

Kognitive und somatische Komponenten der bewussten Fehlerwahrnehmung

Inauguraldissertation
zur
Erlangung des Doktorgrades
der Humanwissenschaftlichen Fakultät
der Universität zu Köln
nach der Promotionsordnung vom 10.05.2010
vorgelegt von

Jan Reinhold Wessel

aus
Heinsberg

Januar 2011

Diese Dissertation wurde von der
Humanwissenschaftlichen Fakultät der Universität zu Köln
im Mai 2011 angenommen.

Für

Ulrike Maria Wessel

* 06. November 1957

† 17. April 2006

Danksagung

Mein Dank gilt allen, die direkt oder indirekt an der Entstehung dieser Dissertation beteiligt waren.

Meiner Betreuerin Prof. Dr. Hilde Haider, für Rat und Tat in den letzten 6 Jahren.

Prof. Dr. Markus Ullsperger für die Anleitung in die Welt der Handlungsüberwachungsforschung.

Den Mitarbeiterinnen und Mitarbeitern der Arbeitsgruppe Kognitive Neurologie des Max-Planck-Instituts für Neurologische Forschung (S. Döring, T.O.J. Gründler, Dr. G. Jocham, Dr. T.A. Klein, Prof. Dr. J.B. Morton, allen studentischen Hilfskräften und im Besonderen Dr. Claudia Danielmeier), für hilfreiche Diskussionen und Anregungen.

Der AG psychologische Methodenlehre (in Persona Dr. A. Jain) und den Studenten des Department Psychologie der Universität zu Köln, für die Zusammenarbeit in der Lehre.

Den Mitarbeiterinnen und Mitarbeitern des Swartz Center for Computational Neuroscience in San Diego (in Persona Dr. S. Makeig), für die Gastfreundschaft während meines Forschungsaufenthaltes.

Allen hier unbenannten Kolleginnen und Kollegen, die auf Konferenzen, bei Vorträgen, im Gespräch oder als Gutachter oder Koautoren der Publikationen hilfreiche Anmerkungen zu den hier präsentierten Arbeiten beitrugen.

Der Gertrud Reemtsma Stiftung für Hirnforschung für die Finanzierung meiner Arbeit an dieser Dissertation. Der Gustav A. Lienert Stiftung für biopsychologische Methodik für die Finanzierung des Forschungsaufenthaltes in San Diego.

Meiner Familie, meinen Freunden, dafür, dass es sie gibt.

INHALTSVERZEICHNIS

1. Einleitung	2
1.1. Bewusste und unbewusste Fehlerregistrierung im Gehirn	3
1.2. Effekte fehlerhafter Handlungen auf das autonome Nervensystem .	4
2. Details der Publikationen	5
2.1. Publikation 1: Unterschiede zwischen bewusst wahrgenommenen und nicht bewusst wahrgenommenen Fehlern im zentralen und autonomen Nervensystem	5
2.2. Publikation 2: Datengetriebene Auswahl von unabhängigen Signalkomponenten des EEG aus einer ICA und die neuronale Basis von fehlerinduziertem kompensatorischen Verhalten.....	13
2.3. Publikation 3: Die Rolle konfligierender Antworttendenzen bei der Amplitudenmodulation der ERN	16
2.4. Publikation 4: Die unterschiedliche Hirnaktivität bei der bewussten und unbewussten Fehlerverarbeitung (Übersichtsartikel).....	19
3. Schlussfolgerung und Ausblick	21
4. Beiträge der Autoren zu den Publikationen	23
5. Literaturverzeichnis	25
6. Anhang	30

„ Deprehendas te oportet, antequam emendes. “
(*Bevor man sich verbessert, ist es nötig, sich zu ertappen*).

Lucius Annaeus Seneca, Epistulae morales ad lucilium, Brief XXVIII

Die Fähigkeit zur Überwachung des eigenen Verhaltens ist essentiell für zielgerichtetes und adaptives Handeln. Fehlerhafte Handlungen in alltäglichen Situationen können erhebliche unerwünschte Folgen haben – dies gilt im besonderen für solche Fehler, die von der handelnden Person nicht bemerkt werden.

Auch wenn sich im Verlaufe dieser Dissertation zeigen wird, dass Senecas exklusive Annahme, nur bewusst wahrgenommene Fehler können zu Korrektur- und Lernerfolg führen, nicht in allen Belangen aufrecht erhalten werden kann, so ist es aus der Alltagserfahrung intuitiv verständlich, dass die Natur der subjektiven Wahrnehmung (und das bewusste (Nicht-)Vorhandensein derselbigen) der Fehlerhaftigkeit einer Handlung einen erheblichen Einfluss auf das weitere Verhalten ausübt.

1. Einleitung

Die kognitive Neurowissenschaft befasst sich nunmehr seit zwei Jahrzehnten intensiv mit der Verarbeitung von Fehlern im menschlichen (und tierischen) Gehirn. Seit der Entdeckung eines ereigniskorrelierten Hirnpotentials (EKP) im menschlichen Elektroenzephalogramm (EEG), welches eine selektive Erhöhung im direkten Anschluss an Handlungsfehler zeigt (der s.g. Error-Related Negativity, Ne, bzw. ERN, Falkenstein, Hohnsbein, Hoormann, & Blanke, 1990; Gehring, Goss, Coles, Meyer, & Donchin, 1993), hat die Erforschung der Prozesse der Handlungsüberwachung mittels der Methoden der kognitiven Neurowissenschaft rapide zugenommen. So wurden zum Zeitpunkt dieser Dissertation mit allen wichtigen Methoden der kognitiven Neurowissenschaft Korrelate der Fehlerverarbeitung / Handlungsüberwachung im zentralen Nervensystem identifiziert – seien es durch funktionelle Magnetresonanztomographie (fMRT) gemessene Veränderungen in der metabolischen Aktivität präfrontaler Hirnrindenareale (Kerns, et al., 2004; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004), Veränderungen in der Feuerrate dopaminerger Neurone im Mittelhirn des Makaken (Schultz, Dayan, & Montague, 1997), oder Reaktionen des autonomen Nervensystems (ANS), quantifiziert durch die Herzrate (HR: Crone, et al., 2003; Danev & de Winter, 1971; Hajcak, McDonald, & Simons, 2003), den Pupillendurchmesser (PD:

Critchley, Tang, Glaser, Butterworth, & Dolan, 2005), oder die galvanische Hautleitfähigkeit (O'Connell, et al., 2007).

1.1. Bewusste und unbewusste Fehlerregistrierung im Gehirn

Ungleich weniger Forschungsarbeiten beschäftigten sich jedoch bisher damit, die Unterschiede in der Hirnaktivität zu charakterisieren, die sich in Abhängigkeit von, oder als Ursache der bewussten Wahrnehmung oder Nichtwahrnehmung fehlerhafter Handlungen messen lassen. Zudem bieten diese Arbeiten ein hochgradig inkohärentes Bild: Eine einflussreiche Publikation aus dem Jahre 2001 (Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001) zeigt eine Dissoziation zwischen den beiden bekannten fehlerbezogenen EKPs, der frühen ERN (die für gewöhnlich in den ersten 100 ms nach einer fehlerhaften Handlung an frontozentralen Elektroden ihre maximale Auslenkung zeigt) und der späteren Error Positivity (Pe, Falkenstein, Hoormann, Christ, & Hohnsbein, 2000; Overbeek, Nieuwenhuis, & Ridderinkhof, 2005). Während erstere keine Größenveränderung in Abhängigkeit von der Fähigkeit der Probanden, ihre Handlung per Tastendruck als fehlerhaft zu deklarieren, zeigte (diese simplifizierende operationale Definition des komplexeren Phänomens „Fähigkeit zum bewussten Abruf von Informationen“ ist die bisher beinahe exklusiv verwandte Definition in der neurowissenschaftlichen Subdisziplin der Kognitionsforschung), so war die Amplitude der P_e äußerst sensitiv für diese Variable. Dieser Befund wurde in der Folge häufig repliziert (Endrass, Reuter, & Kathmann, 2007; O'Connell, et al., 2007; Shalgi, Barkan, & Deouell, 2009) und zum Teil auch in der funktionellen Magnetresonanztomographie berichtet (Hester, Foxe, Molholm, Shpaner, & Garavan, 2005; Klein, et al., 2007, dort quantifiziert durch die hämodynamische Aktivität des posterioren medialen Frontalcortex, pmFC, dem neuronalen Generator der ERN, siehe Dehaene et al., 1994; Ullsperger & von Cramon, 2001; Debener et al., 2005). Jedoch gab es sowohl vor (Scheffers & Coles, 2000) als auch nach (Hester, et al., 2010; Maier, Steinhauser, & Hubner, 2008; Steinhauser & Yeung, 2010) der Publikation dieses Befundes Studien, die diesen Befund im Hinblick auf den Nulleffekt in der ERN-Amplitude nicht replizieren konnten, so dass nach wie vor Unklarheit darüber herrscht, wie die Rolle der ERN bei der Generierung von subjektivem Fehlerbewusstsein geartet ist.

Dies wird begleitet von der immer noch anhaltenden Unklarheit über den exakten Mechanismus, der der Genese der ERN zu Grunde liegt. In den hauptsächlich dominierenden Theorien postulieren die respektiven Autoren verschiedene Funktionen

der ERN, von der einfachen Detektion einer Diskrepanz zwischen ausgeführter und intendierter Reaktion (Mismatch Theory: Coles, Scheffers, & Holroyd, 2001), über die Detektion zweier konfligierender Antworttendenzen (Response-Conflict Theory: Botvinick, Braver, Barch, Carter, & Cohen, 2001; Yeung, Cohen, & Botvinick, 2004), bis hin zur Rolle der ERN als Auslösesignal in einem Verstärkungslernprozess (Reinforcement-Learning Theory: Holroyd & Coles, 2002).

Im Hinblick auf die dargestellte Thematik wird in der vorliegenden Dissertation einerseits versucht, die Rolle der ERN im Hinblick auf die subjektive Fehlerbewusstheit zu beleuchten (Publikationen 1,2 und 4), als auch, einen Beitrag zur Debatte über den Generator-Mechanismus der ERN selber zu leisten (Publikation 3).

