



# HHS Public Access

Author manuscript

*Biol Psychiatry Cogn Neurosci Neuroimaging*. Author manuscript; available in PMC 2024 August 01.

Published in final edited form as:

*Biol Psychiatry Cogn Neurosci Neuroimaging*. 2023 August ; 8(8): 805–814. doi:10.1016/j.bpsc.2022.12.005.

## Reliability and validity of TMS-EEG biomarkers

Sara Parmigiani<sup>1,2</sup>, Jessica M. Ross<sup>2,1</sup>, Christopher Cline<sup>1,2</sup>, Christopher Minasi<sup>1,2</sup>, Juha Gogulski<sup>1,2,3</sup>, Corey J Keller<sup>1,2,8</sup>

<sup>1</sup>Department of Psychiatry and Behavioral Sciences, Stanford University Medical Center, Stanford, CA, 94305, USA

<sup>2</sup>Veterans Affairs Palo Alto Healthcare System, and the Sierra Pacific Mental Illness, Research, Education, and Clinical Center (MIRECC), Palo Alto, CA, 94394, USA

<sup>3</sup>Department of Clinical Neurophysiology, HUS Diagnostic Center, Clinical Neurosciences, Helsinki University Hospital and University of Helsinki, Helsinki, FI-00029 HUS, Finland

### Abstract

Noninvasive brain stimulation and neuroimaging have revolutionized human neuroscience, with a multitude of applications including diagnostic subtyping, treatment optimization, and relapse prediction. It is therefore particularly relevant to identify robust and clinically valuable brain *biomarkers* linking symptoms to their underlying neural mechanisms. Brain biomarkers must be reproducible (*i.e.*, have *internal reliability*) across similar experiments within a laboratory and be generalizable (*i.e.*, have *external reliability*) across experimental setups, laboratories, brain regions, and disease states. However, *reliability* (*internal* and *external*) is not alone sufficient; biomarkers also must have *validity*. *Validity* describes closeness to a true measure of the underlying neural signal or disease state. We propose that these metrics, *reliability and validity*, should be evaluated and optimized before any biomarker is used to inform treatment decisions. Here, we discuss these metrics with respect to causal brain connectivity biomarkers from coupling transcranial magnetic stimulation (TMS) with electroencephalography (EEG). We discuss controversies around TMS-EEG stemming from the multiple large off-target components (noise) and relatively weak genuine brain responses (signal), as is unfortunately often the case in noninvasive human neuroscience. We review the current state of TMS-EEG recordings, which consist of a mix of reliable noise and unreliable signal. We describe methods for evaluating TMS-EEG biomarkers, including how to assess *internal* and *external reliability* across facilities, cognitive states, brain networks, and disorders, and how to *validate* these biomarkers using invasive neural recordings or treatment response. We provide recommendations to increase *reliability* and *validity*, discuss lessons learned, and suggest future directions for the field.

<sup>6</sup> Correspondence : Corey Keller, MD, PhD, Stanford University, Department of Psychiatry and Behavioral Sciences, 401 Quarry Road, Stanford, CA 94305-5797, ckeller1@stanford.edu , Phone: +1 8025786292 .

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

#### Disclosures

Dr. Parmigiani, Ross, Cline, and Minasi report no biomedical financial interests or potential conflicts of interest. Dr. Gogulski receives funding from Orion Research Foundation and Finnish Medical Foundation. Dr. Keller receives funding from the National Institutes of Health, the Burroughs Wellcome Fund, and holds equity in Alto Neuroscience, Inc.

## Keywords

TMS; EEG; TEP; reliability; validity; TMS-EEG

---

### 1. **Reliability and validity in neuroimaging**

Noninvasive brain imaging and stimulation has revolutionized human neuroscience over the past thirty years. Many tools exist to image brain activity including functional MRI (fMRI), magnetoencephalography (MEG), and electroencephalography (EEG). Each modality has different limitations in terms of cost and complexity (fMRI, MEG), limited temporal resolution (fMRI, fNIRS), and limited spatial resolution (MEG, EEG). Pairing these imaging techniques with noninvasive brain stimulation can enable the causal study of responses to perturbation in focal regions of cortex and connected networks. These causal methods include using transcranial direct and alternating current stimulation (tDCS and tACS, respectively) and transcranial magnetic stimulation (TMS), and have been critical in accumulating knowledge of and treating neurological and psychiatric disorders. To use noninvasive brain stimulation techniques as part of a biomarker, stimulation responses must be quantified with a metric related to normal neurophysiology, pathophysiology, or responses to an exposure or intervention (1,2). TMS effects on the motor and visual cortices can be captured with corticospinal (electromyography) and visual perception readouts, but TMS to other regions may be best captured using electrophysiology. The combination of TMS with EEG is particularly promising due to the comparable temporal resolution between TMS and EEG and the possibility of improving EEG's poor spatial specificity using the greater spatial specificity of TMS itself (3).

Given the limitations of each of these noninvasive tools, it is critical that the readouts of stimulation-imaging methods are stable and reproducible (*reliability*) and measure the underlying neural processes of interest (*validity*). Reliability refers to the consistency of a biomarker. *Internal reliability* refers to the ability to reproduce the biomarker within a laboratory, and asks the question: *'how well in the lab and for this experimental setup can the study be reproduced?'* Biomarkers also need to have strong *external reliability*, answering the question: *'how well in other laboratories or clinical environments, with different experimental setups and where operators may be trained differently, can the study yield consistent results?'* There are many instances in which a metric can have *validity* but not *internal reliability* or *external reliability* (Figure 1). We propose that it is necessary to critically evaluate and optimize both the *reliability* and *validity* of any brain biomarker of interest, particularly prior to implementation in aiding diagnosis or treatment.

All too often, there is an understandable desire to immediately use recently-discovered human noninvasive tools rather than to first send them through months of rigorous testing. Testing of a tool is time-consuming and often less immediately impactful, while applying that novel tool to study a specific brain circuit or disorder can yield high impact publications, open up future lines of investigation, and be immediately translatable to the clinic. However, if a tool has either low *reliability* or *validity*, clinical application should not proceed until deemed sufficient by the scientific community. Unfortunately, and particularly

for tools not regulated by the FDA, there is no such rigorous statistical barrier to mainstream use. Instead, rigorous testing is often performed after years of application research (4,5). A lack of rigor in biomarker development can in the end impede scientific progress and cast doubt on these measurement tools, over time weakening the scientific community's view on clinical noninvasive human neuroscience (6).

We propose that prior to clinical use measurements derived from noninvasive tools must be rigorously evaluated and deemed to have high *reliability* and *validity*. The field of *psychometrics* is dedicated to the evaluation of scientific metrics of psychological properties, but is most often applied to clinical and behavioral assessments and not regularly to novel neuroimaging biomarkers (7,8). Thus, we must develop these psychometric-like rigorous testing algorithms for novel noninvasive brain measurements prior to clinical use. Below we outline this process for one relatively new noninvasive brain measurement tool, TMS-EEG, and describe how to rigorously assess and optimize TMS-EEG biomarkers. We hope that other research endeavors follow suit in rethinking, reevaluating, and improving biomarkers from their measurement tool of choice.

## 2. Motor-evoked potentials (MEPs): A gold standard

While noninvasive imaging techniques (fMRI, MEG, EEG) measure various aspects of neural activity, they lack a causal component. Causal techniques such as TMS involve perturbing brain activity and measuring the consequent response. Initial TMS studies focused on stimulation of the primary motor cortex (M1) (1–4)(13), the source of much of our knowledge about the physiological effects of TMS in humans (14–16). Thanks to the somatotopic organization of the M1, a muscle of interest can be activated by applying TMS to a specific portion of the M1 (17). Supra-threshold single TMS pulses over M1 elicit a strong electromyography (EMG) response over skeletal muscle, termed the *motor evoked potential* or MEP. The MEP is linearly correlated with the number of activated corticospinal neurons (14) and thus considered *valid* to track corticospinal excitability. Based on many studies, the MEP has been shown to be highly *reliable*, stable across time within a laboratory (*internal reliability*, (18–22)), and generalizable across laboratories (*external reliability*; (11,23–26)), partially due to its high signal-to-noise (>1000 uV responses with low levels of noise). Of course, some theoretical assumptions and approximations are required, but all things considered, the MEP and its features (*e.g.*, latency, amplitude, morphology) have high *internal* and *external reliability* as well as high *validity* to track excitability of the corticospinal tract. As such, it is in widespread use both in research and clinical practice (27–29). Unfortunately, MEPs can only probe the corticospinal tract, and as a result other tools including TMS-EEG are needed to explore causal relationships in other brain regions.

