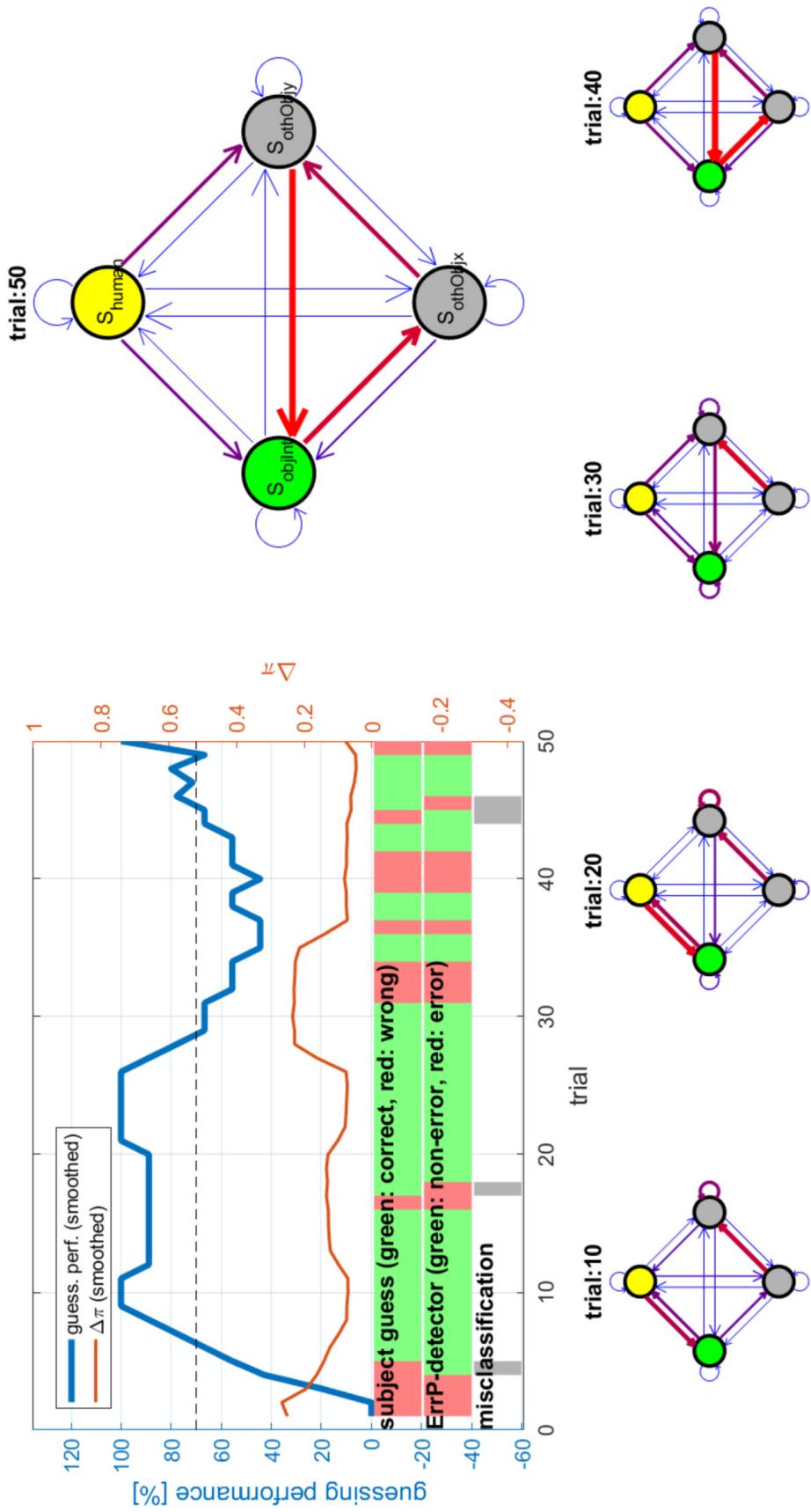
**Figure 41** Exemplary case 2 (s17/CORL-II) of temporary increase of guessing performance followed by decreases (unlearning).
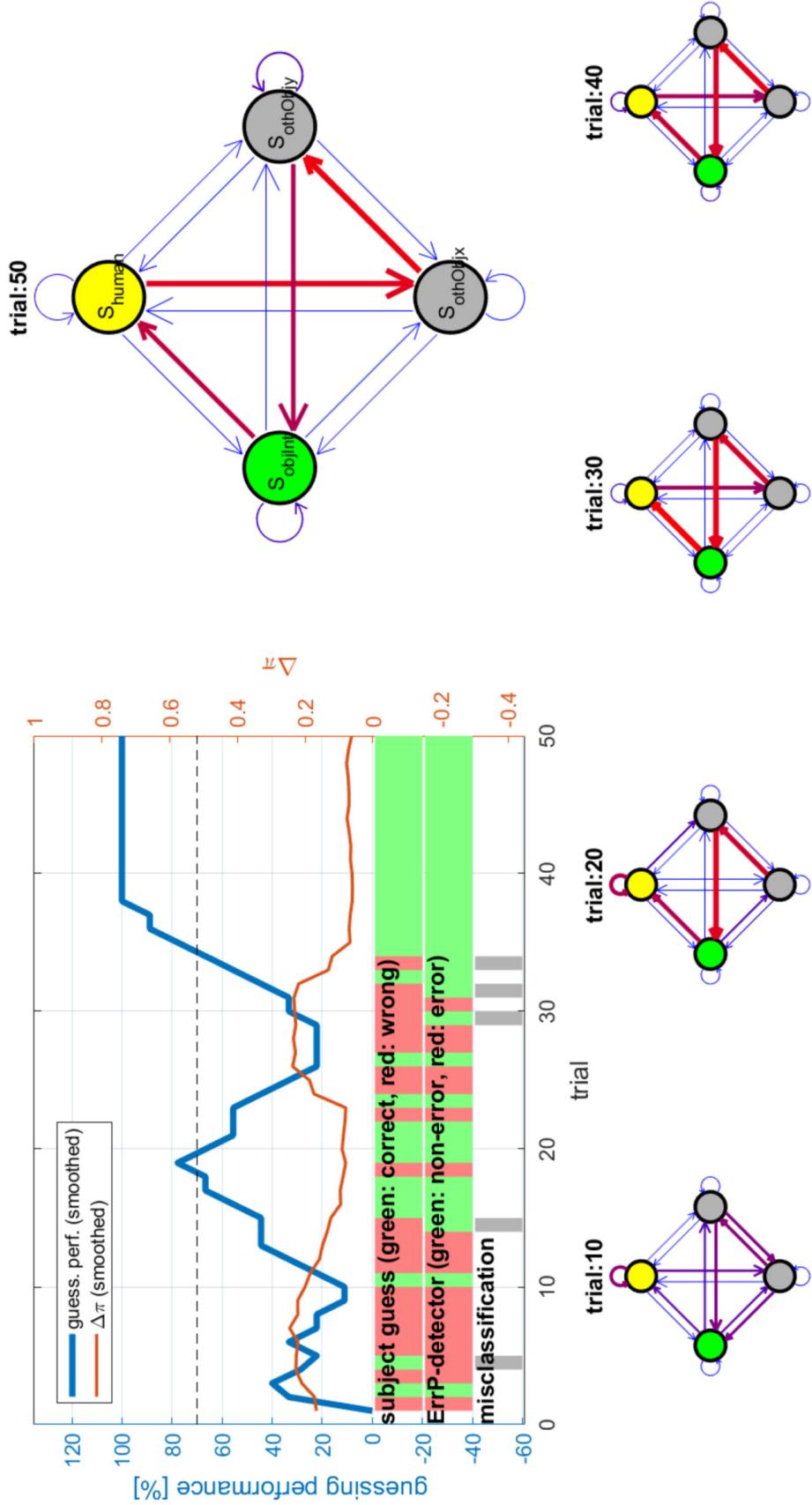
Figure 42 Exemplary case 1 (s08/CORL-IV) of possible policy re-adaptations.
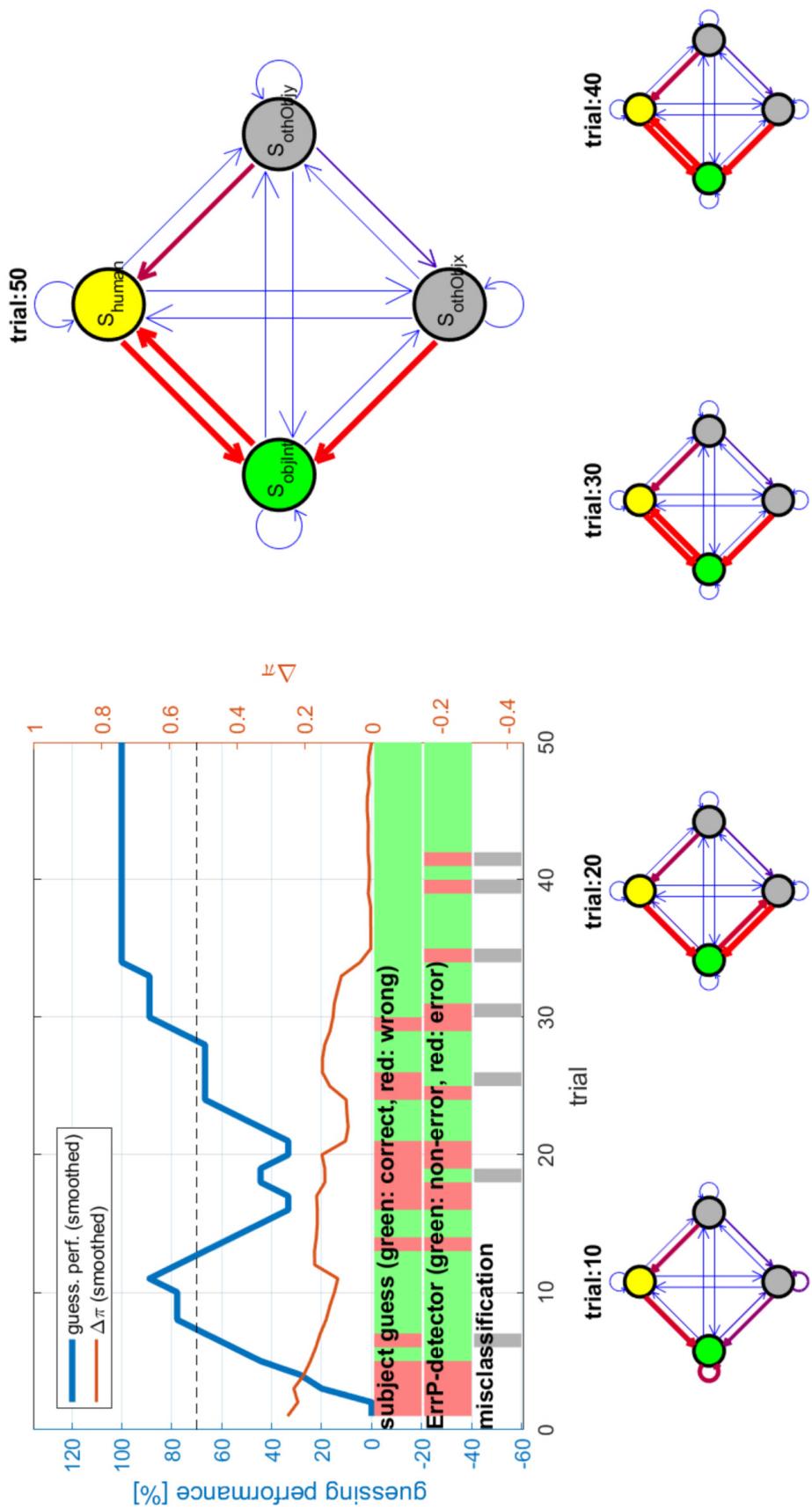
Figure 43 Exemplary case 2 (s09/CORL-II) of possible policy re-adaptations.