1.2. Effekte fehlerhafter Handlungen auf das autonome Nervensystem

Auch außerhalb des Zentralnervensystems, im peripheren oder autonomen Nervensystem, lassen sich Effekte von fehlerhaften Handlungen beobachten (s.o.). Dabei kann es sich einerseits um Manifestationen derselben Prozesse handeln, die sich auch in den Indizes der zentralnervösen Aktivität widerspiegeln, oder aber um Prozesse, die auf einer anderen Verarbeitungsebene stattfinden. Da die verschiedenen Instanzen des autonomen Nervensystems durch teils sehr unterschiedliche kortikale wie subkortikale Schaltkreise verschaltet sind, ist eine genaue Untersuchung so vieler dieser Instanzen wie möglich im Hinblick auf ihre Rolle bei der (un-)bewussten Fehlerverarbeitung von Nöten, um überhaupt die Postulierung einer einheitlichen Theorie der zerebralen Mechanismen der subjektiven Fehlerbewusstheit zu ermöglichen. Bis zu dieser Dissertation gibt es jedoch nur eine einzige Untersuchung, die sich mit der Frage nach den unterschiedlichen Reaktionen des autonomen Nervensystems auf bewusst wahrgenommene und nicht bewusst wahrgenommene Fehlern beschäftigt (O'Connell et al., 2007). Dort wurde der galvanische Hautwiderstand in einer Fehlerbewusstheitsaufgabe (*Error awareness task*, EAT, Hester et al., 2005) gemessen; es zeigte sich eine signifikante relative Erhöhung der elektrodermalen Aktivität (EDA) bei bewusst wahrgenommenen Fehlern im Vergleich zu korrekten Eingaben, bei gleichzeitiger relativer Verringerung der EDA (bis hin zur völligen Absenz selbiger) bei nicht bewusst wahrgenommenen Fehlern.

In der vorliegenden Dissertation soll nun, zusätzlich zu den bisher benannten Zielen, versucht werden, die Rolle des autonomen Nervensystems bei der bewussten bzw. unbewussten Fehlerwahrnehmung genauer zu untersuchen, indem mit einem

multimodalen Ansatz zwei stark unterschiedlich innervierte Instanzen peripherer Aktivität im Hinblick auf ihre Abhängigkeit vom subjektiven Bewusstseitsgrad der Fehlerverarbeitung gemessen werden (Publikation 1). Zudem wird eine Integration der Ergebnisse in eine einheitliche Theorie der (un-)bewussten Fehlerwahrnehmung auf Basis dieser und anderer Ergebnisse vorgenommen (Publikation 4).

2. Details der Publikationen

Das „Herz“ dieser Dissertation bilden 2 Publikationen (Wessel, Danielmeier, & Ullsperger, in press; Wessel & Ullsperger, 2011), die der Autor der Dissertation als Erstautor angefertigt hat. Weiterhin werden zwei Publikationen angeführt (Danielmeier, Wessel, Steinhauser, & Ullsperger, 2009; Ullsperger, Harsay, Wessel, & Ridderinkhof, 2010) an denen der Autor der Dissertation als Koautor mitgewirkt hat.

2.1. Publikation 1: Unterschiede zwischen bewusst wahrgenommenen und nicht bewusst wahrgenommenen Fehlern im zentralen und autonomen Nervensystem

In der ersten der Publikationen (Wessel, Danielmeier & Ullsperger, in press) wurde die Rolle sowohl des autonomen Nervensystems, als auch des zentralen Nervensystems bei der bewussten und unbewussten Fehlerwahrnehmung untersucht. Es werden zwei Experimente präsentiert, bei denen multiple Messinstrumente (EEG, Elektrokardiographie (EKG), Pupillometrie) benutzt wurden, um die Aktivität eines Großteils des Gesamtnervensystems bei der Bearbeitung einer Anti-Sakkaden Aufgabe aufzuzeichnen. Bei der Anti-Sakkaden Aufgabe erscheint ein imperativer Stimulus in einer von zwei Boxen am linken bzw. rechten Rand des Bildschirms; die Probanden sollen ihren Blick bei Erscheinen dieses Stimulus zur anderen Box hinwenden. Fehler in der Aufgabe sind definiert als irrtümliche Sakkaden in Richtung des Stimulus. Die Anti-Sakkaden Aufgabe eignet sich besonders zur Untersuchung der Unterschiede zwischen bewusst wahrgenommenen und nicht bewusst wahrgenommenen Fehlern, da sie, im Gegensatz zu sonst in der Handlungsüberwachungsforschung eingesetzten Aufgaben einen hinreichenden Anteil von nicht bewusst wahrgenommenen Fehlern induziert (siehe Nieuwenhuis et al., 2001). Mehrere neuartige und fortschrittliche Analyseansätze sowohl der Verhaltensdaten (Sakkaden wie Fehlersignalisierung), als auch der physiologischen Daten, unter anderem mittels Independent Component Analysis (ICA, Jutten & Herault, 1991, s.u.), kamen in dieser Studie zum Einsatz und konnten dazu

beitragen, die bisher wenig kommensurable (ERN) oder nur spärlich vorhandene (ANS) Literatur zu bereichern.

In Experiment 1 konnte zunächst gezeigt werden, dass sowohl Herzrate, als auch Pupillendurchmesser hypothesenkonforme Unterschiede zwischen bewusst wahrgenommenen und nicht bewusste wahrgenommenen Fehlern (ab hier: bwF und nbwF) zeigen, i.S. einer stärkeren Aktivität nach bwF im Vergleich zu nbwF: Die nach Fehlern auftretende relative Herzratenverringering (Danev & de Winter 1971, Crone et al., 2003, Hajcak et al., 2003) zeigte sich hypothesenkonform exklusiv für bwF (siehe Abbildung 1).

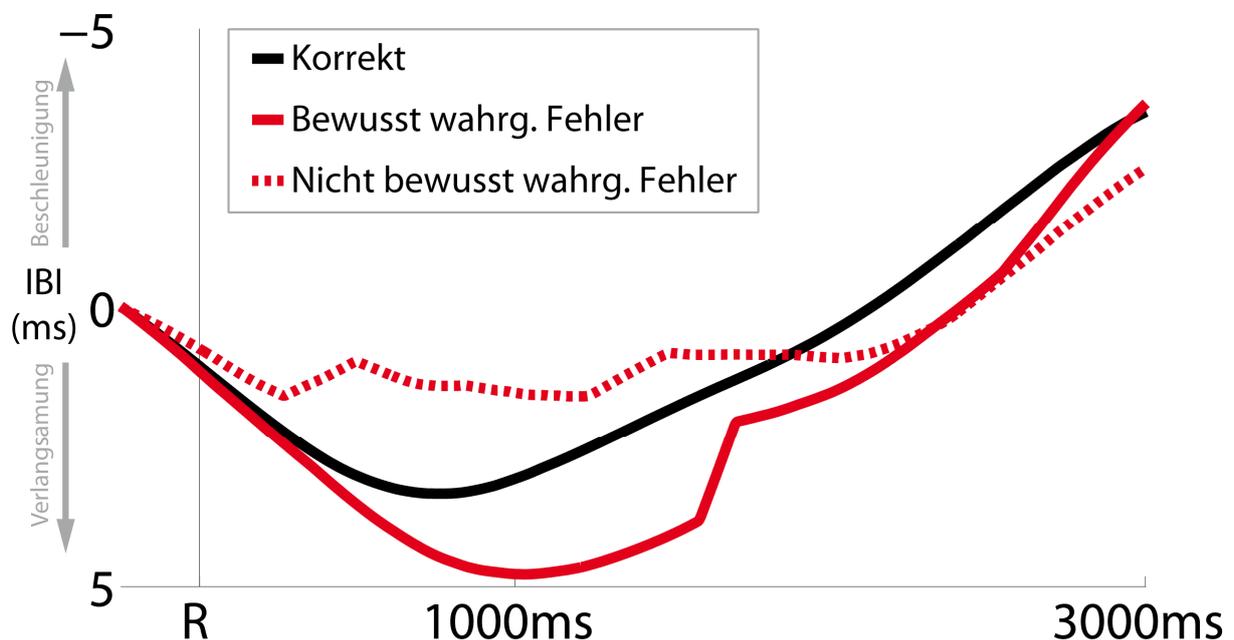


Abbildung 1: Herzraten-Resultate aus Wessel, Danielmeier & Ullsperger (im Druck), Experiment 1. Es zeigt sich ein signifikanter Unterschied zwischen bewusst wahrgenommenen Fehlern und nicht bewusst wahrgenommenen Fehlern, insofern als dass nur erstere die bekannte fehlerinduzierte Herzratenverlangsamung zeigen.

Zudem konnte ebenfalls gezeigt werden, dass die nach Fehlern auftretende relative Pupillendilatation (Critchley et al., 2005) bei bwF im Vergleich zu nbwF signifikant stärker ausgeprägt ist (siehe Abbildung 2). Diese Resultate lassen, zusammen mit den EDA Befunden von O'Connell et al. (2007), auf eine global stärkere Reaktion des autonomen Nervensystems bei bewusst wahrgenommenen Fehlern schließen und ermöglichen eine genauere Theorie der Rolle des ANS bei der Fehlerverarbeitung (s.u.).

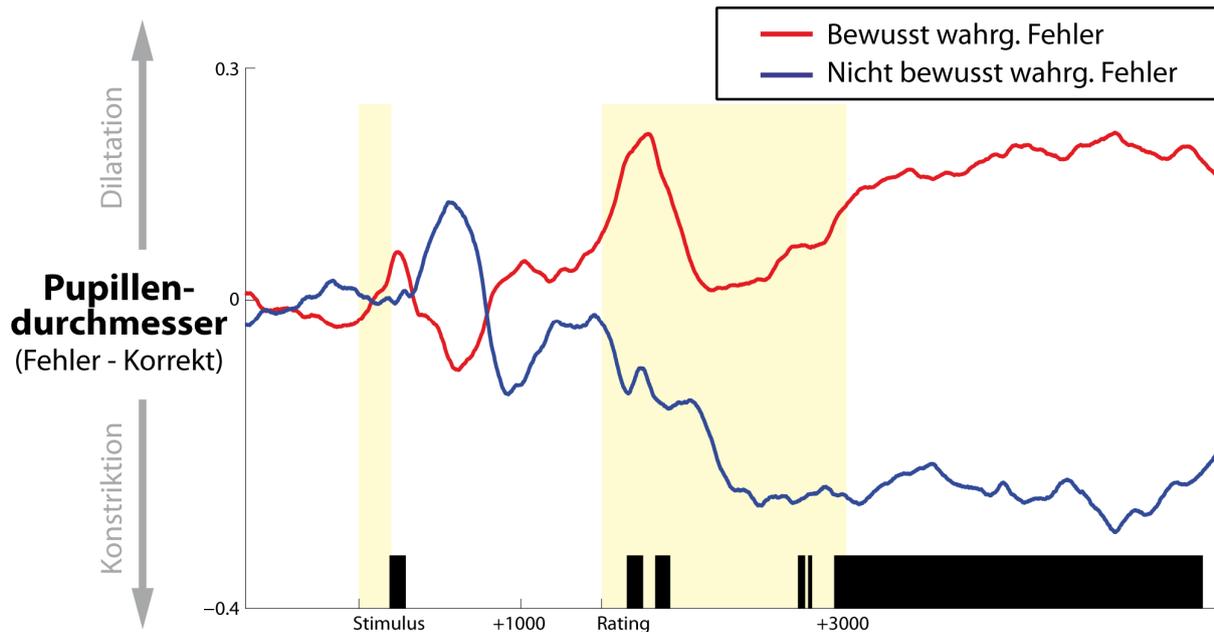


Abbildung 2: Pupillometrie-Resultate aus Wessel, Danielmeier & Ullsperger (im Druck), Experiment 1. Gezeigt ist der (um die Lichtreaktion bereinigte) Pupillendurchmesser, getrennt für beide Fehlertypen. Es handelt sich jeweils um die Differenzwellen zwischen Fehlern und allen korrekten Eingaben. Werte über 0 bedeuten demnach eine relative Pupillendilatation im Vergleich zu den korrekten Eingaben, Werte unter 0 entsprechen einer Pupillenkonstriktion. In den gelb unterlegten Zeiträumen treten jeweils neue Lichtreize auf. Die schwarzen Balken zeigen signifikante Unterschiede zwischen den beiden Fehlertypen an (Monte-Carlo Simulation, 10.000 Iterationen, bei FDR-Korrektur für multiple Vergleiche auf $\alpha = 0,05$). Es zeigt sich vor allem in späteren Latenzbereichen signifikant stärkere Dilatation bei den bewusst wahrgenommenen Fehlern.