## 3. TMS-evoked potentials (TEPs)

Although relatively new, TMS-EEG is a powerful noninvasive neuroimaging tool with strong *internal reliability* (also known as *reproducibility*) (30–32). However, more rigorous *external reliability* and *validity* assessment with subsequent optimization is needed. TMS-evoked potentials (TEPs) are a result of coupling single pulses of TMS with scalp EEG recording. TEPs were first described in 1997 (30). Foundational studies in motor cortex

physiology, cortical excitability related to movement, and stimulation intensity followed (34,35). TEPs became more relevant as a potential clinical tool when they were used to help understand sleep physiology (36,37). In contrast to MEPs, which probe the entire corticospinal tract, TEPs measure the central nervous system response to single pulses of TMS without requiring a motor read-out. TEPs consist of a multiphasic response lasting ~500 ms (31). Initial reports described high *internal reliability* (Figure 2; see (30)), demonstrating that TEPs can be repeatable after a week and are sensitive to changes in stimulation amplitude, site, and angle (30). Indeed, although outside the scope of this review, it should be noted that TEPs are also highly *sensitive* to small but relevant changes in stimulation site, angle, and intensity (30). Subsequently, TEPs were largely viewed as plug-and-play and have since been applied to various cognitive states and connectivity features (38,39), brain disorders (40,41), stimulation sites (42), stimulator devices (43), amplifier types (43), and EEG recording setups (44).

However, shortly after dissemination, the assumed high *external reliability* began to be questioned. Initial studies were performed by stimulating medial structures with minimal muscle artifact (30,36,45), which may explain the high reports of *internal reliability*. It was eventually determined that 1) stimulation to lateral brain regions elicits strong early muscle artifacts that can confound the TEP (46,47), and 2) the sensation and auditory click from a TMS pulse results in non-specific sensory evoked responses (48). In addition, while *sample and hold* amplifiers in early TEP reports reduced artifacts during and directly following the TMS pulse, due to their high cost, less expensive DC amplifiers were often purchased by other groups, introducing additional differences across laboratories. As a result, controversies ensued when TEPs across laboratories reported different levels of *internal reliability* of the TEP. In recent years, multiple groups have re-evaluated the *internal reliability* of TEPs (32,49,50). Previously demonstrated substantial *internal reliability* for later components (>50 ms) of the TEP are now known to consist partially of sensory non-specific off-target effects (48), and weaker *internal reliability* for earlier components (<50 ms) of the TEP are thought to reflect more *valid* components of local cortical excitability (51). Discussion of the *reliability* and *validity* of these early and late components of the TEP collided when similar TEP morphologies were reported after single pulses of sham (placebo) TMS (6) and single pulses of TMS to the shoulder (52). In response, a discussion ensued within the TMS-EEG scientific community questioning the most appropriate methods for obtaining TEPs with high *reliability* and *validity* (Figure 3, (50)). This discussion highlighted the need for 1) rigorous assessment with each new TMS coil, location, angle, and intensity probed, as well as 2) standardization across the TMS-EEG community with respect to *reliability* and *validity* assessments.

To move towards robust measures of causal brain excitability, particularly for extended use to clinical applications, we evaluate in the sections below the off-target and intended neural signals present in the TEP, discuss these in reference to *reliability* and *validity* metrics, and outline methods to assess and enhance these metrics.

#### 4. Off-target components (noise) in the TEP: reliable but not valid

The largest amplitude components in the TEP are usually *reliable* but not *valid*. Because the TEP is the result of single TMS pulses averaged across multiple trials (53,54), it is an aggregate of all electrical signals measurable using EEG. Contributing sources to the TEP have overlapping time courses and are thus susceptible to net displacement effects of individual waves (*i.e.*, constructive and destructive interference). In addition to multiple central neural sources (33,33,42,55), contributions to this aggregate waveform also include non-neural sources (56–60) and peripherally-evoked neural sources unrelated to the direct effects of but time-locked and in response to TMS (6,52,61–64). TMS-evoked central neural sources will be referred to hereafter as *signal*, and non-neural sources, peripherally-evoked neural sources, and off-target neural sources will be referred to as *noise*.

Myriad sources of *noise* confound interpretation of the TEP. Non-neural contributions include 1) pulse artifact from current flow through the coil, 2) recharge artifact from the capacitor between discharges, 3) decay artifact from changes in capacitance between the electrodes and gel as well as between the gel and the scalp, 4) electrode noise from poor contact with the scalp or conductive gel drying or leaking, 5) line noise from equipment in the room (60Hz or 50Hz), 6) evoked and continuous muscle activity, 7) electrocardiogram (EKG), 8) eye blinks, eye movements, microsaccades, and 9) movement of the electrodes (often where the coil is resting on the cap) (57,60,65–67). Peripherally-evoked contributions to the TEP include 1) somatosensory-evoked potentials from the sensation of the TMS pulse on peripheral nerves and 2) auditory-evoked potentials from the auditory click of the TMS pulse (6,48,51,52,61,63,68). Off-target central neural sources can also occur due to 1) inhibitory mechanisms, including those associated with visual field changes from eye blinks, eye movements and microsaccades (69,70), 2) fluctuations in brain state such as awakesness and attention (63,69,71), and 3) pain-related responses (72). Best approaches for handling these many off-target contributions are the topics of numerous other publications and methodological debates (5,60,68,73–76). However, it should be emphasized that these non-neural, peripherally-evoked, and central off-target components of the TEP are often *reliable*, with large amplitudes and long durations (73,77), and therefore should be carefully considered and minimized using a combination of techniques during data collection (36,48,51,61,78,79) and during post-processing (5,52,60,68,77). In summary, there are numerous sources of *reliable* and easily recognizable large amplitude *noise* contributions to the TEP that need to be accounted for when designing a study, minimized during data collection, and if needed removed during data analysis to uncover the effective *signal* of interest.

#### 5. True TEP components (signal): valid but unreliable

While obtaining true neural TEP *signal* is challenging due to the multitude of large magnitude off-target recordings often also present in the TEP (*noise*; see Section 4 and (47,80)), the observation of *valid* components of the TEP from locally-excited brain tissue is possible and can provide important insights into brain physiology and pathology. When TMS is applied to brain regions where off-target noise confounds are minimal, there is strong evidence for the neural basis (*validity*) of TEPs, especially in the early (<50 ms)

components (51). For example, in primary motor cortex (M1), the MEP correlates with the amplitude of the early 15–30 ms response of the TEP (81,82), suggesting that early TEP components reflect cortical excitability. Moreover, local high-frequency TEPs evoked by M1 stimulation and paired-pulse MEPs (conditioned MEPs) share a similar time course, suggesting that they may index similar types of cortical activity (83). Furthermore, TMS to the prefrontal cortex (PFC) produces TEP peaks (84) with amplitudes (85) and latencies (86) that correlate with M1 TEP amplitudes. However, TEPs evoked by PFC and M1 stimulation may reflect different underlying neural mechanisms: early TEP peaks (<50 ms) following left PFC stimulation are insensitive to excitatory NMDA receptor blockade (87), whereas early M1 TEP peaks reflect a balance of GABAergic inhibition and glutamatergic excitation (88). Moreover, corticothalamic frequencies evoked by TMS differ depending on the targeted cortical area (42). In addition to TEP peaks, other TMS-EEG metrics such as interhemispheric signal propagation (ISP) (90) and interhemispheric balance (IHB), which reflect interhemispheric dynamics, and paired-pulse TMS metrics such as long-interval cortical inhibition (LICI) (89) may be useful especially clinical populations for whom MEP measurements are not possible, as in the case of pathological states such as suicidal depression (91). Complexity based metrics also appear to be promising for biomarker development, including but not limited to the Perturbational Complexity Index (PCI) (89,92) and the Similarity Index (SI) (68). TEPs have been examined in various neurological and psychiatric disorders (for review see (54)) these studies have largely focused on later (>100 ms) components of the TEP, where the peripherally-evoked off-target effects are known to be present and have large amplitude, confounding the *validity* of these biomarkers. In summary, although there is evidence suggesting *validity* of the early components of the TEPs, the off-target neural sources and non-neural artifacts surrounding these early components hinders *reliability* and exploration as potential biomarkers of diagnosis and treatment outcome. Thus, increasing the amplitude of these early TEPs and removing artifact and other off-target effects is a critical next step to boost both *reliability* and *validity* of the TEP.