164
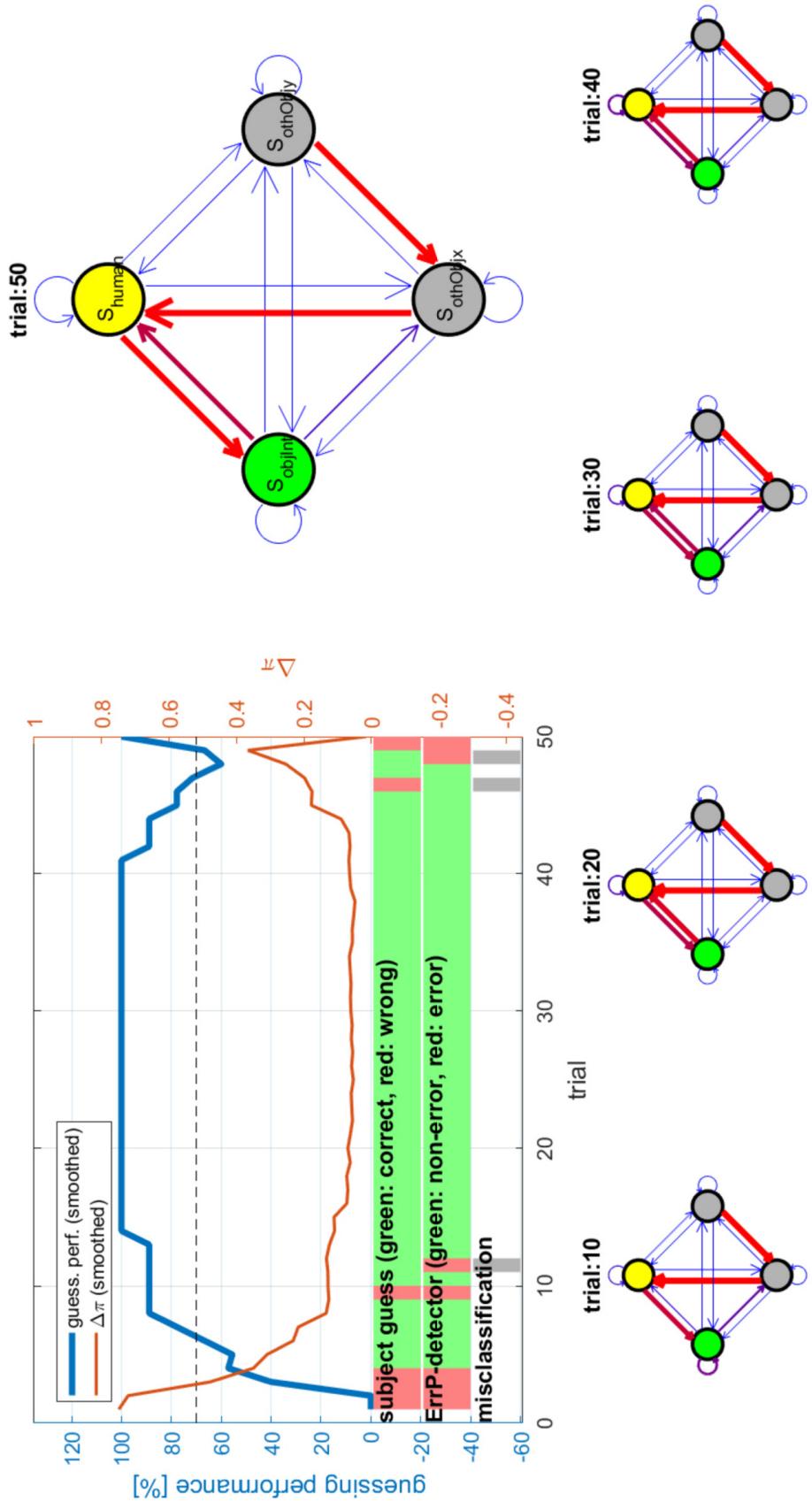
Figure 44 Exemplary case 3 (s14/CORL-I) of possible policy re-adaptations.

165

Supplementary material for chapter 6
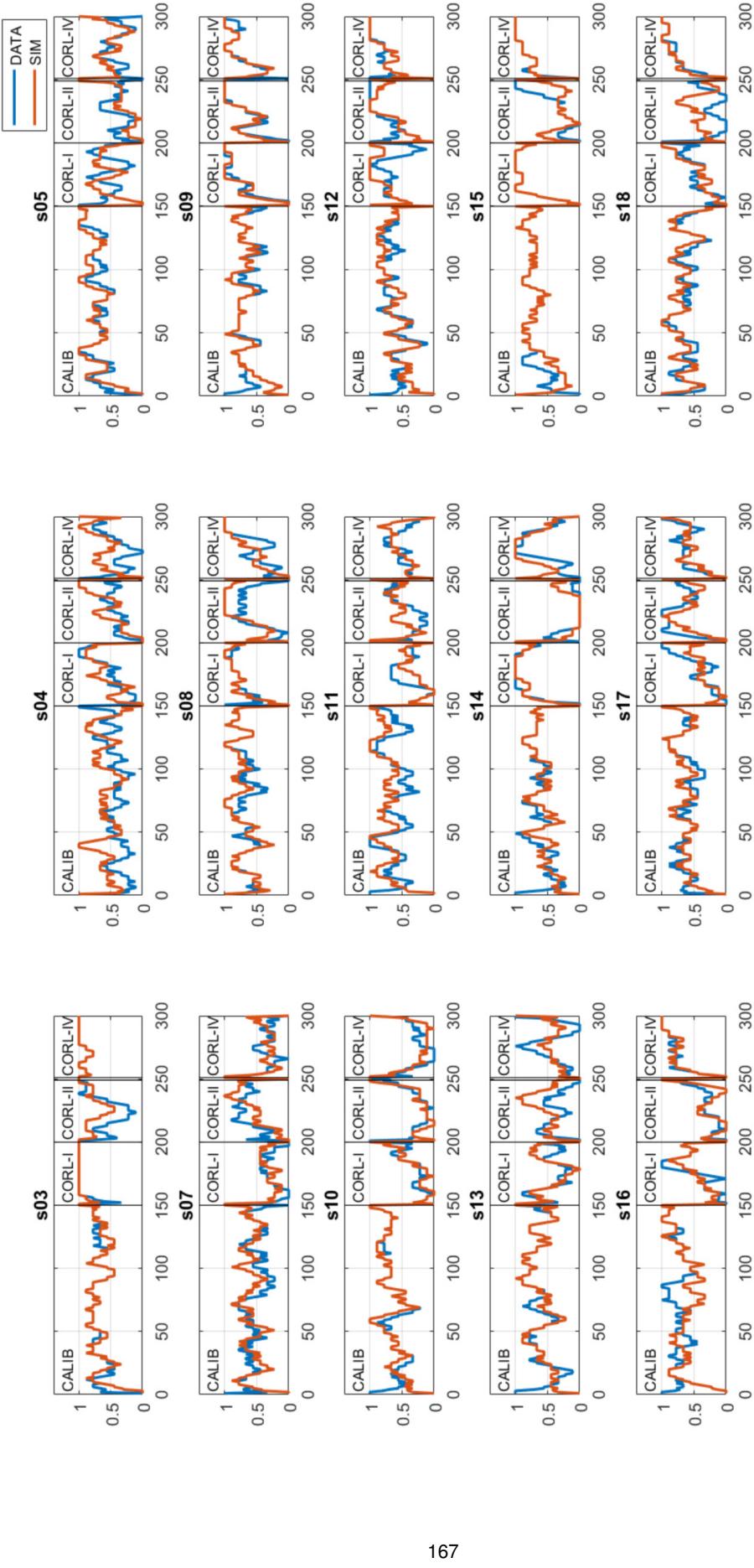
Figure 45 Qualitative results of goodness of fit: Each panel depicts individual results of simulated (orange) and real (blue) guessing performance per participant for CALIB, CORL-I, CORL-II, and CORL-IV (smoothed with a 10-trials kernel for each part of the experiment separately). Overall, the results show a good match between simulation and ground truth, indicating that the model captured relevant behavioral effects observed in the human data. Supplementing Figure 30 with results for all participants.