Zusätzlich zeigten sich im Zentralnervensystem in Abhängigkeit von der subjektiven Fehlerbewusstheit Unterschiede nicht nur in der Amplitude der P_e , sondern auch in der Amplitude der ERN, die, auch wenn kommensurabel mit einigen vorherigen Veröffentlichungen, dem größeren Teil der bisherigen Literatur (Nieuwenhuis et al., 2001, Endrass et al., 2007, O'Connell et al., 2007, Shalgi et al., 2009) widersprechen: Während zwar einige Befunde (Scheffers & Coles, 2000, Maier et al., 2008, Steinhauser & Yeung, 2010) auf eine größere ERN bei bewusst wahrgenommenen Fehlern hindeuten, so ist die bisher vorherrschende Meinung vor allem zurückzuführen auf Nullbefunde aus der ersten Veröffentlichung (Nieuwenhuis et al., 2001), die sich explizit mit der Frage des Einflusses der subjektiven Fehlerbewusstheit auf die fehlerbezogenen EKPs beschäftigt hat. Dort wurde gezeigt, dass die subjektive Fehlerbewusstheit scheinbar

keinen signifikanten Effekt auf die Größe der ERN zeigt. Diesen Nulleffekt konnten wir in Experiment 1 nicht replizieren: Die ERN war nach bewusst wahrgenommenen Fehlern im Vergleich zu nbwF signifikant erhöht (siehe Abbildung 3).

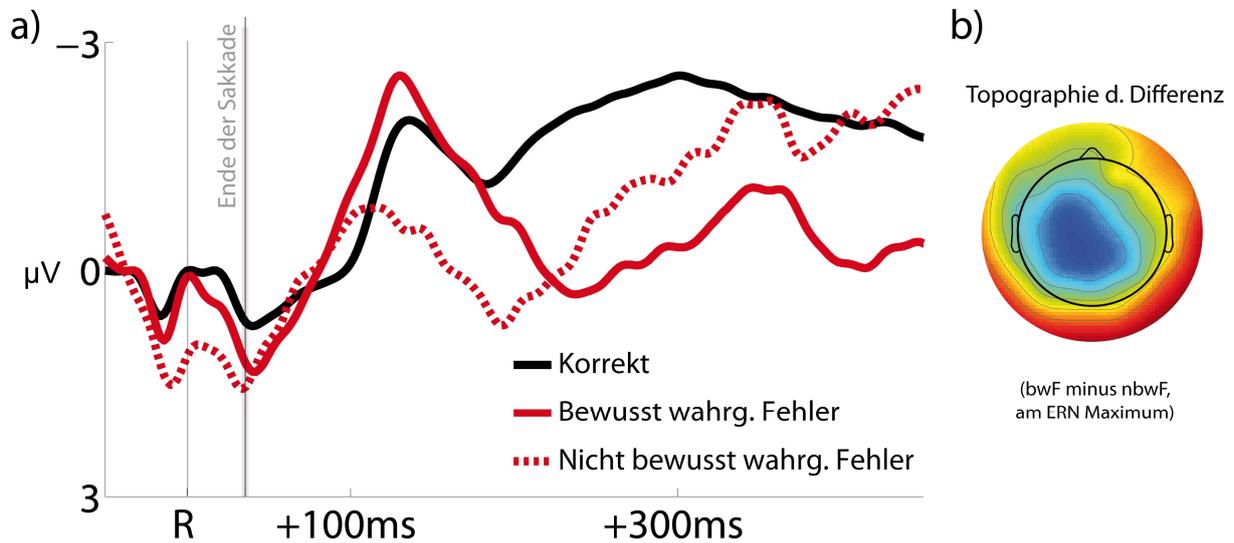


Abbildung 3: ERN-Resultate aus Wessel, Danielmeier & Ullsperger (im Druck), Experiment 1. Es zeigt sich ein signifikanter Effekt der subjektiven Fehlerbewusstheit auf die Amplitude der ERN in a). In b) ist die topographische Verteilung dieser Differenz über den Skalp zu sehen. R = Response (man beachte, dass bei der Anti-Sakkaden Aufgabe „Anfang“ und „Ende“ einer Antwort bestimmbar sind, da es sich nicht um Tastendrucke, sondern Sakkaden handelt. R bezeichnet den Beginn der Sakkade, die Hilfslinie bei 39 ms gibt das durchschnittliche Ende der Sakkaden an, der graue Balken darum das 95% Konfidenzintervall).

Experiment 2 wurde in der Folge explizit dazu entwickelt, um die Gründe für diese Dissonanz zwischen den aktuellen Befunden von Experiment 1 und der bisherigen Literatur genauer zu untersuchen. Das experimentelle Paradigma wurde (wie in Experiment 1) so gestaltet, dass sich eine maximale Kompatibilität (im Hinblick auf die Darbietung und die Art des Stimulusmaterials) mit dieser ersten Publikation ergibt, die den Nulleffekt der Fehlerbewusstheit auf die ERN-Amplitude zunächst demonstriert hat (Nieuwenhuis et al., 2001). Es wurde in Experiment 2 nicht nur darauf geachtet, das gleiche Paradigma zu verwenden wie in Nieuwenhuis et al. (2001), sondern das identische Stimuluslayout sowie Trialtiming aus dieser Arbeit. Um die Dissonanz zwischen den Befunden aus unserem Experiment 1 und der Referenzstudie optimal erklären zu können, wurde als zusätzliche Auswertetechnik eine Operationalisierung

einer präziseren Abstufung des Grades der subjektiven Fehlerbewusstheit vorgenommen, ohne jedoch das Paradigma (Anti-Sakkaden Aufgabe) an sich zu verändern, um einen direkten Vergleich zu ermöglichen: Dies geschah durch einen Mediansplit der Versuchsdurchgänge im Hinblick auf die Zeit, die die Probanden benötigten, um die Entscheidung bezüglich der Akkuratheit ihrer Eingaben einzuschätzen. Die Probanden wurden nach wie vor nach jedem Versuchsdurchgang befragt, ob sie meinen, einen Fehler begangen zu haben, oder nicht. Die Zeit, die die Probanden zur Entscheidung für oder gegen einen Fehler benötigten, wurde als Indikator für den Grad der subjektiv empfundenen Sicherheit der eigenen Einschätzung interpretiert. Durch diese Technik war es möglich, ohne Veränderung der „Bewusstseinsabfrage“ an sich einen zusätzlichen Richtwert für den Grad der Ausprägung der subjektiven Fehlerbewusstheit zu schaffen: Entschieden die Probanden sich sehr schnell, dass sie richtig oder falsch lagen, so sollte diese Entscheidung mit höherer Sicherheit behaftet sein als wenn sie bei der Beantwortung der Frage lange zögerten.

Zunächst konnte auch in diesem Experiment, übereinstimmend mit Experiment 1 dieser Publikation und im Gegensatz zur Referenzpublikation (Nieuwenhuis et al., 2001) gezeigt werden, dass die ERN-Amplitude bei bwF im Vergleich zu nbwF signifikant erhöht ist. Zusätzlich konnte dieser Unterschied fast exklusiv auf das Subset von Versuchsdurchgängen zurückgeführt werden, bei denen die Probanden ihr Verhalten mit hoher Sicherheit (reflektiert in niedriger Entscheidungszeit) als fehlerhaft einschätzten, was die differentielle Rolle der ERN bei der Bewusstwerdung von Verhaltensfehlern unterstreicht: Eine hohe ERN-Amplitude geht einer schnellen Entscheidung für die Fehlerhaftigkeit der zurückliegenden Handlung voraus (siehe Abbildung 4).

Ferner werden in der Diskussion Gründe aufgezeigt, die möglicherweise dazu geführt haben, dass dieser Unterschied in der ERN-Amplitude bisher nicht immer (bzw. nicht immer mit statistischer Signifikanz) demonstriert werden konnte. Dabei sind vor allem zwei (miteinander verbundene) Aspekte von Bedeutung: a) ein möglicher Antwortbias durch das experimentelle Design der ersten Publikation zum Thema (Nieuwenhuis et al., 2001), welcher sich ebenfalls im Großteil der späteren Replikationen mit Nullbefunden findet (O'Connell et al., 2007, Shalgi et al., 2009) und b) der mangelnden statistischen Teststärke zur Absicherung einer Nullhypothese in allen Arbeiten.

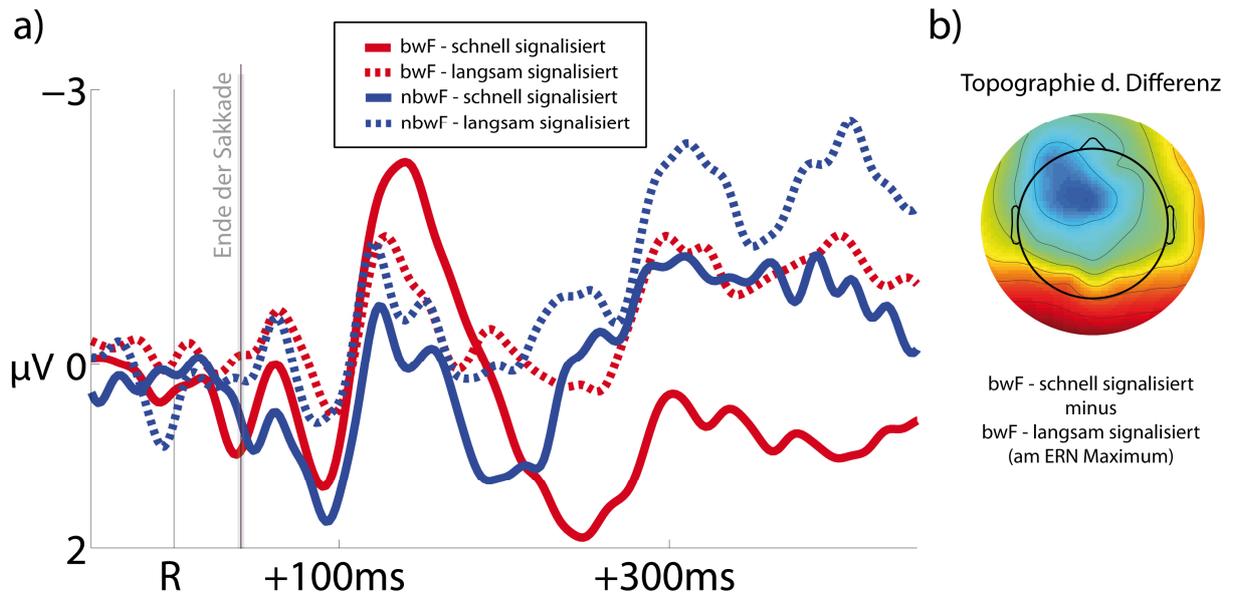


Abbildung 4: ERN-Resultate aus Wessel, Danielmeier & Ullsperger (im Druck), Experiment 2. Gezeigt sind die Ergebnisse der Mediansplit-Analyse der Entscheidungszeiten für die Klassifikation der eigenen Handlung als fehlerhaft oder korrekt. Es zeigt sich, dass der (in Experiment 2 wieder signifikante) Unterschied zwischen bewusst wahrgenommenen und nicht bewusst wahrgenommenen Fehlern in der ERN-Amplitude vor allem auf die große ERN-Amplitude bei schnell signalisierten, bewusst wahrgenommenen Fehlern zurückzuführen ist.