## 6. Evaluation of TEP reliability

A *reliable* TMS-EEG biomarker should be stable across time (93,94) when there is no change in the indicated disease process, and should be sensitive to change when fundamental parameters including instrument, recording set-up, cortical area stimulated, or intensity of stimulation are varied (30). *Reliability* of a biomarker can be evaluated by: (i) measuring the same metric over time within a laboratory (minutes (41), hours (93), or weeks (25,26,76) in a test-retest fashion (25,28,69,74)); (ii) repeated quantification across different instruments or in different laboratories (49,52,72); or (iii) against different preprocessing and analytic pipelines (5,65,68). Described in more detail below, we find that *internal reliability* can be high when using certain set-ups (26) and when applying TMS to medial brain regions with minimal muscle activation. *Reliability* is reduced with the presence of *noise* sources, including when more lateral sites with larger muscle activations are stimulated. Methods to remove sources of *noise* and boost *reliability* include experimentally minimizing sensory contributions online (68) and offline (56), and should be considered in reference to *reliability* metrics.

## 7. Boosting TEP *reliability* during data collection

With a thorough understanding of the factors contributing to *noise* and *signal* in the TEP, various methodological approaches during assessment can be employed to boost the signal to noise ratio (SNR) and improve *reliability*. EEG amplifiers with specialized features (high sampling rate, high dynamic range, slew-rate limiting, and/or sample-and-hold circuitry) can reduce the effect of the primary electromagnetic pulse artifact on the recorded EEG signals (96–98). Delaying stimulator capacitor recharge until after the TEP time period of interest can prevent recharge-related electrical noise from masking relevant EEG signals (99). Minimizing EEG electrode impedance (<5 kOhms) can reduce decay artifact caused by charge buildup at the electrode-gel-skin interface (100). Active electrodes can reduce sensitivity to environmental electrical noise such as 50 or 60 Hz line noise when compared with passive electrodes (101). A thin layer of foam placed between the coil and scalp may reduce artifacts related to bone conduction of the TMS sound (79). Passive noise reduction with earmuffs (48,51) and active noise masking (61,79) minimize auditory sensation and saliency of pulses. In combination with noise reduction and masking, alterations in pulse timing can further reduce off-target EEG components related to sensation and saliency (48). When feasible, rearranging electrode wires can minimize TMS-induced electrical artifact in electrodes near the site of stimulation (44).

Undesirable *noise* and desirable *signal* can be highly sensitive to stimulation location, coil angle, and intensity. By quantifying some aspects of *noise* and *signal* in real-time, it is feasible to individualize stimulation parameters (location, angle, intensity) to directly maximize the SNR of specific TEP features (30,102). Such real-time individualization approaches hold great promise for improving the *reliability* of TEP-based biomarkers by minimizing off-target effects and maximizing local brain responses. These approaches may also improve the *validity* of TEP-based biomarkers by effectively targeting stimulation to more relevant brain circuits with greater fidelity (103).

## 8. Evaluation of TEP *validity* using noninvasive tools

As stated above, early components (<50 ms) of the TEP may represent *valid* metrics of local cortical activation but are currently confounded by multiple sources of *noise*, which in turn reduces both *internal* and *external reliability*. The currently low accuracy of the early TEP with potential for high *validity* is deeply influenced by SNR, making an otherwise valid biomarker invalid (5). In the TEP, underlying signals of local activation can be masked by *noise* (104). It is worth noting that off-target sensory responses and true and valid TEP may not be independent and linearly separable phenomena. Thus, standard methods to remove these sensory responses and non-neural artifacts from genuine brain responses (105) as well as statistical comparisons between real and sham TMS (51,71,79,106–108) may be called into question and reveal the intrinsic challenge about the nature of TEP analysis. These limitations do not necessarily imply that the early TEP cannot be a valid biomarker, but rather that it requires further investigation, mechanistic understanding, sharing of data and protocols, and optimization of experimental setup to maximize SNR. Indeed, if we consider that *validity* describes how well a tool is sampling the desired physiology, in terms of local excitability, data gathered in the last few decades tells a promising while

not complete story of the *possible validity* of early (<50 ms) components of the TEP. Indeed, early TEPs correlate with corticospinal tract excitability (82,109), are reduced in schizophrenia (110,111), mimics characteristic slow waves in NREM sleep, deep sedation, and disorders of consciousness (36,45,112–114), and are modulated in recovery from injury or stroke (115,116). To definitively link TEPs to their underlying neural correlates, however, intracranial brain recordings show promise.

## 9. Evaluation of TEP *validity* using invasive brain recordings

To evaluate the *validity* of a tool, in addition to linking novel neuroimaging metrics with well-known noninvasive measures, it is also important to link to intracranial neurophysiology. Stated another way, using insights gained from intracranial recordings as the ‘ground truth’ may be valuable to inform noninvasive TMS studies and establish the validity of TEP components. Intracranial neurophysiology can be valuable in a number of ways, including 1) TMS-evoked intracranial electrophysiology, 2) electrical stimulation-evoked electrophysiology, and 3) simultaneous invasive and noninvasive brain recordings. First, recent novel work combining *TMS with intracranial brain recordings* suggests that single pulses of TMS modulates both local and downstream brain circuitry that can be captured in the TMS-evoked intracranial response (117). Second, examining the intracranial neural responses to *intracranial electrical stimulation* can also provide valuable ‘ground truth’ information (118–122). As intracranial electrical stimulation is not perceived, the ground truth cortico-cortical response profile to electrical stimulation without sensory confounds can be compared to the noninvasive localization and morphology of TEPs. In this manner, one can identify and isolate TMS-evoked sensory responses from grounded TMS-evoked neural signals that are consistent with intracranial recordings. Finally, in contrast to the previous non-time-locked evaluations of noninvasive and intracranial measurements, a novel combination of simultaneous intracranial and EEG scalp recording after stimulation (13) can provide direct assessment of the neural correlates of noninvasive biomarkers. Caveats with this intracranial approach are that the patient population is currently limited to epileptic surgery patients and that there are physiological and artifactual differences between electrical and magnetic stimulation. In summary, pairing noninvasive and invasive brain stimulation with intracranial recording methods have great potential for improving the *validity* of TEPs and other novel neuroimaging metrics.

## 10. Boosting TEP *validity* using offline analytic methods

In addition to online optimization during data collection, offline analytical methods can be employed to remove *noise* and thus boost the SNR and *validity* of the TEP. There are many preprocessing pipelines that have been developed for this purpose (74–76,123), all of which involve a similar set of preprocessing steps: removal of 1) large TMS pulse artifacts, 2) line noise, 3) other known non-neural artifacts such as from eye movement and blinks, and 4) off-target neural sources (68), such as somatosensory and auditory activations. Removal of this *noise* is approached mainly using a combination of interpolation, filters, and independent component analysis (ICA) based methods (73,124). Interpolation is useful for artifacts with known timing and overwhelmingly high amplitude, such as the primary TMS pulse artifact, but inherently causes loss of data in the interpolated time range. This



is why this technique is only typically used when the artifact overwhelms the overlapping EEG data. Filters break the signal into its spectral components and remove those outside of the specified range to effectively remove line noise, drift, and high-frequency muscle and environmental electrical noise. ICA is a blind source separation technique typically used for removal of eye-blink and saccade artifacts, and by many groups for removal of other off-target activations. ICA is used to decompose the EEG into independent sources of activity that are linearly mixed and is followed with identification of artifactual sources using known spatial and temporal characteristics and removal (73). This process can either be done manually or automatically using machine learning classification algorithms (76,125). In summary, offline removal of known artifacts in the TEP can boost SNR and enhance *validity* of the TEP.