167
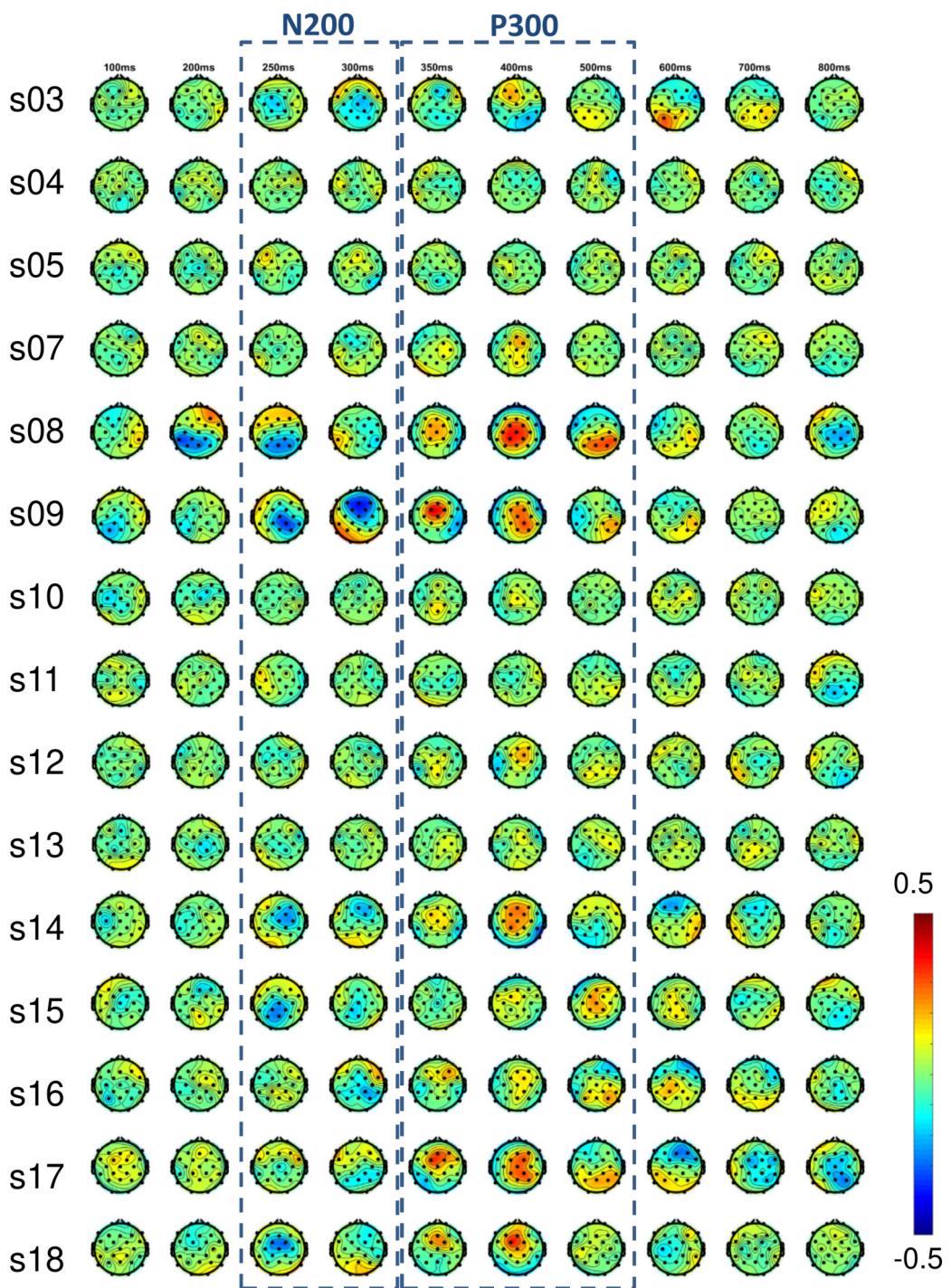
Figure 46 Single-trial spatio-temporal correlation coefficients between absolute model prediction error and ERPs. Supplementing Figure 31 with results for each participant individually.

Appendix B: Ease-of-use EEG technology for practical ErrP decoding

Introduction

Ever since the invention by Hans Berger in 1926 [50], EEG has been established as one of the main technologies for monitoring brain activity. EEG is currently used for diagnosis (e.g. epilepsy) and long-term monitoring (e.g. sleep monitoring) in clinical contexts; for research in human cognitive neuroscience; for BCI applications allowing paralyzed individuals rudimentary communication with their environment; and finally, for BCI-based neuro-rehabilitation of sensorimotor deficits (e.g. stroke). It is widely acknowledged that especially BCI applications could bring most benefit to the patients if they were applicable in the patient's homes [49, 246]. Despite technological advances during the last decades, EEG technology is, however, still rarely used outside the research laboratories and clinical contexts [247]. More recently developed alternative technologies for monitoring brain activity do not provide a solution to this quest: fMRI and positron emission tomography (PET) are very expensive, non-portable and require patients to remain still/immobile during data acquisition. fNIRS is comparably expensive and causes limited wearing comfort (tight mounting of sensors). Invasive technologies, such as intracortical EEG (ECoG), and single- and multi-electrode recordings require surgical implantation and as such are accompanied by corresponding health risks. EEG technology on the other hand is comparably inexpensive and comfortable, non-invasive, portable, and allowing for patient mobility. Therefore, we believe that near future home-use applications involving neuro-recording will likely remain EEG-based. Despite the advantages of EEG technology, there are still several limitations which render the technology impractical for straightforward home-use:

- Handling EEG technology requires professional training, excluding untrained personnel such as nurses, care givers, and users / patients themselves.

- EEG equipment has obtrusive looks which can raise distrust or embarrassment in users / patients.

- EEG signal acquisition is very sensitive to external interference and artifacts resulting in limited data quality especially when recording in uncontrolled environments (outside the laboratory and clinical contexts).

- EEG equipment is still relatively expensive ranging within prices of 5-100K EUR

Along this line, Mihajlovic and colleagues asked an essential question in the title of their 2015 survey paper: "Wearable, Wireless EEG Solutions in daily life applications: What are we missing?" [247]. From their survey, they concluded that future development of EEG-based neuro-technology would benefit from:

- *application-driven design* with limited functionality tailored to the target application.

- *end-user driven development* involving end-users, e.g. patients, as early as possible in the development process.

- *standardization and sharing of EEG data* allowing for benchmarking and facilitating the development of novel processing and decoding methods on larger datasets.

- development of *novel methods to handle EEG artifacts* to enhance data quality of recordings in uncontrolled environments.
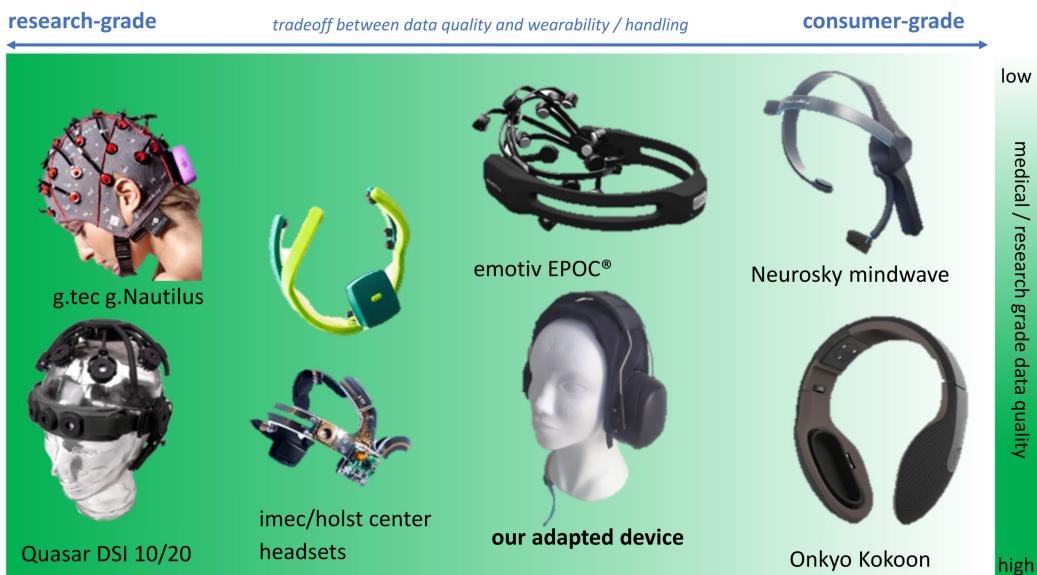
Figure 47 Overview of (in part commercially available) ease-of-use EEG systems: Ease-of-use EEG devices spanning from research-grade systems at the left delivering near clinical-grade data quality for BCI research and applications to consumer-grade systems on the right with limited data quality designed for entertainment, gaming, and neurofeedback applications. Please note, our adapted device (*ICS-Headset*) being located somewhat in the middle between research-grade and consumer-grade devices.

In recent years, several companies and institutions have contributed with such ease-of-use, application-specific EEG systems which in part are already commercially available. Some examples are illustrated in Figure 47, spanning from *research-grade systems* at the left delivering near clinical-grade data quality designed for BCI research and applications to *consumer-grade systems* on the right with limited data quality designed for entertainment, gaming, and neurofeedback applications. When looking at the range of commercially available EEG systems an evident gap between these two categories can be observed: What is missing are EEG systems which provide easy handling and unobtrusive design while still guaranteeing near research-grade data quality.

The prototypes developed at the Chair for Cognitive Systems (TUM) during the last years make a contribution in line with this missing type of EEG systems. An initial functional prototype was developed early 2016 for the recording of sensorimotor rhythms in adolescents and adults diagnosed with CP. An improved follow-up functional prototype was developed later in 2016 and employed in experimental studies with healthy participants. Both functional prototypes were validated with regard to their usability for measuring sensorimotor rhythms; the follow-up functional prototype was furthermore validated with regard to its usability for the measurement and decoding of ErrPs. This chapter presents the technical realization of these prototypes and their functional validation in experimental studies.

An unobtrusive, ease-of-use EEG prototype: The *ICS-Headset*

The development of the initial prototype was motivated by the special requirements of a study of sensorimotor disabilities in patients with motor disorders planned by collaborators from the TUM medical faculty. Envisaged, in particular, were adults and adolescents with a diagnosis of CP, a medical condition characterized by motor deficits caused by damage to the developing brain pre-, peri- or post-natal. The main hypothesis was that sensorimotor rhythms observable in patients' EEG would undergo long-term changes during a 6-weeks training of hand-motor coordination (in form of piano training). Sensorimotor rhythms are characterized by ERS/ERD of the brain mu- and beta rhythm and constitute a neuronal correlate of motor action [248]. ERS/ERD was here measured during the execution of a hand motor task adapted from the SRTT (serial reaction time task) [249].