Bezüglich des möglichen Antwortbias ist die Instruktion in der Studie von Nieuwenhuis et al. (2001) zu beachten: Die Probanden wurden instruiert, bei der Abfrage zur Fehlerbewusstheit eine Taste zu drücken, wenn sie der Meinung waren, einen Fehler gemacht zu haben, und im Gegenzug *keine* Taste zu drücken, wenn sie der Meinung waren, keinen Fehler gemacht zu haben. Dies kann nun zwei Effekte haben, die die Befundlage potentiell zu Gunsten eines Nullbefundes im Hinblick auf die ERN beeinflusst haben können: Zunächst hatten die Probanden in den benannten Studien nach Beginn eines jeden Versuchsdurchgangs nur 1400 bis 2250 ms Zeit, sowohl die durch die Aufgabe geforderte Sakkade auszuführen, als auch bei Bedarf die anschließende Taste für die Fehlersignalisierung zu betätigen. Sollten die Probanden für beides zusammen länger als diese Zeit benötigt haben, wurde der entsprechende Versuchsdurchgang klassifiziert als „Proband bemerkte keinen Fehler“. Dies führt potentiell dazu, dass Versuchsdurchgänge, bei denen Probanden einen Fehler vermuten, jedoch (zu) lange für ihre bewusste Entscheidung brauchen, fälschlicherweise als „nicht bewusste Fehler“ klassifiziert werden. Da nach unseren Hypothesen und Befunden eine größere ERN das

bewusste Bemerkung eines Fehlers begünstigt, hätten solche Durchgänge eine verhältnismäßig große ERN und würden das mittlere ERN-EKP bei den „nicht bewusst wahrgenommenen Fehlern“ laut dieser Klassifikation vergrößern. Der Nulleffekt der subjektiven Fehlerbewusstheit würde damit über Gebühr begünstigt. Der zweite Effekt, der mit dieser Instruktion einhergeht, ist der Tatsache geschuldet, dass das Drücken einer Taste mit mehr mentalem und physischem Aufwand verbunden ist als das „Nichtstun“ bei wahrgenommenen Nicht-Fehlern, was einen möglichen Antwortbias zu Gunsten der Nichtsignalisierung von Fehlern zur Folge haben könnte. Eine (im Artikel nicht berichtete) Analyse der Verhaltensdaten liefert Evidenz für diese Argumentation: Die „False Alarm Rate“ (FAR, Green & Swets, 1966), ein Parameter aus der Signalentdeckungstheorie, der im Kontext dieser Studien angibt, wie häufig die Probanden fälschlicherweise einen Fehler signalisieren unterscheidet sich nennenswert zwischen den Studien: Während in der vorliegenden Studie die FAR bei 9,7% liegt, lag sie bei Nieuwenhuis et al. bei 1,5%, was auf einen bemerkenswerten (aber mangels weiterer statistischer Werte aus Nieuwenhuis et al. Studie nicht inferenzstatistisch testbaren) Bias zur Nichtsignalisierung eines Fehlers hindeuten kann. In der Konsequenz heißt dies, dass die Probanden bei hoher Unsicherheit über die Akkuratheit ihrer Eingabe (die unserer Hypothese nach durch die größere Amplitude der ERN verursacht wird) eher geneigt sind, ihren Fehler dennoch nicht zu signalisieren, was wiederum den Nullbefund im ERN-EKP begünstigt. Weitere Studien, die diesen Nulleffekt berichteten, benutzten dieselbe Instruktion (O’Connell et al., 2007, Shalgi et al., 2009). In unseren Studien benutzten wir eine Instruktion, bei denen die Probanden in jedem Fall eine Taste drücken sollten, sprich: Taste A für einen wahrgenommenen Fehler, Taste B für eine subjektiv als korrekt empfundene Eingabe. Zudem hatten die Probanden mehr Zeit für ihre Entscheidung (der nächste Durchgang begann erst bei Druck einer der beiden Tasten). Dadurch wird der Antwortbias effektiv eliminiert und eine fragwürdige Einordnung einzelner Durchgänge vermieden, womit auch die Nullhypothese nicht über Gebühr begünstigt wird.

Der zweite Hauptaspekt, der zur Dissonanz zwischen den Befunden zwischen den Studien führen kann, liegt in der statistischen Teststärke. Mit Ausnahme der Studie von Nieuwenhuis et al., (2001) finden alle Studien, trotz teilweise potentiell vorhandenem Antwortbias eine numerisch (aber nicht signifikant) vergrößerte ERN nach bwF (Endrass et al., 2007, Shalgi et al., 2009, O’Connell et al., 2007), oder stärkere Aktivität des Generators der ERN im pmFC nach bwF im Vergleich zu nbwF (Klein et al., 2007,

Hester et al., 2005). Auf Grund der Komplexität (und des damit einhergehenden zeitlichen und monetären Aufwands) der Datenerhebung in den kognitiven Neurowissenschaften sind die Stichprobenumfänge solcher Studien häufig beschränkt und bieten damit nicht genug statistische Teststärke, um kleinere, aber vorhandene Effekte nachweisen zu können. Keine der hier behandelten Studien (inklusive der beiden Experimente aus der vorliegenden Publikation) hat Stichproben mit mehr als 20 Probanden, was dazu führt, dass Nullhypothesen trotz imminent hoher Beta-Fehler Wahrscheinlichkeiten akzeptiert werden: Als Beispiel berichten Shalgi et al. (2009) einen statistisch nicht signifikanten Unterschied zwischen den ERN-Amplituden nach bwF und nbwF zu Gunsten einer größeren ERN nach bwF mit $p = 0,19$, bei einem Stichprobenumfang von 16 Probanden. Ähnliches findet sich in der fMRT-Literatur, wo bei Klein et al. (2007) numerische Unterschiede in der Amplitude der hämodynamischen Antwort im pMFC zu Gunsten stärkerer Aktivität nach bwF mit $p = 0,21$ bei 13 Probanden als nicht signifikanter Effekt verworfen werden müssen. Es ist also durchaus möglich, dass tatsächlich vorhandene Unterschiede der Aktivität des der ERN zu Grunde liegenden Hirnareals auf Grund mangelnder Teststärke nicht gefunden werden. Ein starkes Indiz dafür ist Tatsache, dass neuerliche Untersuchungen mit größeren Stichproben (>20) im fMRT tatsächlich (auch statistisch signifikante) Unterschiede in der Aktivität des pMFC zwischen bwF und nbwF im Einklang mit unseren Befunden finden (Harsay, et al., submitted; Hester, et al., 2010).

Zusätzlich zu diesen, im Artikel genannten Gründen, ist vor allem auch anzumerken, dass der Einsatz der Independent Component Analysis (einer erst seit wenigen Jahren für das EEG etablierten Methode, Delorme & Makeig, 2004; Makeig, Jung, Bell, Ghahremani, & Sejnowski, 1997) zur Korrektur der, auf Grund der Antwortmodalität im Anti-Sakkaden Task überdurchschnittlich stark ausgeprägten, Augenbewegungsartefakte im EEG einen relativen Zugewinn an Validität gebracht haben kann. Die vorliegende Publikation ist die erste dieser Art, die diese Technik zur Artefaktbereinigung benutzt. Diese Methode war zum Zeitpunkt der ersten Veröffentlichung (Nieuwenhuis et al., 2001) noch nicht flächendeckend verfügbar, bringt bei der Korrektur von Augenbewegungen (Hoffmann & Falkenstein, 2008; Mennes, Wouters, Vanrumste, Lagae, & Stiers, 2010) erhebliche Vorteile (wie auch bei anderen typischerweise im EEG gefundenen Artefakten: Delorme, Sejnowski, & Makeig, 2007).

2.2. Publikation 2: Datengetriebene Auswahl von unabhängigen Signalkomponenten des EEG aus einer ICA und die neuronale Basis von fehlerinduziertem kompensatorischen Verhalten

Die zweite Publikation (Wessel & Ullsperger, 2011) thematisiert in der Hauptsache eine neuartige Methode zur objektiveren Anwendung der ICA Methode auf ereigniskorrelierte EEG Daten. Die Independent Component Analysis (Comon, 1994; Jutten & Herault, 1991) ist eine Sammelbezeichnung für mathematische Blindquellentrennungsverfahren, bei denen ein Mischsignal (z.B. eine Audio-Aufnahme, oder eben ein EEG), also eine lineare Summe der Signale mehrerer unabhängiger Generatoren (im Falle des EEGs lokaler Dipolfelder), welches über mehrere Sensoren (im Falle des EEGs Elektroden) aufgezeichnet wurde, auf Basis diverser Merkmale von tatsächlichen Quellensignalen (z.B. Nicht-Gaussianität, stochastische Unabhängigkeit, niedrige Komplexität, siehe Hyvärinen, Karhunen, & Oja, 2001) in eben jene Quellensignale zurücküberführt wird. Durch algebraische Operationen können in der Folge Zeitreihen extrahiert werden, die im Idealfall ausschließlich die Aktivität eines der Generatoren des linear-additiven Mischsignals (im Falle des EEGs also eines lokalen Dipolfeldes, sprich: eines funktionell distinkten Hirnareals) widerspiegeln. ICA-basierte Verfahren wurden in den 1990er Jahren erstmals sowohl auf EEG (Makeig et al., 1997), als auch auf fMRT Daten angewandt (McKeown, et al., 1998). Einer der vielen Vorteile der Anwendung dieser Methode ist die Tatsache, dass die Identifizierung einer mittels ICA ermittelten Signalkomponente bei korrekter Zuordnung zu einem konkreten Hirnprozess von allen anderen Signalquellen isoliert und in der Folge extrahiert werden kann. Damit steigt das Signal-Rausch-Verhältnis (SRV) rapide an, was Auswertungen der Hirnsignale auf Einzeldurchgangsebene erlaubt, die ohne diesen Verarbeitungsschritt nur schwer möglich wären (siehe z.B. Debener, et al., 2005; Eichele, Juvodden, Ullsperger, & Eichele, 2010; Gentsch, Ullsperger, & Ullsperger, 2009; Roger, Benar, Vidal, Hasbroucq, & Burle, 2010). Problematisch ist dabei jedoch genau dieser letzte Schritt der an die ICA anschließenden Inferenz: Welcher Hirnprozess (z.B. ein EKP im EEG) wird in welcher Signalkomponente abgebildet? Bisher wurde diese Inferenz nach Augenscheinvalidität vom Wissenschaftler selber vorgenommen. In dieser Publikation wird nun ein Algorithmus vorgestellt, der vom Autor dieser Dissertation dazu entwickelt wurde, genau diesen Selektionsprozess zu standardisieren und automatisieren. Auf Basis von klaren inferenzstatistischen und stochastischen Kriterien kann dieser

Algorithmus für jeden Zeitraum in einem EKP die Quellsignale aus der ICA ermitteln, welche signifikant mit dieser Aktivität assoziiert sind (siehe Abbildung 5b).