After removing known *noise*, several further decisions regarding how to quantify the TEP are necessary, each of which can affect its *validity*. The TEP waveform is a complex multi-phasic response and quantifying this response is not straightforward. Many decisions are necessary including which peaks to quantify (P30, N45, P60, N100, and/or P200), how to quantify a peak (peak amplitude, peak-to peak-difference, area under the curve, latency, etc.), which electrodes to analyze (single electrode, multiple electrodes, sensor or source space), and whether to evaluate in time or frequency space (42,82). All of these decisions can greatly influence the *validity* of the TEP, since each metric extracts different information encoded in the TEP signal. For example, in assessing TEP change based on late TEP peaks (N100, P200), see also (71), one might find more *reliable* (yet less *valid*) change since later TEP peaks are significantly composed of auditory-evoked potentials caused by the clicking sound produced by the TMS machine (6). Thus, in order to ask targeted questions regarding how the brain responds to the TMS probing of targeted neural circuits, it is essential that the metrics used to quantify TEP signals during analysis encode targeted and *valid* neural responses. Basic statistical techniques comparing the TEP after real and sham rTMS have been utilized to determine whether certain features, such as the late TEP peaks (107), represent *valid* measures of targeted neural activity. However, myriad features can be extracted from the TEP, and basic statistical techniques are not ideal for exploratory investigation across many of them. As such, more sophisticated statistical techniques (such as machine learning methods) could be helpful to describe multiple TEP features in a data-driven manner. These data-driven feature exploration approaches can be helpful for future hypothesis generation. In summary, careful consideration of offline analytic decisions and incorporation of more modern data-driven approaches can greatly influence the *validity* of TEPs.

## 11. Recommendations and future directions

In this paper, we evaluate TMS-EEG, a powerful causal neuroimaging tool, to describe and discuss the concepts of *reliability* and *validity* of potential biomarkers (see also (126) and (127) for discussions regarding TMS-EEG related biomarkers). The TMS field started with a ‘gold standard’ MEP approach that was highly *reliable* (both *internally* and *externally*) and *valid*. As the field incorporated TMS-EEG approaches to evaluate cortico-cortical interactions outside of the motor cortex, initial groups reported high *internal reliability*. However, as TMS-EEG was adopted across labs and translated to different brain regions

and experimental setups, *external reliability* and *validity* were both reduced. We speculate that the trajectory of investigating neurophysiological responses to TMS, from initial high *internal reliability* and *validity* to reduced *external reliability* and *validity*, is all-too-common in noninvasive neuroimaging. We hope that by adopting the recommended approach outlined below, this trajectory can be modified to produce tools with high *reliability* and *validity*.

To help guide further research into the *reliability* and *validity* of novel and established neuroimaging techniques, we provide some general recommendations:

- Although *reliability* and *validity* are related, a biomarker can have high *reliability* but low *validity* or high *validity* but low *reliability* (Figure 1). We suggest that careful consideration of both are necessary before implementing a biomarker to make diagnostic or treatment decisions.
- *Internal reliability* can be enhanced by optimizing the SNR during data collection. Hardware optimization, real-time search for optimal stimulation parameters, and closed-loop methodologies can enormously reduce *noise* and increase *signal* strength.
- Although initial reports of *internal reliability* may be high, we recommend immediate or very early assessment of *external reliability* by collaborating closely with one or more external labs. Further, prompt dissemination and data sharing is of critical importance to ensure consistently high *external reliability* across groups and experimental setups.
- *Validity* is difficult to assess using noninvasive neuroimaging methodologies. Intracranial methods (animal models, human intracranial EEG) can provide ground truth assessment to improve evaluation of biomarker *validity*. Furthermore, simultaneous scalp and intracranial EEG may allow for direct comparisons between noninvasive biomarkers and their neural correlates (13). We strongly recommend that investigators initiate collaborations with labs using these intracranial methods to assess biomarker *validity*.

## Acknowledgements

This work was supported by R01MH126639, R01MH129018, and the Burroughs Wellcome Fund Career Award for Medical Scientists. JR was supported by the Department of Veterans Affairs Office of Academic Affiliations Advanced Fellowship Program in Mental Illness Research and Treatment, the Medical Research Service of the Veterans Affairs Palo Alto Health Care System, and the Department of Veterans Affairs Sierra-Pacific Data Science Fellowship. JG was supported by personal grants from Orion Research Foundation and Finnish Medical Foundation.

## References

1. Califf RM (2018): Biomarker definitions and their applications. *Exp Biol Med* Maywood NJ 243: 213–221.
2. FDA-NIH Biomarker Working Group (2016): BEST (Biomarkers, EndpointS, and Other Tools) Resource. Silver Spring (MD): Food and Drug Administration (US). Retrieved July 13, 2022, from <http://www.ncbi.nlm.nih.gov/books/NBK326791/>
3. Lioumis P, Rosanova M (2022): The role of neuronavigation in TMS-EEG studies: Current applications and future perspectives. *J Neurosci Methods* 380: 109677. [PubMed: 35872153]

4. Rogasch NC, Biabani M, Mutanen TP (2022): Designing and comparing cleaning pipelines for TMS-EEG data: A theoretical overview and practical example. *J Neurosci Methods* 371: 109494. [PubMed: 35143852]
5. Bertazzoli G, Esposito R, Mutanen TP, Ferrari C, Ilmoniemi RJ, Miniussi C, Bortoletto M (2021): The impact of artifact removal approaches on TMS-EEG signal. *NeuroImage* 239: 118272. [PubMed: 34144161]
6. Conde V, Tomasevic L, Akopian I, Stanek K, Saturnino GB, Thielscher A, et al. (2019): The non-transcranial TMS-evoked potential is an inherent source of ambiguity in TMS-EEG studies. *NeuroImage* 185: 300–312. [PubMed: 30347282]
7. Salmond SS (2008): Evaluating the reliability and validity of measurement instruments. *Orthop Nurs* 27: 28–30. [PubMed: 18300685]
8. Kerlinger FN (1966): *Foundations of Behavioral Research*. Holt, Rinehart and Winston: New York.
9. Rothwell JC, Thompson PD, Day BL, Boyd S, Marsden CD (1991): Stimulation of the human motor cortex through the scalp. *Exp Physiol* 76: 159–200. [PubMed: 2059424]
10. Kobayashi M, Pascual-Leone A (2003): Transcranial magnetic stimulation in neurology. *Lancet Neurol* 2: 145–156. [PubMed: 12849236]
11. Rossini PM, Burke D, Chen R, Cohen LG, Daskalakis Z, Di Iorio R, et al. (2015): Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: Basic principles and procedures for routine clinical and research application. An updated report from an I.F.C.N. Committee. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 126: 1071–1107.
12. Thielscher A, Wichmann FA (2009): Determining the cortical target of transcranial magnetic stimulation. *NeuroImage* 47: 1319–1330. [PubMed: 19371785]
13. Parmigiani S, Mikulan E, Russo S, Sarasso S, Zauli FM, Rubino A, et al. (2022): Simultaneous stereo-EEG and high-density scalp EEG recordings to study the effects of intracerebral stimulation parameters. *Brain Stimulat* 15: 664–675.
14. Cattaneo L (2017): Transcranial Magnetic Stimulation. In: Rogers LJ, Vallortigara G, editors. *Lateralized Brain Functions: Methods in Human and Non-Human Species*. New York, NY: Springer, pp 369–406.
15. Huang Y-Z, Edwards MJ, Rounis E, Bhatia KP, Rothwell JC (2005): Theta burst stimulation of the human motor cortex. *Neuron* 45: 201–206. [PubMed: 15664172]
16. Kujirai T, Caramia MD, Rothwell JC, Day BL, Thompson PD, Ferbert A, et al. (1993): Corticocortical inhibition in human motor cortex. *J Physiol* 471: 501–519. [PubMed: 8120818]
17. Rossi S, Pasqualetti P, Tecchio F, Sabato A, Rossini PM (1998): Modulation of corticospinal output to human hand muscles following deprivation of sensory feedback. *NeuroImage* 8: 163–175. [PubMed: 9740759]
18. Proessel F, Beckner ME, Sinnott AM, Eagle SR, LaGoy AD, Conkright WR, et al. (2021): Reliability of corticospinal excitability estimates for the vastus lateralis: Practical considerations for lower limb TMS task selection. *Brain Res* 1761: 147395. [PubMed: 33662340]
19. Welch JF, Argento PJ, Mitchell GS, Fox EJ (2020): Reliability of diaphragmatic motor-evoked potentials induced by transcranial magnetic stimulation. *J Appl Physiol Bethesda Md* 129: 1393–1404.
20. Roth Y, Pell GS, Zangen A (2010): Motor evoked potential latency, motor threshold and electric field measurements as indices of transcranial magnetic stimulation depth. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 121: 255–258; author reply 258–259.
21. Awiszus F (2012): On relative frequency estimation of transcranial magnetic stimulation motor threshold. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 123: 2319–2320.
22. Rossini PM, Rossi S, Pasqualetti P, Tecchio F (1999): Corticospinal excitability modulation to hand muscles during movement imagery. *Cereb Cortex N Y N* 1991 9: 161–167.
23. Ah Sen CB, Fassett HJ, El-Sayes J, Turco CV, Hameer MM, Nelson AJ (2017): Active and resting motor threshold are efficiently obtained with adaptive threshold hunting. *PLoS One* 12: e0186007. [PubMed: 28982146]
24. Groppa S, Oliviero A, Eisen A, Quartarone A, Cohen LG, Mall V, et al. (2012): A practical guide to diagnostic transcranial magnetic stimulation: report of an IFCN committee. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 123: 858–882.