The envisaged patient group present impaired muscular imbalance and increased muscular tonus. Deficits are not progressive but persist throughout the patients' life. The degree of impairment varies depending on the brain areas affected and motor limitations are often accompanied by learning difficulties, attention problems, perceptual impairments, speaking difficulties and/or epilepsy. The multi-symptomatology in CP makes it a challenging condition to research and demands a multidisciplinary approach when investigating possible rehabilitation strategies. From the lifelong medical follow-up of patients it becomes clear that, whilst primary medical care provides vital treatment, patients need for inclusion is likely to benefit strongly also from the collaboration between medical expertise, neuroscience and advances in neurotechnology. The following technical requirements were defined to both address functional specifications in line with the objectives of the planned study as well as design specifications that conform to the special needs of above described patient group.

Technical requirements

Technical adaptations were performed in a commercially available EEG system to conform to the following requirements:

- unobtrusive visual appearance that is not immediately associated with a clinical examination device.

- short preparation time: < 5 minutes.

- comfortable wearing for up to 30 minutes.

- easily adapted to different head sizes (from children to adults).

- electrodes positioning according to the 10-20 system [128] and covering the majority of sensorimotor areas.

- resistance to hygienic treatment.

- positive reception/good acceptance of the device by the participants.
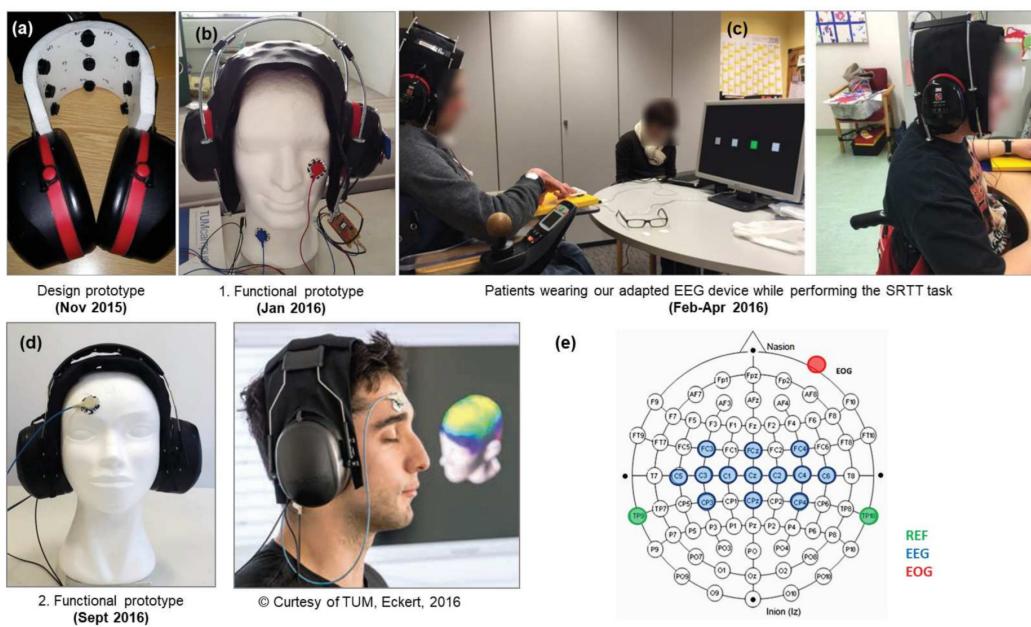
Figure 48 EEG prototypes developed at the Chair for Cognitive System (TUM): History of development of our adapted EEG device (*ICS-headset*). 13 semi-dry EEG channels covering sensorimotor areas are embedded into the elastic sheet between the earpieces. Note the optional electrooculogram (EOG) channel attached to the cheek or forehead for capturing vertical or horizontal eye movements. (a) Design prototype; (b) initial functional prototype used in a study with adolescents and adults diagnosed with CP, published in [186]; (c) patients wearing our EEG device while performing the SRTT in the first data collection session of [186]; (d) follow-up functional prototype; (e) Electrodes positioning of our adapted EEG device according to the international 10-20 system [128].

Design concept and technical realization

Standard EEG recording equipment for scientific and medical purposes usually consist of a cap of flexible fabric with electrode placeholders, an amplifier connected to the electrodes via cables, and a computer connected to the amplifier recording and storing the data (e.g. Brain Products actiChamp system). The cap setup allows for flexible and precise positioning of the electrodes and is comfortable to wear. However, despite the high data precision achieved with

these systems, some technical aspects make them non-optimal for recordings in the given clinical populations. Most such systems use gel-based electrodes which require long setup times, exceeding the patient's endurance, and are as such not usable. Alternative systems use dry electrodes (e.g. Guger Technologies g$^{\circledR}$.Nautilus system) which significantly reduce preparation time. However, dry electrodes require firm contact pressure of the electrode on the scalp to yield good conductivity and as such reduce wearing comfort. Moreover, dry electrodes are more susceptible to noise and measurement artifacts than gel-based electrodes. At last, the chin strap for closing and tightening the cap can cause feelings of suffocation and might not be tolerated by some patients. Also, the chinstrap makes signal acquisition prone to artifacts resulting from head and face movements as well as talking. Patients with CP, especially adolescents, are unlikely to sit still for long time. Some systems do not make use of a cap setup, but rather a frame or flexible headband design, such as the QUASAR HMS, or the Emotiv Systems' EPOC. These systems however were also inadequate for the envisaged group, partially because of their obtrusive appearance, partially because of limited flexibility for electrode positioning. None of the commercially available systems fulfilled all requirements which motivated the need for a customized design.

As a compromise between flexible cap and stiff frame setup which does not require a chin strap we decided for a design mimicking headphones (see Figure 48). Headphones have natural unobtrusive appearance as people are familiar with using them for music listening. Furthermore, headphones apply contact pressure around the ears which has proven comfortable to wear and allow for tight sit. In addition, the headphones allow the recording electrodes to be embedded within their headband and consequently to be naturally positioned over motor areas, necessary to assess sensorimotor-related brain activity. As for sensors and electronics we decided to re-use a worn-off Emotiv EPOC device. Despite the original purpose of the Emotiv EPOC for gaming and entertainment, there have been numerous scientific papers making use of the system. The system has been validated with regard to the use for scientific purposes, such as in [185] and [250]. These works demonstrated that despite lower signal quality, the system delivers usable data in a wide spectrum of applications. The Emotiv EPOC sensor technology makes use of a semi-dry solution, namely felt pads soaked with saline solution establishing conductivity between the proband's scalp and the gold-coated electrodes. The electrodes require a certain contact pressure but the soft felt pads allow for comfortable wearing. The Emotiv EPOC headset is a one-size for all frame with a fixed positioning of the electrodes. Its original layout has mainly channels over the pre-frontal areas.

We freed sensors and electronics from the original plastic frame and included the raw modules into our headphones setup. A flexible two-layer sheet made of washable fabric was mounted in between the left and right earpiece, which are held together by two curved metal bars. The two-layer sheet embeds and hides 13 (of original 14) EEG electrodes and the connecting cables to the electronics. One channel was spared for capturing vertical eye-movement via EOG signals (see Figure 48). This allows for either online or post-hoc eye-movement artifact reduction. The right earpiece was used as the housing for electronics, leaving enough

space to fill it with rubber foam for sustaining sound insulation and ensuring no contact of the proband's ear with the electronics. The electronics communicates wirelessly with a transceiver connected via USB to a recording computer. The original reference channels were replaced by longer cables connectable to the proband's left and right mastoids via ECG (electrocardiogram) patches. Miscellaneous parts to assemble the whole system were 3D-printed. The Emotiv EPOC hardware is compatible with open source recording software, such as the Open-Vibe framework [212]. This allows for convenient access to raw signals and high flexibility in experiment design and implementation. Proband preparation takes < 5 minutes; after initial preparation, electrode impedances stay stable for at least 30 minutes. A simple mechanism allows for quick ($<$ 5 minutes) adaptation of the headset to different headsizes (approx. 52-58 cm in diameter). For hygienic reasons, felt pads can be replaced and the electrode sheet as well as the earpieces be sanitized. More information is provided in Table 21.

Usability for measurement of sensorimotor rhythms

The usability of the initial functional prototype (see Figure 48(b)) for measuring sensorimotor-related EEG activity was validated in several experimental tasks including executed and imagined hand/finger-movements. For a systematic assessment of data quality, all experiments were conducted twice: (1) with a research grade EEG system and (2) with our adapted device (*ICS-headset*, initial prototoype). Data was collected from one healthy participant performing the SRTT in two separate sessions. In the experiment, the participant was presented 1 out of 4 possible targets on a computer screen and had to respond with a corresponding right hand key press. In total 40 trials were collected per data set with an inter-trial pause of 10 s. Each trial consisted of the visual cue presentation and the subsequent participant response. The first session was recorded using a Brain Products actiChamp 32-channel gel-based active electrodes setup with 500 Hz sampling rate. All leads were referenced to the average of left and right mastoid and impedances were kept $\leq$5kΩ. The second session took place on a different day and was recorded using our adapted EEG device with 128 Hz sampling rate. Also here, all leads were referenced to the average of the left and right mastoid; electrode connectivity was tested using the Emotiv System's TestBench™ software. Sensors were adjusted until connectivity reached the "green" level (corresponds to an impedance of $\leq$220kΩ according to a test performed by Badcock and colleagues in 2015 [251]. All data processing was carried out in MATLAB®, in part using functions provided by the EEGLAB toolbox [137]. First, the data from the actiChamp device was downsampled to 128 Hz to make it comparable to the data collected with the adapted device. Then, each dataset was highpass filtered using a zero phase Hamming Windowed sinc FIR filter with cutoff frequency of 0.5 Hz. The data was then segmented into epochs, time-locked to the moment of key press. No epochs were rejected for further analyses.