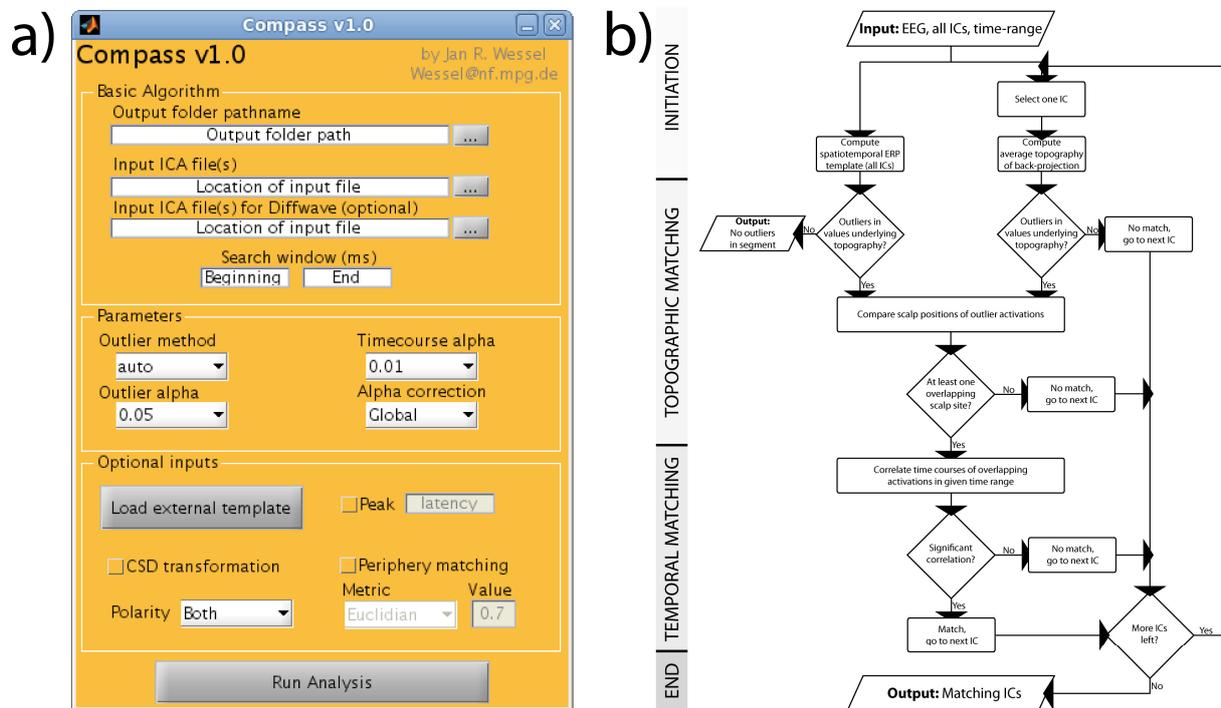


Abbildung 5: a) Grafische Benutzeroberfläche für die MATLAB-basierte Implementation des Algorithmus aus Wessel & Ullsperger (im Druck). b) Programmablaufplan für den Algorithmus (Vergrößerung von 5b siehe Publikation)

Nach einer ausführlichen Schilderung des entwickelten Algorithmus und einer Prüfung seiner Validität durch eine umfangreiche Simulation artifizierlicher Daten, wird jedoch zudem versucht, mittels Anwendung des selbigen Algorithmus einen Befund zu generieren, der für die Forschung zur (un-)bewussten Fehlerwahrnehmung von essentieller Bedeutung sein kann: Im Anschluss an fehlerhafte Handlungen kommt es zu adaptiven, kompensatorischen Verhaltensweisen, die sich unter anderem an der Reaktionszeit in den Versuchsdurchgängen nach fehlerhaften Eingaben zeigen lassen. Eine dieser kompensatorischen Reaktionen, dass s.g. Post-Error Slowing (PES, Rabbitt, 1967), eine relative Reaktionszeitverkürzung nach Fehlern im Vergleich zu korrekten Eingaben, wird nach einhelliger Befundlage ausschließlich nach bewusst wahrgenommenen Fehlern beobachtet (u.a. Nieuwenhuis et al., 2001, Endrass et al., 2007, Wessel, Danielmeier & Ullsperger, in press). Der Nachweis eines direkten Zusammenhangs der ERN-Amplitude mit dem Ausmaß des PES wäre daher ein weiteres Indiz für die direkte Rolle der ERN in der bewussten Fehlerwahrnehmung.

Es ist jedoch nicht eindeutig belegt, ob die Amplitude der ERN tatsächlich, wie häufig auf Grund ihrer neuronalen Generatoren hypothetisiert wird, das Ausmaß des Post-Error Slowing bedingt. Neuroimaging Studien zeigen, dass die metabolische Aktivität in den Hirnarealen (pmFC), die der ERN zu Grunde liegen, das Ausmaß des PES beeinflussen (Garavan, Ross, Murphy, Roche, & Stein, 2002; Kerns et al., 2004; Klein et al., 2007). Auch Tierstudien deuten auf einen solchen Zusammenhang hin (Narayanan & Laubach, 2006). Eine aktuelle Publikation (Debener et al., 2005) konnte diesen Zusammenhang im Menschen erstmalig auf Einzeldurchgangsebene demonstrieren, indem gezeigt wurde, dass die Amplitude der ERN bei Fehlern das Ausmaß des PES im nächsten Trial vorhersagen kann. Um, zur Generierung sinnvoller Rückschlüsse auf Einzeldurchgangsebene, entgegen der Mittelungslogik der EKPs (siehe z.B. Handy, 2005), ein hinreichendes Signal-Rausch-Verhältnis gewährleisten zu können, wurde von Debener et al. ein ICA-basierter Algorithmus zur Blindquellentrennung benutzt. Durch exklusive Selektion derjenigen Signalkomponente des EEGs, die das ERN-Signal repräsentiert (s.o.), generierten die Autoren ein hinreichendes SRV. Diese Technik hat vor allem im Bereich der Handlungsüberwachungsforschung in den letzten Jahren zu enormen Erkenntnisgewinn geführt (u.a. Gentsch et al., 2009, Roger et al., 2010, Eichele et al., 2010). Problematisch war dabei bisher jedoch die hochgradig subjektive Natur der Komponentenauswahl (s.o.): Dadurch, dass die Auswahl der Signalkomponente aus der ICA, die die ERN repräsentiert, zumeist per Augenscheinvalidität visuell vom Wissenschaftler selber getätigt wurde (auf Basis diverser a priori definierter Auswahlkriterien), gibt es, bedingt durch die hohe Zahl der Freiheitsgrade, bei einer solchen Auswahl potentielle Bedenken bezüglich der Objektivität und damit der Reliabilität solcher Untersuchungen.

In der vorliegenden Publikation wird nicht nur eine hochgradig datengetriebene Methode zur Auswahl solcher Signalkomponenten demonstriert und an Hand einer Simulation validiert, sondern es wird auch gezeigt, dass sich mit dieser verbesserten Technik die prädiktive Validität der ERN-Amplitude auf das Ausmaß des folgenden PES bestätigt: Eine größere ERN im Anschluss an einen Handlungsfehler geht mit einem größeren PES im nächsten Versuchsdurchgang einher. Dies konnte in einer Re-Analyse von Daten aus einer Flankierreizaufgabe (siehe Publikation 3, Danielmeier, Wessel, Steinhauser & Ullsperger, 2009) in dieser Publikation gezeigt werden. Im Zusammenspiel mit den Befunden u.a. aus Publikation 1 bezüglich der Exklusivität des PES für bewusst wahrgenommene Fehler ist dies als weitere, indirekte Evidenz für die

entscheidende Rolle der ERN bei der Bewusstwerdung von Handlungsfehlern zu interpretieren.

Über diese Befunde hinaus erwies sich die präsentierte Methode in der Simulation als ebenso valide für die Identifikation der unabhängigen Signalkomponenten von anderen ereigniskorrelierten Potentialen als der ERN. Die Methode und die dazu entwickelte grafische Benutzeroberfläche für MATLAB (siehe Abbildung 5a) ist damit für alle thematischen Felder der kognitiven Neurowissenschaft, weit über die Forschung zur Handlungsüberwachung hinaus, von enormem Potential.

2.3. Publikation 3: Die Rolle konfligierender Antworttendenzen bei der Amplitudenmodulation der ERN

Diese Publikation, an der der Autor der Dissertation als Ko-Autor beteiligt war, beschäftigt sich mit dem genaueren Mechanismus, der der Genese der ERN zu Grunde liegt.

Aus den hier präsentierten Arbeiten (im Besonderen Publikation 1) wird in der Folge ein einheitliches Modell zur Entstehung der bewussten Fehlerwahrnehmung im Gehirn abgeleitet (siehe Publikation 4), in dem den beiden fehlerbezogenen EKPs (ERN und P_e) unterschiedliche Funktionen bei der Fehlerregistrierung zukommen. In diesem „Evidenzakkumulationsmodell der bewussten Fehlerwahrnehmung“ werden multiple Instanzen potentieller Evidenz angenommen, die dem kognitiven System zur Detektion der Fehlerhaftigkeit einer Handlung zur Verfügung stehen. Dazu gehören neben physiologischen Signalen, wie der motorischen Efferenzkopie und der Proprio- und Interozeption, auch „objektive“ externale und internale Merkmale für die Fehlerhaftigkeit einer Handlung, wie der sensorische Input der Handlungseffekte (external, dieser konnte in den Verhaltensdaten aus Publikation 1 eindeutig als unterschiedlich zwischen den beiden Fehlertypen nachgewiesen werden) und der so genannte *Response-Conflict* (internal, s.o.). Nach dem hier präsentierten Evidenzakkumulationsmodell der bewussten Fehlerwahrnehmung spiegeln sich diese „objektiven“ Merkmale in der Amplitude der ERN wider (siehe Publikation 1), wohingegen das Zustandekommen der späteren P_e , im Einklang mit anderen Publikationen (z.B. Nieuwenhuis et al., 2001, Steinhauser & Yeung, 2010), als Ergebnis von Evaluations- oder Entscheidungsprozessen erklärt wird.