25. Bigoni C, Cadic-Melchior A, Vassiliadis P, Morishita T, Hummel FC (2022): An automatized method to determine latencies of motor-evoked potentials under physiological and pathophysiological conditions. *J Neural Eng* 19. 10.1088/1741-2552/ac636c
26. Nagle KJ, Emerson RG, Adams DC, Heyer EJ, Roye DP, Schwab FJ, et al. (1996): Intraoperative monitoring of motor evoked potentials: a review of 116 cases. *Neurology* 47: 999–1004. [PubMed: 8857734]
27. Prabhu G, Voss M, Brochier T, Cattaneo L, Haggard P, Lemon R (2007): Excitability of human motor cortex inputs prior to grasp. *J Physiol* 581: 189–201. [PubMed: 17332001]
28. Bologna M, Suppa A, Conte A, Latorre A, Rothwell JC, Berardelli A (2016): Are studies of motor cortex plasticity relevant in human patients with Parkinson's disease? *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 127: 50–59.
29. Tolmacheva A, Savolainen S, Kirveskari E, Lioumis P, Kuusela L, Brandstack N, et al. (2017): Long-Term Paired Associative Stimulation Enhances Motor Output of the Tetraplegic Hand. *J Neurotrauma* 34: 2668–2674. [PubMed: 28635523]
30. Casarotto S, Romero Lauro LJ, Bellina V, Casali AG, Rosanova M, Pigorini A, et al. (2010): EEG responses to TMS are sensitive to changes in the perturbation parameters and repeatable over time. *PLoS One* 5: e10281. [PubMed: 20421968]
31. Lioumis P, Kici D, Savolainen P, Mäkelä JP, Kähkönen S (2009): Reproducibility of TMS-Evoked EEG responses. *Hum Brain Mapp* 30: 1387–1396. [PubMed: 18537115]
32. Kerwin LJ, Keller CJ, Wu W, Narayan M, Etkin A (2018): Test-retest reliability of transcranial magnetic stimulation EEG evoked potentials. *Brain Stimulat* 11: 536–544.
33. Ilmoniemi RJ, Virtanen J, Ruohonen J, Karhu J, Aronen HJ, Näätänen R, Katila T (1997): Neuronal responses to magnetic stimulation reveal cortical reactivity and connectivity. *Neuroreport* 8: 3537–3540. [PubMed: 9427322]
34. N Vv, K D, K S, I Rj (2003): Modulation of electroencephalographic responses to transcranial magnetic stimulation: evidence for changes in cortical excitability related to movement. *Eur J Neurosci* 18. 10.1046/j.1460-9568.2003.02858.x
35. Komssi S, Kähkönen S, Ilmoniemi RJ (2004): The effect of stimulus intensity on brain responses evoked by transcranial magnetic stimulation. *Hum Brain Mapp* 21: 154–164. [PubMed: 14755835]
36. Massimini M, Ferrarelli F, Huber R, Esser SK, Singh H, Tononi G (2005): Breakdown of cortical effective connectivity during sleep. *Science* 309: 2228–2232. [PubMed: 16195466]
37. Massimini M, Ferrarelli F, Esser SK, Riedner BA, Huber R, Murphy M, et al. (2007): Triggering sleep slow waves by transcranial magnetic stimulation. *Proc Natl Acad Sci U S A* 104: 8496–8501. [PubMed: 17483481]
38. Zazio A, Bortoletto M, Ruzzoli M, Miniussi C, Veniero D (2019): Perceptual and Physiological Consequences of Dark Adaptation: A TMS-EEG Study. *Brain Topogr* 32: 773–782. [PubMed: 31076949]
39. Bortoletto M, Veniero D, Thut G, Miniussi C (2015): The contribution of TMS-EEG coregistration in the exploration of the human cortical connectome. *Neurosci Biobehav Rev* 49: 114–124. [PubMed: 25541459]
40. Voineskos D, Blumberger DM, Zomorodi R, Rogasch NC, Farzan F, Fousias G, et al. (2019): Altered Transcranial Magnetic Stimulation-Electroencephalographic Markers of Inhibition and Excitation in the Dorsolateral Prefrontal Cortex in Major Depressive Disorder. *Biol Psychiatry* 85: 477–486. [PubMed: 30503506]
41. Canali P, Sarasso S, Rosanova M, Casarotto S, Sferrazza-Papa G, Gosseries O, et al. (2015): Shared reduction of oscillatory natural frequencies in bipolar disorder, major depressive disorder and schizophrenia. *J Affect Disord* 184: 111–115. [PubMed: 26074020]
42. Rosanova M, Casali A, Bellina V, Resta F, Mariotti M, Massimini M (2009): Natural frequencies of human corticothalamic circuits. *J Neurosci Off J Soc Neurosci* 29: 7679–7685.
43. Varone G, Hussain Z, Sheikh Z, Howard A, Boulila W, Mahmud M, et al. (2021): Real-Time Artifacts Reduction during TMS-EEG Co-Registration: A Comprehensive Review on Technologies and Procedures. *Sensors* 21: 637. [PubMed: 33477526]