Table 21 Facts and figures of the *ICS-headset*. Some figures (*) were taken from the Emotiv System's EPOC+ specification, see `https://emotiv.com/store/compare/`

Sensors and channels

Number of channels	**18** (13 EEG, 1 EOG, 2 Reference + 2 axis gyrometer)
Sensor technology*	Semi-dry saline soaked felt pads
EEG channel labels (10-20 system)	FCz, Cz, CPz, FC3, C1, C3, C5, CP3, CP5, C2, C4, C6, CP4, EOG, REF1, REF2

Electronics and signal acquisition

Sampling rate*	2048 Hz internal, filtered and downsampled to **256 Hz**
Frequency response*	**0.16 - 43 Hz**
Resolution*	**14 bit** per channel (0.51 μV)
Wireless data transmission*	Emotiv System's proprietary 2.4GHz wireless (custom USB receiver)

Usability

Internal battery power*	Li-poly battery, 680 mAh, $>$ **12 hours**
Maximum distance to wireless transceiver	up to **10 meters**, 1 meter recommended to avoid data packet loss
Stability of electrode conductivity	tested up to **30 minutes** without re-applying saline solution
Proband preparation time	approx. **5 minutes**
Applicability with respect to head size	approx. **52-58cm** diameter
Adjustment to different head sizes	$<$ **5 minutes**
Price per device	in total approx. **US$799** (Emotiv EPOC US$ 699 + $<$US$100 miscellaneous)
Costs for spare parts per data recording / proband	approx. **US$1** (3 single use ECG patches, 13 felt pads)

ERS/ERD was computed from the mu-bandpass and beta-bandpass filtered signals according to Pfurtscheller's method [248]. The upper panels of Figure 49 show the average ERS/ERD time-courses of channel C3 comparing both recording sessions. Furthermore, we performed a time-frequency analysis of channel C3 using the wavelet-based ERSP (event-related spectral perturbation) technique [252], see Figure 49, lower panels. For ERSP computation we used the epoch [-3...3] s with respect to key press. Wavelet parametrization was set to 3 cycles for the lowest frequency (2 Hz) expanding gradually towards half of the number of cycles for the highest frequency (30 Hz). As a result, we observe clear mu-power suppression and traces of beta-power suppression in both datasets. There are two main differences between the two datasets: (1) Motor preparation related ERD onset with respect to the movement onset appeared earlier in the first dataset compared to the second data set. (2) The recovery time (event-related synchronization after motor execution) is longer in the second data set (around 1500 ms) compared to the first data set (around 1000 ms). Whether or not these variations resulted from the different measurement setups or the daily constitution of the participant cannot be stated with certainty. In any case, the reduction in mu power derived from signals collected over motor areas that is expected to occur during preparation of movement is visible in both data sets. This observation, together with the practical advantages of the new EEG system described above, validates and supports the use of the new adapted EEG system to assess changes in mu-power ERS/ERD over motor areas. The results are reported in further detail in [253].
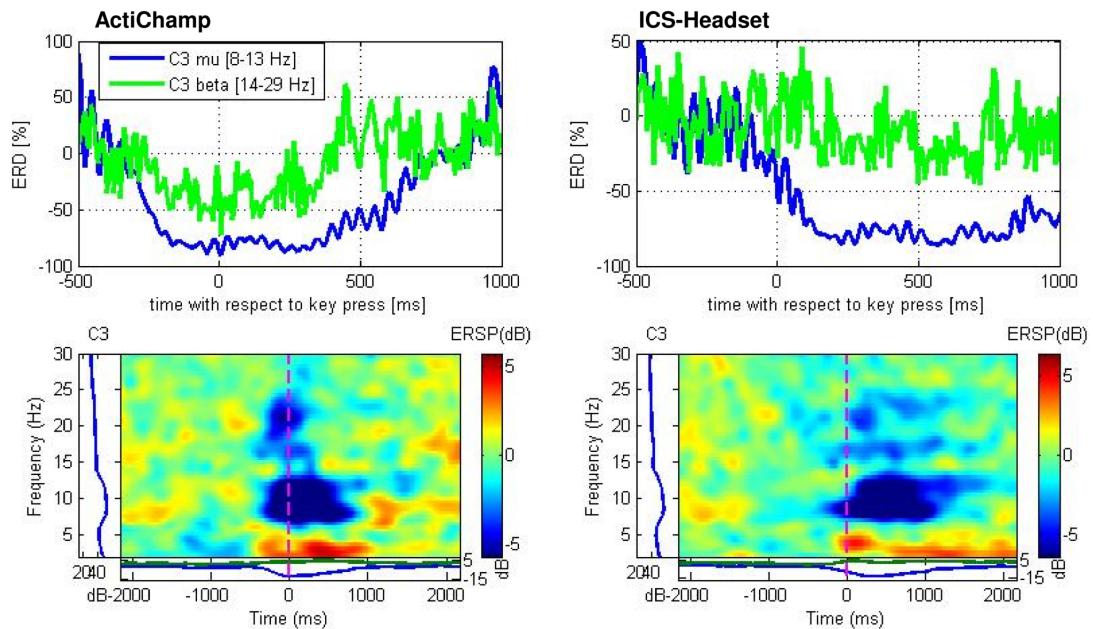
Figure 49 Motor related ERD acquired with research-grade versus ease-of-use recording hardware: Top panels: Analysis of mu- and beta-power ERD in channel C3 according to [248]. Bottom panels: Time-frequency representation of channel C3 time-locked to the participant's key press using the wavelet-based ERSP technique.

The same adapted SRTT was then employed to assess, with the adapted device, hand motor-related ERD in 13 adolescents and adults with and without a diagnosis of CP. Measurements were performed in a normal office at the day rehabilitation centers attended by the patients. Data analysis revealed visibility of effects of motor-related ERD in most patients. Consistent

changes of these effects across recording sessions which could be related to improvements of hand-motor coordination were however not found. More detailed results are reported in [186].

A short questionnaire has been delivered to some of the participants after the EEG recordings to collect the first impressions on the device. Most patients reported the device to be comfortable (5 out of 5 persons), not heavy (3 out of 5) and not looking harmful or dangerous (5 out of 5). None felt bothered by the system, but some were nevertheless aware of it during the recordings (2 out of 5). Researchers on the other hand were very pleased with the easy way the system can be set up and with the short preparation time required, an aspect that they found very helpful when testing all the patients, not only the younger ones.

Usability for measurement and decoding of ErrPs

In order to assess the usability of our adapted EEG device for the measurement and decoding of ErrPs, we compare the observability of ErrPs in two separate datasets. In both datasets, participants performed the same experimental task, but signals were acquired with different recording setups. Subjects were asked to perform an adapted version of the cursor task (see Chapter 4, [45]). Adaptations to the experimental protocol were as follows: feedback in form of cursor-movement was delayed by 500 ms, time locked to the participants' key press response. This adaptation was introduced to temporally disentangle residual motor-related effects due to the key press action from effects related to feedback processing. Furthermore, the experiment time was reduced to 150 trials per participant with an average of 53 error-trials and 97 non-error trials (error-probability of $p_{err} = 35\%$). Dataset ACTICHAMP[1] consists of 18 participants (age: 28.9±4.7, 9 females, 9 males) in total; EEG signals were acquired with a research-grade EEG system (Brain Products actiChamp amplifier with 32-channel gel-based electrodes). Dataset ICS-HEADSET[2] consists of 13 participants (age: 29.8±5.0, 4 females, 9 males)[3]; EEG signals were acquired with the follow-up prototype of the *ICS headset* (see Figure 48(d)). Further information about the recording setups of both datasets are listed in Table 22.

In terms of data pre-processing, both dataset were treated identical whenever possible. All EEG data preprocessing was carried out in MATLAB®, in part using functions provided by the EEGLAB toolbox [137]: In order to remove high frequency and power-line noise, we first filtered the signals of the EEG and EOG channels using a zero phase Hamming windowed sinc FIR band-pass filter with cutoff frequencies of 1 Hz and 20 Hz. Next, we identified and interpolated contaminated EEG channels using kurtosis with a threshold of 5%. EOG activity

---

[1] Available for download here: `https://github.com/stefan-ehrlich/dataset-errp-coadaptation/tree/master/data_cursor`

[2] Available for download here: `https://github.com/stefan-ehrlich/dataset-ICS-EEG-headset/tree/master/data_cursor`

[3] Both datasets were acquired with a different set of participants during different time periods. Although some participants might have taken part in both data collection sessions, no particular care was taken to balance the datasets or identify common participants as this was not expected to affect the results and overall conclusions reported here.

Table 22 Overview and comparison of ACTICHAMP and ICS-HEADSET dataset.

Dataset	ACTICHAMP	ICS-HEADSET
EEG hardware	Brain Products actiChamp system	Emotiv EPOC+ customized headset
Electrodes	32 active gel-based	14 passive semi-dry
EEG channels	27 channels (FP1, FP2, F3, F4, F7, F8, FC1, FC2, FC5, FC6, C3, C4, T7, T8, CP5, CP6, P3, P4, P7, P8, TP9, TP10, O1, O2, Fz, Cz, Pz)	13 channels (FCz, Cz, CPz, FC3, C1, C3, C5, CP3, CP5,C2, C4, C6, CP4)
EOG channels	3 channels, gel-based (forehead, left and right outer canthi)	1 channel, ECG patch (right outer canthi)
Referencing	TP9 + TP10, gel-based (average mastoids)	TP9 + TP10, ECG patches (average mastoids)
Data transmission	cable-based (USB)	wireless data transmission
Sampling rate	1024 Hz	256 Hz
Impedance level	$\leq 10 k\Omega$	$\leq 220 k\Omega$

in the EEG signals (eye-blink and lateral eye movements) was corrected using a method suggested by Schlögl and colleagues [179][4]. Afterwards, EEG signals were re-referenced to CAR to further reduce signal contamination due to external noise sources. The data was further segmented into data epochs [-500...1000] ms for non-error- and error-events time-locked to the moment of feedback presentation (cursor-movement).