In der vorliegenden Publikation wurde nun versucht zu zeigen, dass die ERN-Amplitude tatsächlich direkt von der Größe des Konfliktes zweier gleichzeitig aktiver

Antworttendenzen (Botvinick et al., 2001) abhängt. Um zwischen dieser und anderen Theorien für die Genese der ERN (s.o.) differenzieren zu können, wurde für diese Studie eine besonders kontraintuitive Prädiktion aus dieser so genannten „*Response-Conflict*“ (RC) Theorie abgeleitet: Prä- und post-*Response-Conflict* sollten durch die Manipulation des Abstandes der Distraktorstimuli in einer Flankierreizaufgabe voneinander dissoziiert werden können. Laut der RC-Theorie ist die Größe der ERN direkt bedingt durch den *Response-Conflict*, der sich aus dem Produkt zweier gleichzeitig aktiver Antworttendenzen ergibt (meist einer richtigen und einer falschen). Die Flankierreizaufgabe (Eriksen & Eriksen, 1974) ist ein klassisches Paradigma in der Handlungsüberwachungsforschung, die sich auch für die Erforschung des *Response-Conflicts* bewährt hat. In der hier verwandten Version werden die Probanden instruiert, einem zentral dargebotenen Zielreiz (Pfeil nach links oder Pfeil nach rechts) folgend die entsprechende Taste auf einer Antwortbox zu betätigen. Kurz vor der Darbietung des Zielreizes werden ober- und unterhalb Flankierreize dargeboten: Pfeile, die entweder in dieselbe Richtung wie der Zielreiz weisen (kongruenter Stimulus) oder in die entgegengesetzte Richtung (inkongruenter Stimulus). Durch das Vorhandensein solcher inkongruenter Stimuli und die schnelle Taktung der Aufgabe, die die Probanden unter Zeitdruck absolvieren, machen die Probanden bei der Bearbeitung der Aufgabe im Normalfall eine hinreichende Menge von Fehlern, um eine statistische Testung fehlerbezogener Potentiale zu ermöglichen. Die Aufgabenstellung ist entsprechend gut für die Handlungsüberwachungsforschung geeignet. Besonders die inkompatiblen Versuchsdurchgänge eignen sich, auf Grund des Vorhandenseins von Pfeilen in beide Richtungen (von denen nur eine zu beachten ist) zur Erforschung des RC. Die zentrale Manipulation in der hier verwandten Variante der Flankierreizaufgabe war die Distanz der Distraktoren vom zentral dargebotenen Zielreiz.

Die *Response-Conflict* Theorie sagt für diese Manipulation kontraintuitiver Weise einen geringeren post-RC für inkompatible Distraktorstimuli voraus, die sich näher am imperativen Stimulus befinden, als für solche Versuchsdurchgänge, in denen die inkompatiblen Reize räumlich weiter entfernt dargeboten werden. Für den prä-RC ist die RC Hypothese genau umgekehrt: Distraktoren, die sich näher am Zielreiz befinden, erzeugen eine stärkere Antworttendenz zur falschen Taste, eine Vorhersage, die insofern intuitiver erscheint, als dass näher am Zielreiz dargebotene Reize visuell zentrierter sind und damit mehr im Fokus der Aufmerksamkeit auf den Zielreiz stehen. Laut der RC-Theorie gilt dies jedoch nur für den prä-RC. Beim post-RC führt die

Tatsache, dass weiter entfernte Flankierreize weniger attentionales Gewicht erhalten, zu einer, im Verhältnis zu Stimuli mit nahen Distraktoren, stärkeren Aktivierung der richtigen Antworttendenz. Da der *Response-Conflict* sich laut der RC-Theorie aus dem Produkt der beiden gleichzeitig aktiven Antworttendenzen ergibt, sagt die RC-Theorie eine größere ERN nach Fehlern mit weiter entfernten Distraktoren voraus, wohingegen für das N2 EKP (ein stimulusbezogenen EKP, welches den prä-RC abbildet) die entgegengesetzte Vorhersage gilt.

Diese Vorhersagen für die Größe der ERN und des N2 EKPs (Yeung et al., 2004), wurden in der aktuellen Publikation mit Hilfe einer komputationalen Simulation des RC Modells an Hand von Verhaltensdaten aus einem Vorexperiment (Experiment 1 in dieser Publikation) generiert und dann mittels EEG (Experiment 2) getestet. Der motorische *Response-Conflict* konnte dabei zusätzlich durch das Lateralisierte Bereitschaftspotential (Coles, 1989; Vaughan, Costa, & Ritter, 1968) im Sinne eines Manipulation Checks physiologisch quantifiziert werden.

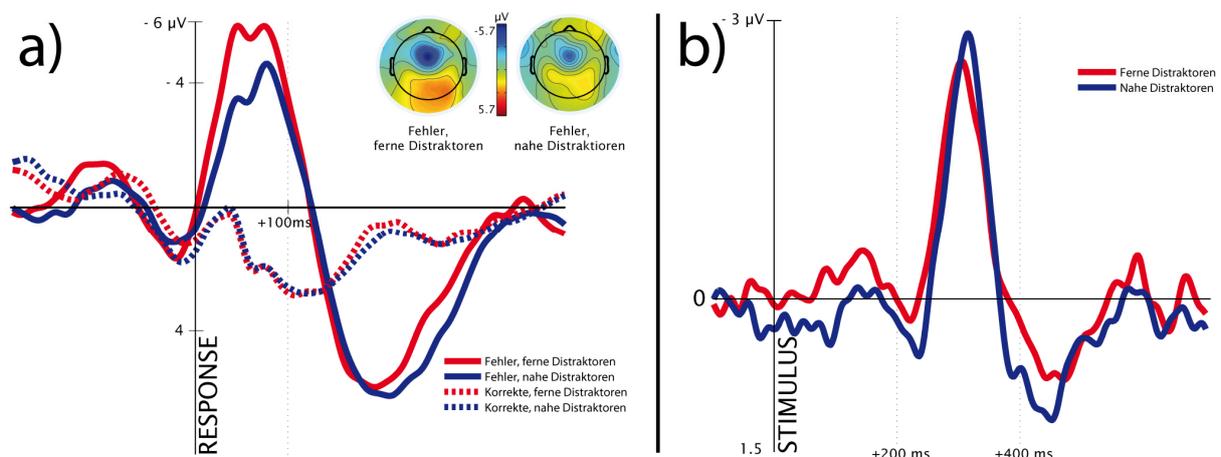


Abbildung 6: ERN-Resultate aus Danielmeier, Wessel, Steinhauser & Ullsperger (2009), Experiment 2. Es zeigt sich im Einklang mit den Hypothesen a) eine signifikant größere ERN-Amplitude nach Fehler bei großen Distanzen zwischen Flankierreizen und Zielreiz und b) eine geringe N2-Amplitude bei großen Distanzen (die N2-Grafik zeigt Differenzpotentiale zwischen inkompatiblen und kompatiblen Versuchsdurchgängen des jeweiligen Stimulustyps, um den RC-Effekt zu betonen)

Es zeigte sich, dass alle physiologischen Daten mit den Prädiktionen des *Response-Conflict* Modells übereinstimmen (siehe Abbildung 6), wohingegen konkurrierende Modelle der ERN-Genese keine (oder falsche) Prädiktionen hinsichtlich der

Manipulation machen, oder die Erklärung der Befunde wenig sparsam (i.S.v. Ockham's Rasiermesser) erscheint. Es wird daraus gefolgert, dass das Ausmaß des *Response-Conflicts* tatsächlich einen direkten Einfluss auf die Größe der ERN, und damit (siehe Publikation 1 und 2) sowohl auf die Entstehung der subjektiven Fehlerbewusstheit, als auch auf fehlerbezogene adaptive Verhaltensänderungen hat.

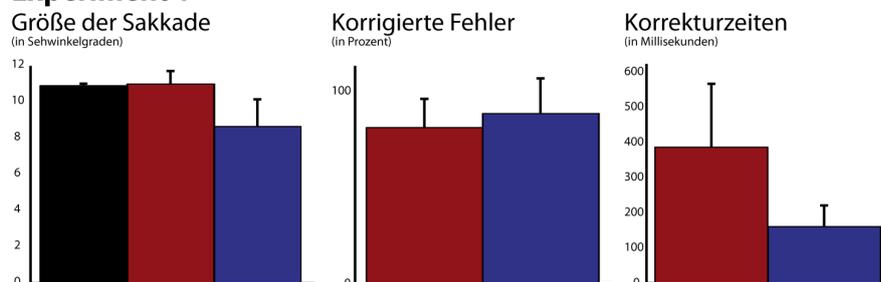
2.4. Publikation 4: Die unterschiedliche Hirnaktivität bei der bewussten und unbewussten Fehlerverarbeitung (Übersichtsartikel)

Bei dieser Publikation handelt es sich um einen Übersichtsartikel über die neurowissenschaftliche Literatur zu den Unterschieden zwischen bewusst und nicht bewusst wahrgenommenen Fehlern, mit einem besonderen Fokus auf den anterioren insulären Kortex. Gesteigerte Aktivität in dieser Hirnregion wird regelmäßig in Experimenten zur Handlungsüberwachung berichtet (für eine Metaanalyse siehe Klein et al., 2007, eine weitere Metaanalyse findet sich in dieser Publikation), jedoch gibt es bisher nur zwei Neuroimaging-Untersuchungen, die differentielle Eigenschaften dieser Hirnregion und des pMFC bei der Fehlerverarbeitung berichten (Magno, Foxe, Molholm, Robertson, & Garavan, 2006; Klein et al., 2007). In letzterer Veröffentlichung wurde gezeigt, dass der anteriore Teil des linken insulären Cortex bei bwF im Vergleich zu nbwF signifikant mehr Aktivität (im Sinne der mittels fMRT gemessenen hämodynamischen Antwort) aufweist. Vor diesem Hintergrund beleuchtet der Artikel zunächst die allgemeinen physiologischen Korrelate der Handlungsüberwachung, um dann die bisherigen Ansätze zur Untersuchung der Fehlerbewusstheit im neurowissenschaftlichen Bereich im Hinblick auf ihre Methodik und Ergebnisse zu beschreiben. In der Folge wird ebenfalls das bereits genannte (in diesem Artikel erstmalig formulierte) Evidenzakkumulationsmodell der bewussten Fehlerwahrnehmung beschrieben (siehe Abbildung 3 im o.g. Artikel), wobei ein zusätzliches Augenmerk auf eine potentielle Rolle der unspezifischen attentionalen Orientierungsreaktion (OR: Sokolov, 1960) bei der Evidenzakkumulation nach Fehlern gelegt wird. Hier werden vor allem die ANS Befunde von Publikation 1 zur Argumentation herangezogen, insofern als dass die differentiellen Effekte der subjektiven Fehlerbewusstheit, die dort im Hinblick auf das ANS gefunden wurden, im phänomenologischen Einklang mit der OR stehen: Die stärkere Aktivität des ANS nach bewusst wahrgenommenen Fehlern entspricht phänomenologisch einer stärker ausgeprägten Orientierungsreaktion. Ob diese OR nun Ursache oder Folge der