44. Sekiguchi H, Takeuchi S, Kadota H, Kohno Y, Nakajima Y (2011): TMS-induced artifacts on EEG can be reduced by rearrangement of the electrode's lead wire before recording. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 122: 984–990.
45. Huber R, Esser SK, Ferrarelli F, Massimini M, Peterson MJ, Tononi G (2007): TMS-induced cortical potentiation during wakefulness locally increases slow wave activity during sleep. *PLoS One* 2: e276. [PubMed: 17342210]
46. Gosseries O, Sarasso S, Casarotto S, Boly M, Schnakers C, Napolitani M, et al. (2015): On the cerebral origin of EEG responses to TMS: insights from severe cortical lesions. *Brain Stimulat* 8: 142–149.
47. Mutanen T, Mäki H, Ilmoniemi RJ (2013): The effect of stimulus parameters on TMS-EEG muscle artifacts. *Brain Stimulat* 6: 371–376.
48. Ross JM, Sarkar M, Keller CJ (2022): Experimental suppression of transcranial magnetic stimulation-electroencephalography sensory potentials. *Hum Brain Mapp*. 10.1002/hbm.25990
49. Ozdemir RA, Boucher P, Fried PJ, Momi D, Jannati A, Pascual-Leone A, et al. (2021): Reproducibility of cortical response modulation induced by intermittent and continuous theta-burst stimulation of the human motor cortex. *Brain Stimulat* 14: 949–964.
50. Belardinelli P, Biabani M, Blumberger DM, Bortoletto M, Casarotto S, David O, et al. (2019): Reproducibility in TMS-EEG studies: A call for data sharing, standard procedures and effective experimental control. *Brain Stimulat* 12: 787–790.
51. Rocchi L, Di Santo A, Brown K, Ibáñez J, Casula E, Rawji V, et al. (2021): Disentangling EEG responses to TMS due to cortical and peripheral activations. *Brain Stimulat* 14: 4–18.
52. Biabani M, Fornito A, Mutanen TP, Morrow J, Rogasch NC (2019): Characterizing and minimizing the contribution of sensory inputs to TMS-evoked potentials. *Brain Stimulat* 12: 1537–1552.
53. Rogasch NC, Fitzgerald PB (2013): Assessing cortical network properties using TMS-EEG. *Hum Brain Mapp* 34: 1652–1669. [PubMed: 22378543]
54. Tremblay S, Rogasch NC, Premoli I, Blumberger DM, Casarotto S, Chen R, et al. (2019): Clinical utility and prospective of TMS-EEG. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 130: 802–844.
55. Paus T, Sipila PK, Strafella AP (2001): Synchronization of neuronal activity in the human primary motor cortex by transcranial magnetic stimulation: an EEG study. *J Neurophysiol* 86: 1983–1990. [PubMed: 11600655]
56. TMS-EEG co-registration: on TMS-induced artifact - PubMed (n.d.): Retrieved July 9, 2022, from <https://pubmed-ncbi-nlm-nih-gov.laneproxy.stanford.edu/19535291/>
57. Recovering TMS-evoked EEG responses masked by muscle artifacts - PubMed (n.d.): Retrieved July 9, 2022, from <https://pubmed-ncbi-nlm-nih-gov.laneproxy.stanford.edu/27291496/>
58. Automatic and robust noise suppression in EEG and MEG : The SOUND algorithm - PubMed (n.d.): Retrieved July 9, 2022, from <https://pubmed-ncbi-nlm-nih-gov.laneproxy.stanford.edu/29061529/>
59. Wu W, Keller CJ, Rogasch NC, Longwell P, Shpigel E, Rolle CE, Etkin A (2018): ARTIST: A fully automated artifact rejection algorithm for single-pulse TMS-EEG data. *Hum Brain Mapp* 39: 1607–1625. [PubMed: 29331054]
60. Rogasch NC, Sullivan C, Thomson RH, Rose NS, Bailey NW, Fitzgerald PB, et al. (2017): Analysing concurrent transcranial magnetic stimulation and electroencephalographic data: A review and introduction to the open-source TESA software. *NeuroImage* 147: 934–951. [PubMed: 27771347]
61. Russo S, Sarasso S, Puglisi GE, Dal Palù D, Pigorini A, Casarotto S, et al. (2022): TAAC - TMS Adaptable Auditory Control: A universal tool to mask TMS clicks. *J Neurosci Methods* 370: 109491. [PubMed: 35101524]
62. A structured ICA-based process for removing auditory evoked potentials - PubMed (n.d.): Retrieved July 9, 2022, from <https://pubmed-ncbi-nlm-nih-gov.laneproxy.stanford.edu/35082350/>
63. Novembre G, Pawar VM, Kilintari M, Bufacchi RJ, Guo Y, Rothwell JC, Iannetti GD (2019): The effect of salient stimuli on neural oscillations, isometric force, and their coupling. *NeuroImage* 198: 221–230. [PubMed: 31085301]

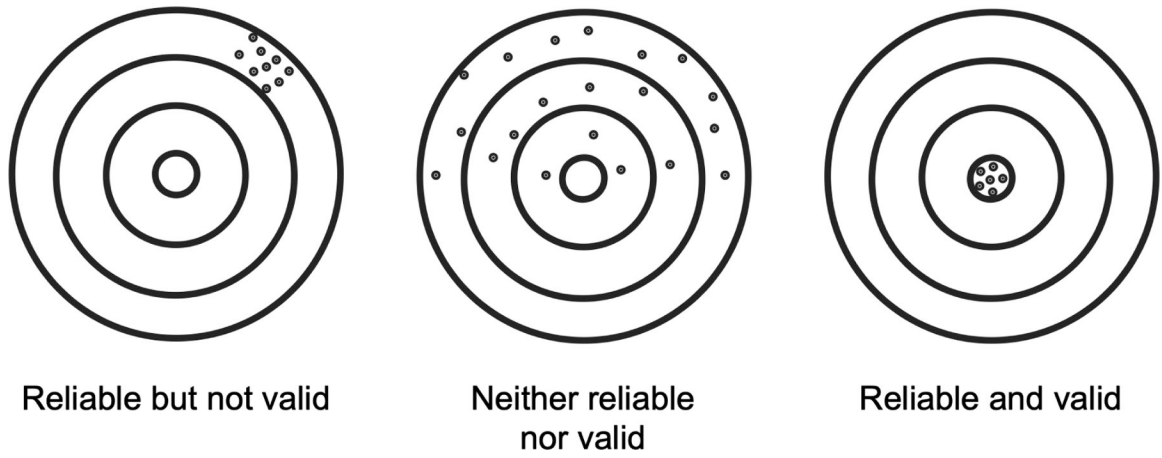
64. Disentangling EEG responses to TMS due to cortical and peripheral activations - PubMed (n.d.): Retrieved July 9, 2022, from <https://pubmed-ncbi-nlm-nih-gov.laneproxy.stanford.edu/33127580/>
65. Veniero D, Bortoletto M, Miniussi C (2009): TMS-EEG co-registration: on TMS-induced artifact. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 120: 1392–1399.
66. Mutanen TP, Metsomaa J, Liljander S, Ilmoniemi RJ (2018): Automatic and robust noise suppression in EEG and MEG: The SOUND algorithm. *NeuroImage* 166: 135–151. [PubMed: 29061529]
67. Wu W, Keller CJ, Rogasch NC, Longwell P, Shpigel E, Rolle CE, Etkin A (2018): ARTIST: A fully automated artifact rejection algorithm for single-pulse TMS-EEG data. *Hum Brain Mapp* 39: 1607–1625. [PubMed: 29331054]
68. Ross JM, Ozdemir RA, Lian SJ, Fried PJ, Schmitt EM, Inouye SK, et al. (2022): A structured ICA-based process for removing auditory evoked potentials. *Sci Rep* 12: 1391. [PubMed: 35082350]
69. Meyberg S, Werkle-Bergner M, Sommer W, Dimigen O (2015): Microsaccade-related brain potentials signal the focus of visuospatial attention. *NeuroImage* 104: 79–88. [PubMed: 25285375]
70. Yuval-Greenberg S, Tomer O, Keren AS, Nelken I, Deouell LY (2008): Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron* 58: 429–441. [PubMed: 18466752]
71. Herring JD, Thut G, Jensen O, Bergmann TO (2015): Attention Modulates TMS-Locked Alpha Oscillations in the Visual Cortex. *J Neurosci Off J Soc Neurosci* 35: 14435–14447.
72. Mouraux A, Diukova A, Lee MC, Wise RG, Iannetti GD (2011): A multisensory investigation of the functional significance of the “pain matrix.” *NeuroImage* 54: 2237–2249. [PubMed: 20932917]
73. Rogasch NC, Thomson RH, Farzan F, Fitzgibbon BM, Bailey NW, Hernandez-Pavon JC, et al. (2014): Removing artefacts from TMS-EEG recordings using independent component analysis: importance for assessing prefrontal and motor cortex network properties. *NeuroImage* 101: 425–439. [PubMed: 25067813]
74. Mutanen TP, Biabani M, Sarvas J, Ilmoniemi RJ, Rogasch NC (2020): Source-based artifact-rejection techniques available in TESA, an open-source TMS-EEG toolbox. *Brain Stimulat* 13: 1349–1351.
75. Atluri S, Frehlich M, Mei Y, Garcia Dominguez L, Rogasch NC, Wong W, et al. (2016): TMSEEG: A MATLAB-Based Graphical User Interface for Processing Electrophysiological Signals during Transcranial Magnetic Stimulation. *Front Neural Circuits* 10: 78. [PubMed: 27774054]
76. Wu W, Keller CJ, Rogasch NC, Longwell P, Shpigel E, Rolle CE, Etkin A (2018): ARTIST: A fully automated artifact rejection algorithm for single-pulse TMS-EEG data. *Hum Brain Mapp* 39: 1607–1625. [PubMed: 29331054]
77. Ilmoniemi RJ, Kici D (2010): Methodology for combined TMS and EEG. *Brain Topogr* 22: 233–248. [PubMed: 20012350]
78. Tchumatchenko T, Reichenbach T (2014): A cochlear-bone wave can yield a hearing sensation as well as otoacoustic emission. *Nat Commun* 5: 4160. [PubMed: 24954736]
79. ter Braack EM, de Vos CC, van Putten MJAM (2015): Masking the Auditory Evoked Potential in TMS-EEG: A Comparison of Various Methods. *Brain Topogr* 28: 520–528. [PubMed: 23996091]
80. Mutanen T, Nieminen JO, Ilmoniemi RJ (2013): TMS-evoked changes in brain-state dynamics quantified by using EEG data. *Front Hum Neurosci* 7: 155. [PubMed: 23630486]
81. Mäki H, Ilmoniemi RJ (2010): The relationship between peripheral and early cortical activation induced by transcranial magnetic stimulation. *Neurosci Lett* 478: 24–28. [PubMed: 20435086]
82. Fecchio M, Pigorini A, Comanducci A, Sarasso S, Casarotto S, Premoli I, et al. (2017): The spectral features of EEG responses to transcranial magnetic stimulation of the primary motor cortex depend on the amplitude of the motor evoked potentials. *PloS One* 12: e0184910. [PubMed: 28910407]
83. Biabani M, Fornito A, Coxon JP, Fulcher BD, Rogasch NC (2021): The correspondence between EMG and EEG measures of changes in cortical excitability following transcranial magnetic stimulation. *J Physiol* 599: 2907–2932. [PubMed: 33599980]