Figure 50 shows a comparison of the ERP time-courses over fronto-central channels time-locked to the onset of feedback presentation for each class of events (blue: non-error, red: error) and the difference grand average (black: error minus non-error), including the topographic representation of the difference grand average at relevant time points. Furthermore, between class discriminative power is depicted as r^2-values. In both datasets, the difference grand average shows the ErrP-characteristic N2-P3 complex; spatio-temporal shapes highly resemble between dataset. The ERP amplitudes (effect strength) as well as the peak r^2-values are lower in dataset ICS-HEADSET compared to dataset ACTICHAMP. This result was expected due to lower signal quality delivered by the Emotiv EPOC+ hardware. In terms of the temporal shape of the ErrP difference wave, we observed a latency shift of approx. 100 ms in the ICS-HEADSET compared to the ACTICHAMP dataset. The N200 component appeared around 350 ms (instead of around 250 ms) and the P300 component around 450 ms (instead of around 350 ms) time locked to the feedback. This latency shift is likely related to the delay caused by wireless data transmission of the *ICS-headset*. Delays due of around 40±20ms were also found in another study comparing the wireless MUSE EEG headset with

---

[4] Here, for dataset ACTICHAMP 3 EOG channels; for dataset ICS-HEADSET 1 EOG channel was used.

the actiChamp EEG system [254]. Despite these variations, the results confirmed feasibility to successfully observe ErrPs using the *ICS-headset*.
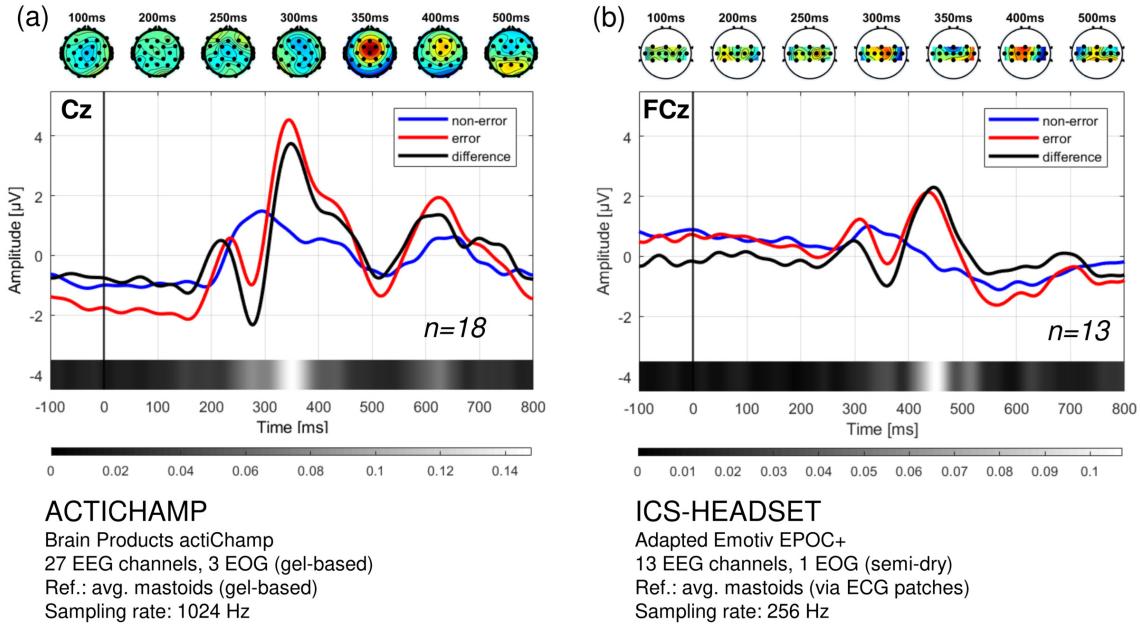

Figure 50 Grand average ERP time-courses over fronto-central channels time-locked to the onset of feedback presentation for each class of events (blue: non-error, red: error) and the difference grand average (black: error minus non-error). The r^2-values for between non-error and machine-error are depicted below the plot, with brighter colors indicating higher values. The difference grand average is furthermore depicted as topographic plots at specific time points above the plot. (a) Results obtained from dataset ACTICHAMP ($n = 18$) and (b) results obtained from dataset ICS-HEADSET ($n = 13$). (Figure style adopted from [141]).

In addition, we compared the single-trial decodability of observed ErrPs using within-participant CV. Here, we followed the procedure of our previous work (see Chapter 5, [46]). First, temporal features were extracted as the arithmetic mean of the signal amplitude in pre-defined windows relative to the moment of feedback presentation. Due to the observed latency shift caused by the wireless data transmission of the *ICS-headset* we decided to extend the time windows[5] to capture also later effects. This resulted in a total of 378 temporal features per epoch (27 channels x 14 windows) for dataset ACTICHAMP and a total of 182 temporal features per epoch (13 channels x 14 windows) for dataset ICS-HEADSET. The features were then used to train a regularized version of the LDA classifier [180] (identical to the method described in Chapter 5, Section 5.2.3.3). Table 22 shows individual participant results for dataset ACTICHAMP with an average cross-validation accuracy of $82.9 \pm 7.1\%$ across 18 participants. These results are in line with the expected decoding performance of 75-85% reported by others and our earlier studies. Table 23 shows individual participant results for dataset ICS-HEADSET with an average cross-validation accuracy of $70.8 \pm 10.6\%$. As expected, the results are lower, but significantly above chance-level[6] in the majority of participants (10 out of 13). Interestingly, a few participants show accuracies close to or above 80% (e.g. s01b, s09b,

---

[5] windows: [150...250] ms, [200...300] ms, [250...350] ms, [300...400 ms, [350...450] ms, [400...500] ms, [450...550] ms, [500...600] ms, [550...650] ms, [600...700] ms, [650...750] ms, [700...800] ms, [750...850] ms, [800-900] ms.

[6] $Acc_{chance} = 60.37\%$ based on inverse cumulative binomial distribution with number of trials $nTrials = 150 * p_{err}$ and $p_{err} = 0.35$; probability of success $p_{success} = 0.5$; confidence threshold $p = 0.05$

s11b, s13b) comparable to accuracies obtained with the actiChamp system. The comparably high between participant standard-deviation is currently unexplained and will be investigated in future works. Possible reasons could be related to variations in the headset placement during the experiment, or variation of participants' head-size or hair-density. Nonetheless, these results confirm successful decoding of ErrPs using the *ICS-headset*.

Table 23 Results of ErrP decoding rates (CV accuracies in [%]) for the ACTICHAMP dataset (32 active gel-based channels using Brain Products actiChamp amplifier).

	s01a	s02a	s03a	s04a	s05a	s06a	s07a	s08a	s09a	s10a
Acc	71.7	86.7	89.1	79.4	75.0	79.2	84.0	92.1	86.4	89.8
λ	0.95	0.6	0.95	0.75	0.95	0.95	0.95	0.95	0.85	0.95

	s11a	s12a	s13a	s14a	s15a	s16a	s17a	s18a	AVG$\pm$SD	
Acc	88.3	78.7	73.6	92.2	81.7	69.9	85.0	89.7	**82.9**$\pm$7.1	
λ	0.55	0.95	0.95	0.95	1.0	0.95	0.95	0.95		

Table 24 Results of ErrP decoding rates (CV accuracies in [%]) for the ICS-HEADSET dataset (data acquisition with 16 passive semi-dry channels *ICS headset*, 2. prototype).

	s01b	s02b	s03b	s04b	s05b	s06b	s07b	s08b	s09b	s10b
Acc	78.7	68.7	55.9	71.0	51.9	58.6	67.1	76.1	79.0	72.9
λ	0.85	0.95	0.65	0.65	0.15	1.0	0.4	0.9	0.85	0.95

	s11b	s12b	s13b	AVG$\pm$SD
Acc	82.7	69.6	88.3	**70.8**$\pm$10.6
λ	0.85	0.8	0.6	

Summary

This chapter described the development and validation of an ease-of-use, unobtrusive EEG device (*ICS-headset*). The device was originally developed having in mind the investigation of motor-related sensorimotor rhythms in patients with CP. To address the needs and requirements of this target group, our device features short preparation times (5 minutes), comfortable wearing up to 30 minutes, unobtrusive looks in the design of headphones for music listening, and electrodes placed over sensorimotor areas. Experimenters reported smooth data collection and overall acceptance of the system among patients. The changes in motor-related ERD over time during executed and imagined movement meet the observations described in the literature, this way supporting the functionality of our adapted EEG device for the assessment of sensorimotor-related measures of brain activity. Most pertinent to this thesis is that the device also proved useful for the measurement and decoding of ErrPs. Decoding accuracies obtained with the *ICS-headset* reached $Acc = 70.8 \pm 10.6\%$ on average across 13

participants with above chance-level accuracies in 10 out of 13 participants. Overall, this confirms usability of the *ICS-headset* for both capturing sensorimotor rhythms and error-related potentials. As such, the device could be a useful trade-off against research-grade EEG systems when targeting applications where practicality is favored over signal quality and decoding performance.

Appendix C: Copyright permissions

Copyright Clearance Center

RightsLink®

Home   Create Account   Help

IEEE
Requesting permission to reuse content from an IEEE publication

Title: When to engage in interaction — And how? EEG-based enhancement of robot's ability to sense social signals in HRI

Conference Proceedings: 2014 IEEE-RAS International Conference on Humanoid Robots

Author: Stefan Ehrlich

Publisher: IEEE

Date: Nov. 2014

Copyright © 2014, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

BACK   CLOSE WINDOW

Title: A neuro-based method for detecting context-dependent erroneous robot action

Conference Proceedings: 2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)

Author: Stefan Ehrlich

Publisher: IEEE

Date: Nov. 2016

Copyright © 2016, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

A computational model of human decision making and learning for assessment of co-adaptation in neuro-adaptive human-robot interaction

Conference Proceedings:
2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)

Author: Stefan K. Ehrlich

Publisher: IEEE

Date: Oct. 2019

Copyright © 2019, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

BACK   CLOSE WINDOW

From:	SCITEPRESS Enquiries <info@scitepress.org>
Sent:	Montag, 24. Juni 2019 12:27
To:	Ehrlich, Stefan
Subject:	AW: get copyright permission

Dear Stefan,

We don't have that stated online. It is stated only on the copyrights.

As my colleague already said we authorize the paper's reprinting for your thesis, as long as all the bibliography information from its publication is there too. Any paper version can be there although we prefer to see there the actual published version.

--

Kind regards,

Cláudia Ferreira

SCITEPRESS Team

SCITEPRESS Office

Avenida de S. Francisco Xavier Lote 7 Cv. C, 2900-616 Setubal - Portugal

Tel.: +351 265 520 184/5

Fax: +351 265 520 186

http://www.scitepress.org

On Wednesday, June 19th 2019, 3:55 pm CEST (+0200), Ehrlich, Stefan wrote:

Dear Vitor,

thank you for your quick reply. Could you point me to a webpage or online document where this is explicitly stated?

Best and thanks,

Stefan

On Wednesday, June 19th 2019, 2:06 pm CEST (+0200), SCITEPRESS Enquiries (info@scitepress.org) wrote:

Dear Stefan,

In this case, the same thing mentioned earlier applies to your PhD thesis, so you can reuse parts as long as correctly cited its origin.