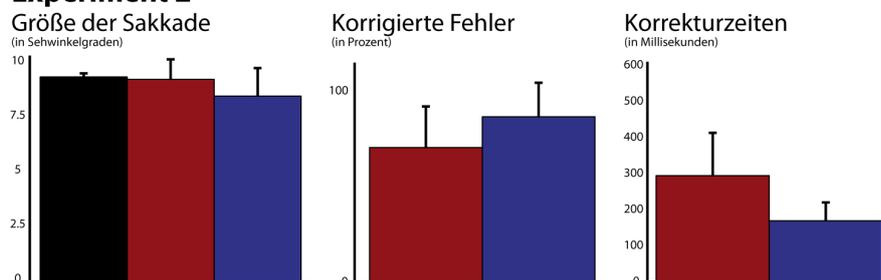
Fehlerbewusstheit ist, lässt sich ohne weitergehende Untersuchungen mit explizit darauf abgestimmten Manipulationen (Läsionsstudien, Transkranielle Magnetstimulation, Medikamentenstudien) nicht beantworten. Dieser Forschungsansatz wird, als direkte Folge der hier präsentierten Arbeiten, in laufenden Untersuchungen bereits verfolgt (siehe Schlussfolgerung und Ausblick).

Auch die behaviouralen Resultate von Wessel, Danielmeier & Ullsperger (in press) sind als direkte Evidenz für die Evidenzakkumulationstheorie zu interpretieren: Einer der großen Vorteile der Anti-Sakkaden Aufgabe ist die Möglichkeit, bei Anwendung geeigneter Algorithmen zur Bemessung von Sakkaden, die objektive „Größe“ von Fehlern zu bemessen, in dem man die Amplitude der Sakkade in die falsche Richtung bestimmt. Diese in Wessel, Danielmeier & Ullsperger (in press) erstmalig bei einer Handlungsüberwachungsstudie mittels Anti-Sakkaden Aufgabe vorgenommene Analyse zeigt deutlich, dass sich, wie die Evidenzakkumulationstheorie vorhersagt, nicht nur Merkmale wie die ERN oder die Aktivität des ANS zwischen bewusst wahrgenommenen und nicht bewusst wahrgenommenen Fehlern unterscheidet.

Experiment 1



Experiment 2



Korrekte
 Bewusst wahrg. Fehler
 Nicht bewusst wahrg. Fehler

Abbildung 7: Verhaltensdaten aus Wessel, Danielmeier & Ullsperger (im Druck). Die Fehlerbalken zeigen eine Standardabweichung.

Tatsächlich unterscheiden sich auch die objektiven, visuellen und propriozeptiven Informationen, die dem kognitiven System bei der Klassifizierung einer Handlung als fehlerhaft oder korrekt zur Verfügung stehen: Während die bewusst wahrgenommenen Fehler der Probanden in beiden Telexperimenten der Studie die identische Amplitude aufwiesen, wie die korrekten Eingaben, so war die Amplitude der nicht bewusst wahrgenommenen Fehler signifikant kleiner (siehe Abbildung 7).

Zudem gibt es signifikante Unterschiede in der spontanen Fehlerkorrektur zwischen beiden Fehlertypen: Nicht nur wurden signifikant mehr der nicht bewusst wahrgenommenen Fehler spontan korrigiert¹, sondern vor allem waren die Korrekturzeiten (i.e. der zeitliche Abstand zwischen dem Ende der fehlerhaften Sakkade und dem Beginn der folgenden Korrektursakkade) bei den nicht bewusst wahrgenommenen Fehlern im Schnitt weniger als halb so lang (siehe Abbildung 7). Dies bedeutet im Gesamten, dass bei bewusst wahrgenommenen Fehlern die Probanden größere Sakkaden in die falsche Richtung machen, diese häufiger nicht spontan korrigieren oder (bei späterer Korrektur) signifikant länger in die falsche Richtung schauen. All diese sensorische Evidenz steht dem kognitiven System bei der Klassifikation einer Handlung als fehlerhaft zur Verfügung. Das Evidenzakkumulationsmodell berücksichtigt sowohl diese Evidenz, als auch die differentiellen Effekte im Nervensystem (ERN, Error Positivity, ANS, Efferenzkopie, etc.). Eine weitere Herausforderung für nachfolgende Studien wird es sein, genau zu quantifizieren, welche Hirnstrukturen bei der Kodierung dieser objektiven Effekte eine Rolle spielen. Die Interaktion zwischen den primären und sekundären sensorischen Kortizes und dem pMFC ist dabei von besonderem Interesse.

3. Schlussfolgerung und Ausblick

Die hier vorliegenden Arbeiten erfüllen viererlei Zwecke:

- 1) Sie lösen den Widerspruch im Bezug auf die Rolle der ERN bei der bewussten Gewährleistung von Handlungsfehlern auf, indem **a)** gezeigt werden konnte, dass Unterschiede in der ERN-Amplitude zwischen bewusst wahrgenommenen und nicht bewusst wahrgenommenen fehlerhaften Handlungen, im Gegensatz zu vorherigen Befunden, auch in

¹ Ein Effekt, der auch in Endrass et al. (2007) beobachtet wurde, jedoch nicht in Nieuwenhuis et al. (2001), da dort auch Pro-Sakkaden zu einem nach dem Versuchsdurchgang erscheinenden „Zielkreuz“ auf der richtigen Seite als „spontane Korrekturen“ klassifiziert wurden,

Anti-Sakkaden Aufgaben zu finden sind (Publikation 1) und **b)** gezeigt werden konnte, dass exklusiv nach bewusst wahrgenommenen Fehlern auftretende adaptive Prozesse wie das Post-Error Slowing direkt durch die Amplitude der ERN vorhergesagt werden können (Publikation 2).

- 2) Sie ordnen diese und andere Befunde in eine umfassende Theorie der Evidenzakkumulation ein, die einen Großteil der zum Zeitpunkt der Dissertation erschienenen Literatur integriert (Publikation 4). Dazu gehört im Besonderen auch die Befunde über differentielle Effekte im ANS bei der Bewusstwerdung von Handlungsfehlern (Publikationen 1 und 4),
- 3) Sie motivieren neue Hypothesen über die Rolle der Orientierungsreaktion bei der Bewusstwerdung von Fehlern (Publikationen 1 und 4). Über die Rolle der Orientierungsreaktion bei fehlerinduzierten adaptiven Prozessen wie PES wurde kürzlich ebenfalls in Verhaltensstudien berichtet (Noetebart et al., 2009), ein Ansatz, der mit den hier präsentierten Resultaten hochkompatibel ist.
- 4) Sie etablieren ein zum Zwecke der Bearbeitung der Fragestellungen dieser Dissertation neu entwickeltes datengetriebenes Verfahren zur Identifikation von unabhängigen Signalkomponenten des EEGs bei Benutzung der Independent Component Analysis, welches weit über die hier präsentierten Anwendungen und Fragestellungen hinaus für die neurowissenschaftliche Forschung von Nutzen sein kann (Publikation 2).

Die hier gezeigten Befunde motivieren ein großes Spektrum weiterer Forschungsfragen, sowohl im Bereich der Handlungsüberwachung, als auch darüber hinaus. Hypothesen, die sich direkt aus dieser Arbeit ableiten lassen, betreffen beispielsweise die Natur etwaiger Attributionsprozesse (Wilbert, 2006), die im Zusammenhang mit der Evidenzakkumulationstheorie bei der Emergenz von Fehlerbewusstheit eine Rolle spielen können, die Natur des Zusammenhangs zwischen der Orientierungsreaktion, der Gewährwerdung von Handlungsfehlern und den Signalen des peripheren und zentralen Nervensystems, und methodische Ansätze, die, wie hier angewandt, zu enormen Erkenntniszuwachs im Bereich der kognitiven Neurowissenschaft führen können.

4. Beiträge der Autoren zu den Publikationen

Publikation 1

Wessel JR, Danielmeier C, Ullsperger M (im Druck). Error awareness revisited: Accumulation of multi-modal evidence from central and autonomic nervous systems. *Journal of Cognitive Neuroscience*

Jan Reinhold Wessel designte die Hypothesen und den Ablauf für Experiment 2, programmierte Experiment 2, sammelte (mit studentischen und technischen Hilfskräften) die Daten für Experiment 2, wertete alle Daten beider Experimente aus (Verhalten, Elektroofokulogramm, Elektroenzephalogramm, Pupillometrie, Elektrokardiogramm) und schrieb das Manuskript.

Claudia Danielmeier designte die Hypothesen und den Ablauf für Experiment 1, programmierte Experiment 1, sammelte (mit studentischen und technischen Hilfskräften) die Daten für Experiment 1, und arbeitete an der Erstellung des Manuskriptes mit.

Markus Ullsperger designte die Hypothesen und den Ablauf für Experiment 1, supervidierte das Projekt und arbeitete an der Erstellung des Manuskriptes mit.

Publikation 2

Wessel JR, Ullsperger M (2011). Selection of independent components comprising event-related potentials: A data-driven approach for greater objectivity. *NeuroImage* 54(3), 2105-2115

Jan Reinhold Wessel plante und designte den Algorithmus, programmierte die MATLAB Algorithmen, designte und programmierte die grafische Oberfläche, plante und berechnete die Simulation des artifiziellen EEGs, plante die Hypothesen und berechnete die Validierungsstudie und schrieb das Manuskript.

Markus Ullsperger supervidierte das Projekt und arbeitete an der Erstellung des Manuskriptes mit.

Publikation 3

Danielmeier C, **Wessel JR**, Steinhauser M, Ullsperger M (2009). Modulation of the error-related negativity by response-conflict. *Psychophysiology* 46(6):1288-98

Claudia Danielmeier designte die Hypothesen und den Ablauf für Experiment 1 und 2, programmierte Experiment 1 und 2, erhob die Daten für Experiment 1 und 2, wertete Experiment 1 (behaviourales Experiment) aus und schrieb das Manuskript.

Jan Reinhold Wessel wertete alle elektroenzephalographischen Daten aus (Experiment 2) und arbeitete an der Erstellung des Manuskriptes mit.

Marco Steinhauser designte, programmierte und wertete die Computersimulation aus und arbeitete an der Erstellung des Manuskriptes mit.

Markus Ullsperger designte die Hypothesen und den Ablauf für Experiment 1 und 2, supervidierte das Projekt und arbeitete an der Erstellung des Manuskriptes mit.

Publikation 4

Ullsperger M, Harsay HA, **Wessel JR**, Ridderinkhof KR (2010). Conscious perception of errors and its relation to the anterior insula. *Brain Structure & Function* 214: 629 – 643.

Alle Koautoren arbeiteten an der Erstellung des Manuskriptes mit. **Jan Reinhold Wessel** schrieb vor allem die Teile „Autonomic nervous correlates of error awareness“ (S. 635 – 636) und „A perspective on integrating the AIC, the ANS, and error awareness: The orienting response account“ (S. 636 – 638).

5. Literaturverzeichnis

- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol Rev*, *108*(3), 624-652.
- Coles, M. G. (1989). Modern mind-brain reading: psychophysiology, physiology, and cognition. *Psychophysiology*, *26*(3), 251-269.
- Coles, M. G., Scheffers, M. K., & Holroyd, C. B. (2001). Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. *Biol Psychol*, *56*(3), 173-189.
- Comon, P. (1994). Independent Component Analysis, a New Concept. *Signal Processing*, *36*(3), 287-314.
- Critchley, H. D., Tang, J., Glaser, D., Butterworth, B., & Dolan, R. J. (2005). Anterior cingulate activity during error and autonomic response. *Neuroimage*, *27*(4), 885-895.
- Crone, E. A., van der Veen, F. M., van der Molen, M. W., Somsen, R. J., van Beek, B., & Jennings, J. R. (2003). Cardiac concomitants of feedback processing. *Biol Psychol*, *64*(1-2), 143-156.
- Danev, S. G., & de Winter, C. R. (1971). Heart rate deceleration after erroneous responses. A phenomenon complicating the use of heart rate variability for assessing mental load. *Psychol Forsch*, *35*(1), 27-34.
- Danielmeier, C., Wessel, J. R., Steinhauser, M., & Ullsperger, M. (2009). Modulation of the error-related negativity by response conflict. *Psychophysiology*, *46*(6), 1288-1298.
- Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-Trial Coupling of Concurrent Electroencephalogram and Functional Magnetic Resonance Imaging Identifies the Dynamics of Performance Monitoring. *The Journal of Neuroscience*, *25*(50), 11730-11737.
- Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, *5*(5), 303-305.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*, *134*(1), 9-21.

- Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage*, *34*(4), 1443-1449.
- Eichele, H., Juvodden, H. T., Ullsperger, M., & Eichele, T. (2010). Mal-adaptation of event-related EEG responses preceding performance errors. *Front Hum Neurosci*, *4*.
- Endrass, T., Reuter, B., & Kathmann, N. (2007). ERP correlates of conscious error recognition: aware and unaware errors in an antisaccade task. *Eur J Neurosci*, *26*(6), 1714-1720.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143-149.
- Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1990). Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In C. H. M. Brunia, A. W. K. Gaillard & A. Kok (Eds.), *Psychophysiological Brain Research* (Vol. 1, pp. 192-195): Tilburg University Press.
- Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biol Psychol*, *51*(2-3), 87-107.
- Garavan, H., Ross, T. J., Murphy, K., Roche, R. A., & Stein, E. A. (2002). Dissociable executive functions in the dynamic control of behavior: inhibition, error detection, and correction. *Neuroimage*, *17*(4), 1820-1829.
- Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*(6), 385-390.
- Gentsch, A., Ullsperger, P., & Ullsperger, M. (2009). Dissociable medial frontal negativities from a common monitoring system for self- and externally caused failure of goal achievement. *Neuroimage*, *47*(4), 2023-2030.
- Green, D., & Swets, J. (1966). Signal detection theory and psychophysics.
- Hajcak, G., McDonald, N., & Simons, R. F. (2003). To err is autonomic: error-related brain potentials, ANS activity, and post-error compensatory behavior. *Psychophysiology*, *40*(6), 895-903.
- Handy, T. C. (2005). *Event-related potentials: A methods handbook*. Cambridge, Mass. [u.a.: MIT Press.
- Harsay, H. A., Cohen, M. X., Spaan, M., Weeda, W. D., Nieuwenhuis, S., & Ridderinkhof, K. R. (submitted). Error awareness: Pupil dilation predicts opposing brain network dynamics during error awareness.

- Hester, R., Foxe, J. J., Molholm, S., Shpaner, M., & Garavan, H. (2005). Neural mechanisms involved in error processing: a comparison of errors made with and without awareness. *Neuroimage*, *27*(3), 602-608.
- Hester, R., Nandam, L. S., Wagner, J., Nathan, P. J., Mattingley, J. B., & Bellgrove, M. A. (2010). Improvement of Error Awareness and Modulation of Error-related BOLD Activity vis Single-dose Dopamine (Methylphenidate) in Healthy Adults. *Front. Hum. Neurosci. Conference Abstract: The 20th Annual Rotman Research Institute Conference, The frontal lobes*.
- Hoffmann, S., & Falkenstein, M. (2008). The correction of eye blink artefacts in the EEG: a comparison of two prominent methods. *PLoS One*, *3*(8), e3004.
- Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev*, *109*(4), 679-709.
- Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent Component Analysis*. New York: Wiley.
- Jutten, C., & Herault, J. (1991). Blind Separation of Sources .1. An Adaptive Algorithm Based on Neuromimetic Architecture. *Signal Processing*, *24*(1), 1-10.
- Kerns, J. G., Cohen, J. D., MacDonald, r., A.W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, *303*(5660), 1023-1026.
- Klein, T. A., Endrass, T., Kathmann, N., Neumann, J., von Cramon, D. Y., & Ullsperger, M. (2007). Neural correlates of error awareness. *Neuroimage*, *34*(4), 1774-1781.
- Magno, E., Foxe, J. J., Molholm, S., Robertson, I. H., & Garavan, H. (2006). The anterior cingulate and error avoidance. *J Neurosci*, *26*(18), 4769-4773.
- Maier, M., Steinhauser, M., & Hubner, R. (2008). Is the error-related negativity amplitude related to error detectability? Evidence from effects of different error types. *J Cogn Neurosci*, *20*(12), 2263-2273.
- Makeig, S., Jung, T. P., Bell, A. J., Ghahremani, D., & Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proc Natl Acad Sci U S A*, *94*(20), 10979-10984.
- McKeown, M. J., Makeig, S., Brown, G. G., Jung, T. P., Kindermann, S. S., Bell, A. J., et al. (1998). Analysis of fMRI data by blind separation into independent spatial components. *Hum Brain Mapp*, *6*(3), 160-188.

- Mennes, M., Wouters, H., Vanrumste, B., Lagae, L., & Stiers, P. (2010). Validation of ICA as a tool to remove eye movement artifacts from EEG/ERP. *Psychophysiology*.
- Narayanan, N. S., & Laubach, M. (2006). Top-down control of motor cortex ensembles by dorsomedial prefrontal cortex. *Neuron*, 52(5), 921-931.
- Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: evidence from an antisaccade task. *Psychophysiology*, 38(5), 752-760.
- O'Connell, R. G., Dockree, P. M., Bellgrove, M. A., Kelly, S. P., Hester, R., Garavan, H., et al. (2007). The role of cingulate cortex in the detection of errors with and without awareness: a high-density electrical mapping study. *Eur J Neurosci*, 25(8), 2571-2579.
- Overbeek, T. J. M., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable components of error processing: On the functional significance of the Pe vis-a-vis the ERN/Ne. [Print Electronic; Print]. *Journal of Psychophysiology*, 19(4), 319-329.
- Rabbitt, P. (1967). Time to detect errors as a function of factors affecting choice-response time. *Acta Psychol (Amst)*, 27, 131-142.
- Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, 306(5695), 443-447.
- Roger, C., Benar, C. G., Vidal, F., Hasbroucq, T., & Burle, B. (2010). Rostral Cingulate Zone and correct response monitoring: ICA and source localization evidences for the unicity of correct- and error-negativities. *Neuroimage*.
- Scheffers, M. K., & Coles, M. G. (2000). Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors. *J Exp Psychol Hum Percept Perform*, 26(1), 141-151.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593-1599.
- Shalgi, S., Barkan, I., & Deouell, L. Y. (2009). On the positive side of error processing: error-awareness positivity revisited. *Eur J Neurosci*, 29(7), 1522-1532.
- Sokolov, E. N. (1960). Neuronal models and the orienting reflex. In M. A. B. Brazier (Ed.), *The central nervous system and behavior*. New York: J. Macy.
- Steinhauser, M., & Yeung, N. (2010). Decision Processes in Human Performance Monitoring. *Journal of Neuroscience*, 30(46), 15643-15653.

- Ullsperger, M., Harsay, H. A., Wessel, J. R., & Ridderinkhof, K. R. (2010). Conscious perception of errors and its relation to the anterior insula. *Brain Struct Funct*, 214(5-6), 629-643.
- Ullsperger, M., & von Cramon, D. Y. (2001). Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *Neuroimage*, 14(6), 1387-1401.
- Vaughan, H. G., Jr., Costa, L. D., & Ritter, W. (1968). Topography of the human motor potential. *Electroencephalogr Clin Neurophysiol*, 25(1), 1-10.
- Wessel, J. R., Danielmeier, C., & Ullsperger, M. (in press). Error awareness revisited: Accumulation of multi-modal evidence from central and autonomic nervous systems. *Journal of Cognitive Neuroscience*.
- Wessel, J. R., & Ullsperger, M. (2011). Selection of independent components representing event-related brain potentials: A data-driven approach for greater objectivity. *Neuroimage*, 54(3), 2105-2115.
- Wilbert, J. (2006). *Die Etikettierung eines Verhaltens als fehlerhaft: Eine empirische Untersuchung zur bewussten Gewährwerdung eigener Verhaltensfehler*. Lengerich u.a: Pabst Science Publishers.
- Yeung, N., Cohen, J. D., & Botvinick, M. M. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev*, 111(4), 931-959.

6. Anhang

Publikation 1 (S. 1 – 39)

Wessel JR, Danielmeier C, Ullsperger M (im Druck). Error awareness revisited: Accumulation of multi-modal evidence from central and autonomic nervous systems. *Journal of Cognitive Neuroscience*

Supplementary Materials (S. 40 – 45)

Publikation 2 (S. 46 – 56)

Wessel JR, Ullsperger M (2011). Selection of independent components comprising event-related potentials: A data-driven approach for greater objectivity. *NeuroImage* 54(3), 2105-2115

Supplementary Materials (S. 57 – 72)

Publikation 3 (S. 73 – 83)

Danielmeier C, **Wessel JR**, Steinhauser M, Ullsperger M (2009). Modulation of the error-related negativity by response-conflict. *Psychophysiology* 46(6):1288-98

Publikation 4 (S. 84 – 98)

Ullsperger M, Harsay HA, **Wessel JR**, Ridderinkhof KR (2010). Conscious perception of errors and its relation to the anterior insula. *Brain Structure & Function* 214: 629 – 643.