84. Kähkönen S, Komssi S, Wilenius J, Ilmoniemi RJ (2005): Prefrontal transcranial magnetic stimulation produces intensity-dependent EEG responses in humans. *NeuroImage* 24: 955–960. [PubMed: 15670672]
85. Kähkönen S, Wilenius J, Komssi S, Ilmoniemi RJ (2004): Distinct differences in cortical reactivity of motor and prefrontal cortices to magnetic stimulation. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 115: 583–588.
86. Lioumis P, Kici D, Savolainen P, Mäkelä JP, Kähkönen S (2009): Reproducibility of TMS-Evoked EEG responses. *Hum Brain Mapp* 30: 1387–1396. [PubMed: 18537115]
87. Rogasch NC, Zipser C, Darmani G, Mutanen TP, Biabani M, Zrenner C, et al. (2020): The effects of NMDA receptor blockade on TMS-evoked EEG potentials from prefrontal and parietal cortex. *Sci Rep* 10: 3168. [PubMed: 32081901]
88. Belardinelli P, König F, Liang C, Premoli I, Desideri D, Müller-Dahlhaus F, et al. (2021): TMS-EEG signatures of glutamatergic neurotransmission in human cortex. *Sci Rep* 11: 8159. [PubMed: 33854132]
89. Sun Y, Farzan F, Mulsant BH, Rajji TK, Fitzgerald PB, Barr MS, et al. (2016): Indicators for Remission of Suicidal Ideation Following Magnetic Seizure Therapy in Patients With Treatment-Resistant Depression. *JAMA Psychiatry* 73: 337–345. [PubMed: 26981889]
90. Hui J, Zomorodi R, Lioumis P, Salavati B, Rajji TK, Chen R, et al. (2020): Pharmacological mechanisms of interhemispheric signal propagation: a TMS-EEG study. *Neuropsychopharmacol Off Publ Am Coll Neuropsychopharmacol* 45: 932–939.
91. Casula EP, Maiella M, Pellicciari MC, Porrazzini F, D'Acunto A, Rocchi L, Koch G (2020): Novel TMS-EEG indexes to investigate interhemispheric dynamics in humans. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 131: 70–77.
92. Casali AG, Gosseries O, Rosanova M, Boly M, Sarasso S, Casali KR, et al. (2013): A theoretically based index of consciousness independent of sensory processing and behavior. *Sci Transl Med* 5: 198ra105.
93. Powden CJ, Hoch JM, Hoch MC (2015): Reliability and minimal detectable change of the weight-bearing lunge test: A systematic review. *Man Ther* 20: 524–532. [PubMed: 25704110]
94. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. (2010): The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil* 19: 539–549.
95. Ozdemir RA, Tadayon E, Boucher P, Momi D, Karakhanyan KA, Fox MD, et al. (2020): Individualized perturbation of the human connectome reveals reproducible biomarkers of network dynamics relevant to cognition. *Proc Natl Acad Sci U S A* 117: 8115–8125. [PubMed: 32193345]
96. Tomasevic L, Takemi M, Siebner HR (2017): Synchronizing the transcranial magnetic pulse with electroencephalographic recordings effectively reduces inter-trial variability of the pulse artefact. *PloS One* 12: e0185154. [PubMed: 28934300]
97. Thut G, Ives JR, Kampmann F, Pastor MA, Pascual-Leone A (2005): A new device and protocol for combining TMS and online recordings of EEG and evoked potentials. *J Neurosci Methods* 141: 207–217. [PubMed: 15661302]
98. White WL (2010): Why I hate the index finger. *Hand N Y N* 5: 461–465.
99. Rogasch NC, Thomson RH, Daskalakis ZJ, Fitzgerald PB (2013): Short-latency artifacts associated with concurrent TMS-EEG. *Brain Stimulat* 6: 868–876.
100. Julkunen P, Pääkkönen A, Hukkanen T, Könönen M, Tiihonen P, Vanhatalo S, Karhu J (2008): Efficient reduction of stimulus artefact in TMS-EEG by epithelial short-circuiting by mini-punctures. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 119: 475–481.
101. Mancuso M, Sveva V, Cruciani A, Brown K, Ibáñez J, Rawji V, et al. (2021): Transcranial Evoked Potentials Can Be Reliably Recorded with Active Electrodes. *Brain Sci* 11: 145. [PubMed: 33499330]
102. Casarotto S, Fecchio M, Rosanova M, Varone G, D'Ambrosio S, Sarasso S, et al. (2022): The rt-TEP tool: real-time visualization of TMS-Evoked Potentials to maximize cortical activation and minimize artifacts. *J Neurosci Methods* 370: 109486. [PubMed: 35074394]

103. Tervo AE, Nieminen JO, Lioumis P, Metsomaa J, Souza VH, Sinisalo H, et al. (2022): Closed-loop optimization of transcranial magnetic stimulation with electroencephalography feedback. *Brain Stimulat* 15: 523–531.
104. Gordon PC, Jovellar DB, Song Y, Zrenner C, Belardinelli P, Siebner HR, Ziemann U (2021): Recording brain responses to TMS of primary motor cortex by EEG - utility of an optimized sham procedure. *NeuroImage* 245: 118708. [PubMed: 34743050]
105. Biabani M, Fornito A, Mutanen TP, Morrow J, Rogasch NC (2019): Characterizing and minimizing the contribution of sensory inputs to TMS-evoked potentials. *Brain Stimulat* 12: 1537–1552.
106. Raffin E, Harquel S, Passera B, Chauvin A, Bougerol T, David O (2020): Probing regional cortical excitability via input-output properties using transcranial magnetic stimulation and electroencephalography coupling. *Hum Brain Mapp* 41: 2741–2761. [PubMed: 32379389]
107. Gordon PC, Desideri D, Belardinelli P, Zrenner C, Ziemann U (2018): Comparison of cortical EEG responses to realistic sham versus real TMS of human motor cortex. *Brain Stimulat* 11: 1322–1330.
108. Du X, Choa F-S, Summerfelt A, Rowland LM, Chiappelli J, Kochunov P, Hong LE (2017): N100 as a generic cortical electrophysiological marker based on decomposition of TMS-evoked potentials across five anatomic locations. *Exp Brain Res* 235: 69–81. [PubMed: 27628235]
109. Nikulin VV, Kici D, Kähkönen S, Ilmoniemi RJ (2003): Modulation of electroencephalographic responses to transcranial magnetic stimulation: evidence for changes in cortical excitability related to movement. *Eur J Neurosci* 18: 1206–1212. [PubMed: 12956719]
110. Ferrarelli F, Phillips ML (2021): Examining and Modulating Neural Circuits in Psychiatric Disorders With Transcranial Magnetic Stimulation and Electroencephalography: Present Practices and Future Developments. *Am J Psychiatry* 178: 400–413. [PubMed: 33653120]
111. Ferrarelli F, Massimini M, Peterson MJ, Riedner BA, Lazar M, Murphy MJ, et al. (2008): Reduced evoked gamma oscillations in the frontal cortex in schizophrenia patients: a TMS/EEG study. *Am J Psychiatry* 165: 996–1005. [PubMed: 18483133]
112. Massimini M, Ferrarelli F, Murphy M, Huber R, Riedner B, Casarotto S, Tononi G (2010): Cortical reactivity and effective connectivity during REM sleep in humans. *Cogn Neurosci* 1: 176–183. [PubMed: 20823938]
113. Ferrarelli F, Massimini M, Sarasso S, Casali A, Riedner BA, Angelini G, et al. (2010): Breakdown in cortical effective connectivity during midazolam-induced loss of consciousness. *Proc Natl Acad Sci U S A* 107: 2681–2686. [PubMed: 20133802]
114. Bodart O, Fecchio M, Massimini M, Wannez S, Virgillito A, Casarotto S, et al. (2018): Meditation-induced modulation of brain response to transcranial magnetic stimulation. *Brain Stimulat* 11: 1397–1400.
115. Local sleep-like cortical reactivity in the awake brain after focal injury - PubMed (n.d.): Retrieved July 12, 2022, from <https://pubmed.ncbi.nlm.nih.gov/33188680/>
116. Rosanova M, Gosseries O, Casarotto S, Boly M, Casali AG, Bruno M-A, et al. (2012): Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. *Brain J Neurol* 135: 1308–1320.
117. Wang JB, Bruss JE, Oya H, Uitermarkt BD, Trapp NT, Gander PE, et al. (2022, January 21): Effects of transcranial magnetic stimulation on the human brain recorded with intracranial electrocorticography: First-in-human study. *bioRxiv*, p 2022.01.18.476811.
118. Keller CJ, Honey CJ, Mégevand P, Entz L, Ulbert I, Mehta AD (2014): Mapping human brain networks with cortico-cortical evoked potentials. *Philos Trans R Soc Lond B Biol Sci* 369. 10.1098/rstb.2013.0528
119. Keller CJ, Bickel S, Entz L, Ulbert I, Milham MP, Kelly C, Mehta AD (2011): Intrinsic functional architecture predicts electrically evoked responses in the human brain. *Proc Natl Acad Sci U S A* 108: 10308–10313. [PubMed: 21636787]
120. Keller CJ, Honey CJ, Entz L, Bickel S, Groppe DM, Toth E, et al. (2014): Corticocortical evoked potentials reveal projectors and integrators in human brain networks. *J Neurosci Off J Soc Neurosci* 34: 9152–9163.