On Wednesday, June 19th 2019, 1:48 pm CEST (+0200), Ehrlich, Stefan wrote:

Dear Vitor Pedrosa,

thank you for your mail. I would like to reuse parts of the following publication in my doctoral thesis:

Ehrlich, S., Alves-Pinto, A., Lampe, R., & Cheng, G. (2017). A simple and practical sensorimotor EEG device for recording in patients with special needs.In *Neurotechnix2017, CogNeuroEng 2017, Symposium on Cognitive Neural Engineering*.

Please let me know if there is anything to consider, if no copyright permission is required.

Best regards,

Stefan

On Wednesday, June 19th 2019, 1:25 pm CEST (+0200), SCITEPRESS Enquiries (info@scitepress.org) wrote:

Dear Stefan K. Ehrlich,

Thank you for your email.
Could you please let me know the copyright permissions that you are thinking in requesting?

If you just want to add the paper to your website or you institute website, that's no problem as long as long as all the bibliography information from its publication is there too. Any paper version can be there although we prefer to see there the actual published version.

On Monday, June 17th 2019, 7:18 pm CEST (+0200), schoenleitnerflorian@web.de wrote:

Dear SiTePress team,

I am the author of the paper:

"*A Simple and Practical Sensorimotor EEG Device for Recording in Patients with Special Needs*"

In Proceedings of the 5th International Congress on Neurotechnology, Electronics and Informatics - CogNeuroEng, 73-79, 2017, Funchal, Madeira, Portugal

and would like to know how to get copyright permissions.

Thank you for your help,

Stefan K. Ehrlich

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Apr 08, 2019

This Agreement between Technische Universität München -- Stefan Ehrlich ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4564310392164
License date	Apr 08, 2019
Licensed Content Publisher	Springer Nature
Licensed Content Publication	International Journal of Social Robotics
Licensed Content Title	A Feasibility Study for Validating Robot Actions Using EEG-Based Error-Related Potentials
Licensed Content Author	Stefan K. Ehrlich, Gordon Cheng
Licensed Content Date	Jan 1, 2018
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	full article/chapter
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	yes
Title	Neuro-adaptive human-robot interaction using error-related potentials
Institution name	Chair for Cognitive Systems, Department for Electronic and Computer Engineering, Technische Universität München
Expected presentation date	Oct 2019
Requestor Location	Technische Universität München Arcisstraße 21
	München, 80290 Germany Attn: Technische Universität München
Total	0.00 EUR

Terms and Conditions

Springer Nature Terms and Conditions for RightsLink Permissions
Springer Nature Customer Service Centre GmbH (the Licensor) hereby grants you a non-exclusive, world-wide licence to reproduce the material and for the purpose and requirements specified in the attached copy of your order form, and for no other use, subject to the conditions below:

1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in

the published version).

If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Where **print only** permission has been granted for a fee, separate permission must be obtained for any additional electronic re-use.

3. Permission granted **free of charge** for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

4. A licence for 'post on a website' is valid for 12 months from the licence date. This licence does not cover use of full text articles on websites.

5. Where **'reuse in a dissertation/thesis'** has been selected the following terms apply: Print rights of the final author's accepted manuscript (for clarity, NOT the published version) for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).

6. Permission granted for books and journals is granted for the lifetime of the first edition and does not apply to second and subsequent editions (except where the first edition permission was granted free of charge or for signatories to the STM Permissions Guidelines http://www.stm-assoc.org/copyright-legal-affairs/permissions/permissions-guidelines/), and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence.

7. Rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com/bookpermissions@springernature.com for these rights.

8. The Licensor's permission must be acknowledged next to the licensed material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

9. Use of the material for incidental promotional use, minor editing privileges (this does not include cropping, adapting, omitting material or any other changes that affect the meaning, intention or moral rights of the author) and copies for the disabled are permitted under this licence.

10. Minor adaptations of single figures (changes of format, colour and style) do not require the Licensor's approval. However, the adaptation should be credited as shown in Appendix below.

Appendix — Acknowledgements:

For Journal Content:
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

For **Advance Online Publication papers:**
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)

For Adaptations/Translations:
Adapted/Translated by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

<u>Note: For any republication from the British Journal of Cancer, the following credit line style applies:</u>

Reprinted/adapted/translated by permission from [**the Licensor**]: on behalf of Cancer Research UK: : [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

For **Advance Online Publication** papers:
Reprinted by permission from The [**the Licensor**]: on behalf of Cancer Research UK: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj. [JOURNAL ACRONYM])

For Book content:
Reprinted/adapted by permission from [**the Licensor**]: [**Book Publisher** (e.g. Palgrave Macmillan, Springer etc) [**Book Title**] by [**Book author**(s)] [**COPYRIGHT**] (year of publication)

Other Conditions:

Version 1.1

Questions? <u>customercare@copyright.com</u> or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

IOP Publishing LICENSE
TERMS AND CONDITIONS

Jun 18, 2019

This is a License Agreement between Technische Universität München -- Stefan Ehrlich ("You") and IOP Publishing ("IOP Publishing") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by IOP Publishing, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	4611961006176
License date	Apr 08, 2019
Licensed content publisher	IOP Publishing
Licensed content title	Journal of Neural Engineering
Licensed content date	Jan 1, 2004
Type of Use	Thesis/Dissertation
Requestor type	Author of requested content
Format	Print, Electronic
Portion	chapter/article
The requesting person/organization is:	Stefan Ehrlich
Title or numeric reference of the portion(s)	Full article
Title of the article or chapter the portion is from	N/A
Editor of portion(s)	N/A
Author of portion(s)	N/A
Volume of serial or monograph.	N/A
Page range of the portion	
Publication date of portion	28.09.2018
Rights for	Main product and any product related to main product
Duration of use	Life of current edition
Creation of copies for the disabled	no
With minor editing privileges	yes
For distribution to	Worldwide
In the following language(s)	Original language of publication
With incidental promotional use	no
The lifetime unit quantity of new product	Up to 499
Title	Neuro-adaptive human-robot interaction using error-related potentials
Institution name	Chair for Cognitive Systems, Department for Electronic and Computer Engineering, Technische Universität München

Expected presentation date	Oct 2019
Billing Type	Invoice
Billing Address	Technische Universität München Arcisstraße 21
	München, Germany 80290 Attn: Technische Universität München
Total (may include CCC user fee)	0.00 USD

Terms and Conditions

TERMS AND CONDITIONS
The following terms are individual to this publisher:

These special terms and conditions are in addition to the standard terms and conditions for CCC's Republication Service and, together with those standard terms and conditions, govern the use of the Works.

As the "User" you will make all reasonable efforts to contact the author(s) of the article which the Work is to be reused from, to seek consent for your intended use. Contacting one author who is acting expressly as authorised agent for their co-author(s) is acceptable.

User will reproduce the following wording prominently alongside the Work:

- the source of the Work, including author, article title, title of journal, volume number, issue number (if relevant), page range (or first page if this is the only information available) and date of first publication. This information can be contained in a footnote or reference note; and
- a link back to the article (via DOI); and
- if practicable, and IN ALL CASES for new works published under any of the Creative Commons licences, the words "© IOP Publishing. Reproduced with permission. All rights reserved"

Without the express permission of the author(s) and the Rightsholder of the article from which the Work is to be reused, User shall not use it in any way which, in the opinion of the Rightsholder, could: (i) distort or alter the author(s)' original intention(s) and meaning; (ii) be prejudicial to the honour or reputation of the author(s); and/or (iii) imply endorsement by the author(s) and/or the Rightsholder.

This licence does not apply to any article which is credited to another source and which does not have the copyright line '© IOP Publishing Ltd'. User must check the copyright line of the article from which the Work is to be reused to check that IOP Publishing Ltd has all the necessary rights to be able to grant permission. User is solely responsible for identifying and obtaining separate licences and permissions from the copyright owner for reuse of any such third party material/figures which the Rightsholder is not the copyright owner of. The Rightsholder shall not reimburse any fees which User pays for a republication license for such third party content.

This licence does not apply to any material/figure which is credited to another source in the Rightsholder's publication or has been obtained from a third party. User must check the Version of Record of the article from which the Work is to be reused, to check whether any of the material in the Work is third party material. Third party citations and/or copyright notices and/or permissions statements may not be included in any other version of the article from which the Work is to be reused and so cannot be relied upon by the User. User is solely responsible for identifying and obtaining separate licences and permissions from the copyright owner for reuse of any such third party material/figures where the Rightsholder is not the copyright owner. The Rightsholder shall not reimburse any fees which User pays for a republication license for such third party content.

User and CCC acknowledge that the Rightsholder may, from time to time, make changes or additions to these special terms and conditions without express notification, provided that

these shall not apply to permissions already secured and paid for by User prior to such change or addition.

User acknowledges that the Rightsholder (which includes companies within its group and third parties for whom it publishes its titles) may make use of personal data collected through the service in the course of their business.

If User is the author of the Work, User may automatically have the right to reuse it under the rights granted back when User transferred the copyright in the article to the Rightsholder. User should check the copyright form and the relevant author rights policy to check whether permission is required. If User is the author of the Work and does require permission for proposed reuse of the Work, User should select 'Author of requested content' as the Requestor Type. The Rightsholder shall not reimburse any fees which User pays for a republication license.

If User is the author of the article which User wishes to reuse in User's thesis or dissertation, the republication licence covers the right to include the Accepted Manuscript version (not the Version of Record) of the article. User must include citation details and, for online use, a link to the Version of Record of the article on the Rightsholder's website. User may need to obtain separate permission for any third party content included within the article. User must check this with the copyright owner of such third party content. User may not include the article in a thesis or dissertation which is published by ProQuest. Any other commercial use of User's thesis or dissertation containing the article would also need to be expressly notified in writing to the Rightsholder at the time of request and would require separate written permission from the Rightsholder.

User does not need to request permission for Work which has been published under a CC BY licence. User must check the Version of Record of the CC BY article from which the Work is to be reused, to check whether any of the material in the Work is third party material and so not published under the CC BY licence. User is solely responsible for identifying and obtaining separate licences and permissions from the copyright owner for reuse of any such third party material/figures. The Rightsholder shall not reimburse any fees which User pays for such licences and permissions.

As well as CCC, the Rightsholder shall have the right to bring any legal action that it deems necessary to enforce its rights should it consider that the Work infringes those rights in any way.

For STM Signatories ONLY (as agreed as part of the STM Guidelines)

Any licence granted for a particular edition of a Work will apply also to subsequent editions of it and for editions in other languages, provided such editions are for the Work as a whole in situ and do not involve the separate exploitation of the permitted illustrations or excerpts.

Other Terms and Conditions:

Permission granted for reuse of article (J. Neural Eng. 15 066014 https://doi.org/10.1088 /1741-2552/aae069) in your thesis.

STANDARD TERMS AND CONDITIONS

1. Description of Service; Defined Terms. This Republication License enables the User to obtain licenses for republication of one or more copyrighted works as described in detail on the relevant Order Confirmation (the "Work(s)"). Copyright Clearance Center, Inc. ("CCC") grants licenses through the Service on behalf of the rightsholder identified on the Order Confirmation (the "Rightsholder"). "Republication", as used herein, generally means the inclusion of a Work, in whole or in part, in a new work or works, also as described on the Order Confirmation. "User", as used herein, means the person or entity making such republication.

2. The terms set forth in the relevant Order Confirmation, and any terms set by the Rightsholder with respect to a particular Work, govern the terms of use of Works in connection with the Service. By using the Service, the person transacting for a republication license on behalf of the User represents and warrants that he/she/it (a) has been duly authorized by the User to accept, and hereby does accept, all such terms and conditions on behalf of User, and (b) shall inform User of all such terms and conditions. In the event such

person is a "freelancer" or other third party independent of User and CCC, such party shall be deemed jointly a "User" for purposes of these terms and conditions. In any event, User shall be deemed to have accepted and agreed to all such terms and conditions if User republishes the Work in any fashion.

3. Scope of License; Limitations and Obligations.

3.1 All Works and all rights therein, including copyright rights, remain the sole and exclusive property of the Rightsholder. The license created by the exchange of an Order Confirmation (and/or any invoice) and payment by User of the full amount set forth on that document includes only those rights expressly set forth in the Order Confirmation and in these terms and conditions, and conveys no other rights in the Work(s) to User. All rights not expressly granted are hereby reserved.

3.2 General Payment Terms: You may pay by credit card or through an account with us payable at the end of the month. If you and we agree that you may establish a standing account with CCC, then the following terms apply: Remit Payment to: Copyright Clearance Center, 29118 Network Place, Chicago, IL 60673-1291. Payments Due: Invoices are payable upon their delivery to you (or upon our notice to you that they are available to you for downloading). After 30 days, outstanding amounts will be subject to a service charge of 1-1/2% per month or, if less, the maximum rate allowed by applicable law. Unless otherwise specifically set forth in the Order Confirmation or in a separate written agreement signed by CCC, invoices are due and payable on "net 30" terms. While User may exercise the rights licensed immediately upon issuance of the Order Confirmation, the license is automatically revoked and is null and void, as if it had never been issued, if complete payment for the license is not received on a timely basis either from User directly or through a payment agent, such as a credit card company.

3.3 Unless otherwise provided in the Order Confirmation, any grant of rights to User (i) is "one-time" (including the editions and product family specified in the license), (ii) is non-exclusive and non-transferable and (iii) is subject to any and all limitations and restrictions (such as, but not limited to, limitations on duration of use or circulation) included in the Order Confirmation or invoice and/or in these terms and conditions. Upon completion of the licensed use, User shall either secure a new permission for further use of the Work(s) or immediately cease any new use of the Work(s) and shall render inaccessible (such as by deleting or by removing or severing links or other locators) any further copies of the Work (except for copies printed on paper in accordance with this license and still in User's stock at the end of such period).

3.4 In the event that the material for which a republication license is sought includes third party materials (such as photographs, illustrations, graphs, inserts and similar materials) which are identified in such material as having been used by permission, User is responsible for identifying, and seeking separate licenses (under this Service or otherwise) for, any of such third party materials; without a separate license, such third party materials may not be used.

3.5 Use of proper copyright notice for a Work is required as a condition of any license granted under the Service. Unless otherwise provided in the Order Confirmation, a proper copyright notice will read substantially as follows: "Republished with permission of [Rightsholder's name], from [Work's title, author, volume, edition number and year of copyright]; permission conveyed through Copyright Clearance Center, Inc. " Such notice must be provided in a reasonably legible font size and must be placed either immediately adjacent to the Work as used (for example, as part of a by-line or footnote but not as a separate electronic link) or in the place where substantially all other credits or notices for the new work containing the republished Work are located. Failure to include the required notice results in loss to the Rightsholder and CCC, and the User shall be liable to pay liquidated damages for each such failure equal to twice the use fee specified in the Order Confirmation, in addition to the use fee itself and any other fees and charges specified.

3.6 User may only make alterations to the Work if and as expressly set forth in the Order Confirmation. No Work may be used in any way that is defamatory, violates the rights of

third parties (including such third parties' rights of copyright, privacy, publicity, or other tangible or intangible property), or is otherwise illegal, sexually explicit or obscene. In addition, User may not conjoin a Work with any other material that may result in damage to the reputation of the Rightsholder. User agrees to inform CCC if it becomes aware of any infringement of any rights in a Work and to cooperate with any reasonable request of CCC or the Rightsholder in connection therewith.

4. Indemnity. User hereby indemnifies and agrees to defend the Rightsholder and CCC, and their respective employees and directors, against all claims, liability, damages, costs and expenses, including legal fees and expenses, arising out of any use of a Work beyond the scope of the rights granted herein, or any use of a Work which has been altered in any unauthorized way by User, including claims of defamation or infringement of rights of copyright, publicity, privacy or other tangible or intangible property.

5. Limitation of Liability. UNDER NO CIRCUMSTANCES WILL CCC OR THE RIGHTSHOLDER BE LIABLE FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL OR INCIDENTAL DAMAGES (INCLUDING WITHOUT LIMITATION DAMAGES FOR LOSS OF BUSINESS PROFITS OR INFORMATION, OR FOR BUSINESS INTERRUPTION) ARISING OUT OF THE USE OR INABILITY TO USE A WORK, EVEN IF ONE OF THEM HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. In any event, the total liability of the Rightsholder and CCC (including their respective employees and directors) shall not exceed the total amount actually paid by User for this license. User assumes full liability for the actions and omissions of its principals, employees, agents, affiliates, successors and assigns.

6. Limited Warranties. THE WORK(S) AND RIGHT(S) ARE PROVIDED "AS IS". CCC HAS THE RIGHT TO GRANT TO USER THE RIGHTS GRANTED IN THE ORDER CONFIRMATION DOCUMENT. CCC AND THE RIGHTSHOLDER DISCLAIM ALL OTHER WARRANTIES RELATING TO THE WORK(S) AND RIGHT(S), EITHER EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. ADDITIONAL RIGHTS MAY BE REQUIRED TO USE ILLUSTRATIONS, GRAPHS, PHOTOGRAPHS, ABSTRACTS, INSERTS OR OTHER PORTIONS OF THE WORK (AS OPPOSED TO THE ENTIRE WORK) IN A MANNER CONTEMPLATED BY USER; USER UNDERSTANDS AND AGREES THAT NEITHER CCC NOR THE RIGHTSHOLDER MAY HAVE SUCH ADDITIONAL RIGHTS TO GRANT.

7. Effect of Breach. Any failure by User to pay any amount when due, or any use by User of a Work beyond the scope of the license set forth in the Order Confirmation and/or these terms and conditions, shall be a material breach of the license created by the Order Confirmation and these terms and conditions. Any breach not cured within 30 days of written notice thereof shall result in immediate termination of such license without further notice. Any unauthorized (but licensable) use of a Work that is terminated immediately upon notice thereof may be liquidated by payment of the Rightsholder's ordinary license price therefor; any unauthorized (and unlicensable) use that is not terminated immediately for any reason (including, for example, because materials containing the Work cannot reasonably be recalled) will be subject to all remedies available at law or in equity, but in no event to a payment of less than three times the Rightsholder's ordinary license price for the most closely analogous licensable use plus Rightsholder's and/or CCC's costs and expenses incurred in collecting such payment.

8. **Miscellaneous.**

8.1 User acknowledges that CCC may, from time to time, make changes or additions to the Service or to these terms and conditions, and CCC reserves the right to send notice to the User by electronic mail or otherwise for the purposes of notifying User of such changes or additions; provided that any such changes or additions shall not apply to permissions already secured and paid for.

8.2 Use of User-related information collected through the Service is governed by CCC's privacy policy, available online here: http://www.copyright.com/content/cc3/en/tools/footer

/privacypolicy.html.

8.3 The licensing transaction described in the Order Confirmation is personal to User. Therefore, User may not assign or transfer to any other person (whether a natural person or an organization of any kind) the license created by the Order Confirmation and these terms and conditions or any rights granted hereunder; provided, however, that User may assign such license in its entirety on written notice to CCC in the event of a transfer of all or substantially all of User's rights in the new material which includes the Work(s) licensed under this Service.

8.4 No amendment or waiver of any terms is binding unless set forth in writing and signed by the parties. The Rightsholder and CCC hereby object to any terms contained in any writing prepared by the User or its principals, employees, agents or affiliates and purporting to govern or otherwise relate to the licensing transaction described in the Order Confirmation, which terms are in any way inconsistent with any terms set forth in the Order Confirmation and/or in these terms and conditions or CCC's standard operating procedures, whether such writing is prepared prior to, simultaneously with or subsequent to the Order Confirmation, and whether such writing appears on a copy of the Order Confirmation or in a separate instrument.

8.5 The licensing transaction described in the Order Confirmation document shall be governed by and construed under the law of the State of New York, USA, without regard to the principles thereof of conflicts of law. Any case, controversy, suit, action, or proceeding arising out of, in connection with, or related to such licensing transaction shall be brought, at CCC's sole discretion, in any federal or state court located in the County of New York, State of New York, USA, or in any federal or state court whose geographical jurisdiction covers the location of the Rightsholder set forth in the Order Confirmation. The parties expressly submit to the personal jurisdiction and venue of each such federal or state court.If you have any comments or questions about the Service or Copyright Clearance Center, please contact us at 978-750-8400 or send an e-mail to info@copyright.com.

v 1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.