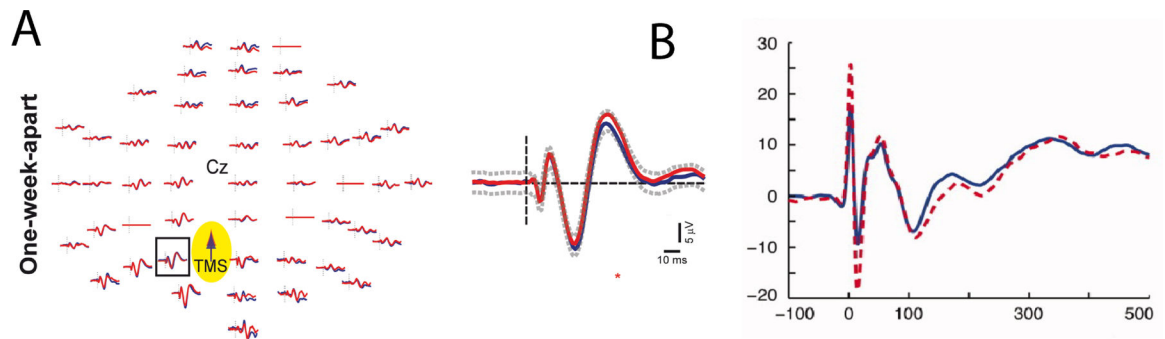


121. Huang Y, Hajnal B, Entz L, Fabó D, Herrero JL, Mehta AD, Keller CJ (2019): Intracortical Dynamics Underlying Repetitive Stimulation Predicts Changes in Network Connectivity. *J Neurosci Off J Soc Neurosci* 39: 6122–6135.
122. Keller CJ, Huang Y, Herrero JL, Fini ME, Du V, Lado FA, et al. (2018): Induction and Quantification of Excitability Changes in Human Cortical Networks. *J Neurosci Off J Soc Neurosci* 38: 5384–5398.
123. Cline CC, Lucas MV, Sun Y, Menezes M, Etkin A (2021): Advanced Artifact Removal for Automated TMS-EEG Data Processing. 2021 10th International IEEE/EMBS Conference on Neural Engineering (NER) 1039–1042.
124. Mutanen TP, Kukkonen M, Nieminen JO, Stenroos M, Sarvas J, Ilmoniemi RJ (2016): Recovering TMS-evoked EEG responses masked by muscle artifacts. *NeuroImage* 139: 157–166. [PubMed: 27291496]
125. Pion-Tonachini L, Kreutz-Delgado K, Makeig S (2019): ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage* 198: 181–197. [PubMed: 31103785]
126. Kallioniemi E, Daskalakis ZJ (2022): Identifying novel biomarkers with TMS-EEG - Methodological possibilities and challenges. *J Neurosci Methods* 377: 109631. [PubMed: 35623474]
127. Cao K-X, Ma M-L, Wang C-Z, Iqbal J, Si J-J, Xue Y-X, Yang J-L (2021): TMS-EEG: An emerging tool to study the neurophysiologic biomarkers of psychiatric disorders. *Neuropharmacology* 197: 108574. [PubMed: 33894219]



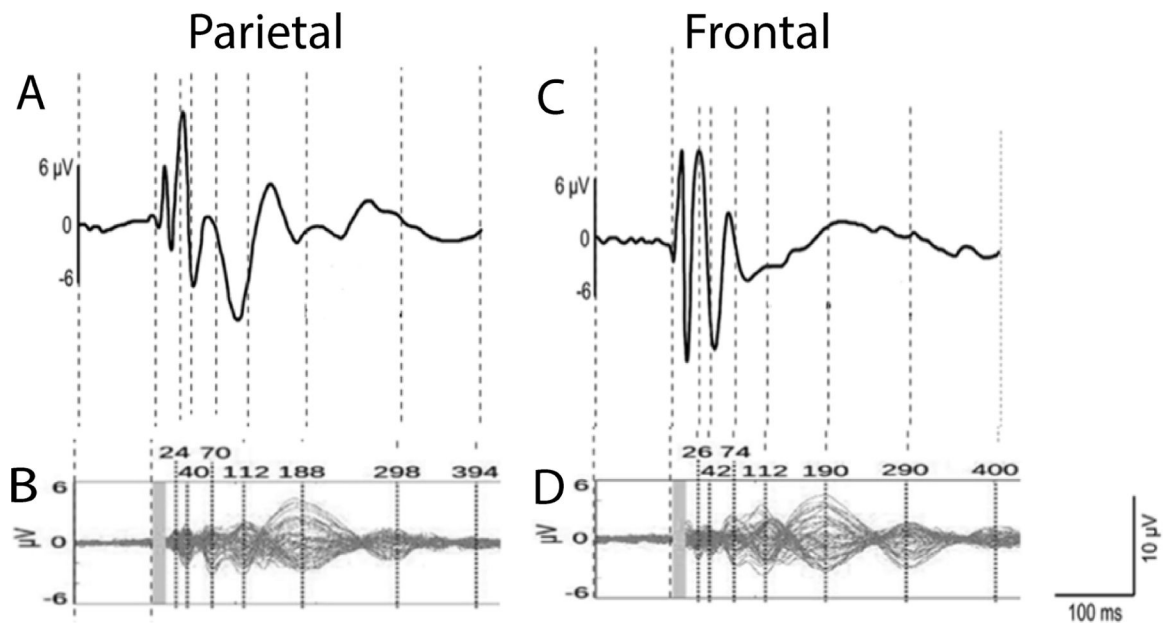
**Figure 1 - Reliability and validity of neuroimaging biomarkers.**

A biomarker should be scrutinized for both reliability and validity because it is possible to have high reliability but low validity or high validity but low reliability.



**Figure 2 - High internal reliability of TEPs.**

In each of these two studies (30,86), TEPs after single pulse TMS to M1 were *internally* consistent across time (one week between experiments). A adapted from (30), where blue line represents the first and red line the second recording one week later. B adapted from (86), where the dashed line represents the first and solid line the second recording one week later.



**Figure 3 - Low external reliability of the TEP.**

Comparison of TEPs from different laboratories. Local TEP responses after stimulation of A,B) parietal cortex and C,D) frontal cortex. A and C are from (42), and B and D are from (6). The figure is adapted from (50).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript