

システム創成工学専攻
情報システムコース

Department of Systems Innovation
Advanced Information System Course

令和二年度 専攻科特別研究論文

ERP を用いた条件式の分岐判断の計測

ERP during a Cognitive Process of
Comparison Operation

指導教員名 上野 秀剛 准教授

論文提出者名 村井 溪介

独立行政法人 国立高等専門学校機構

奈良工業高等専門学校 専攻科

National Institute of Technology, Nara College
Faculty of Advanced Engineering

ERP を用いた条件式の分岐判断の計測

ERP during a Cognitive Process of Comparison Operation

村井 溪介
Keisuke MURAI

独立行政法人 国立高等専門学校機構

奈良工業高等専門学校 専攻科 システム創成工学専攻 情報システムコース

大和郡山市矢田町 22 番地 (〒 639-1080)

National Institute of Technology, Nara College, Faculty of Advanced Engineering

22 Yata-cho, Yamatokoriyama, Nara 639-1080, Japan

Abstract:

Cognitive processes and mental states of software developers have been investigated using electroencephalogram (EEG) in software engineering. However, the feasibility of event-related potential (ERP) analyses for program comprehension were rarely examined. In this study, the aim is to discover the ERPs related to program comprehension. Conducted an experiment to capture ERPs with ten software developers during a cognitive process of comparison operation as one of the simple program comprehension processes. As results, ERP waveforms were observed in the EEG data of four subjects. In addition, the inter-stimulus difference in peak amplitudes and peak latency is confirmed.

Keywords: Event-Related Potential, ERP, SE, Conditional Expressions, Program Comprehension, EEG;

関連業績リスト

1. Keisuke Murai, Hidetake Uwano, Yoshiharu Ikutani, and Takatomi Kubo, "Discovery of Event-Related Potentials during A Cognitive Process of Comparison Operation", In Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2020, December 2020.

目次

1.	序論	1
2.	関連研究	3
3.	脳波	5
3.1	周波数成分	5
3.2	計測方法	6
3.3	事象関連電位	7
4.	実験方法	10
4.1	被験者	10
4.2	実験デザイン	10
4.3	データ収集と処理方法	11
4.4	研究課題	12
5.	結果と考察	14
5.1	RQ1: ERP の検出	14
5.2	RQ2: True と False の ERP の違い	16
6.	問題と改善	20
7.	まとめ	22
	参考文献	24

目次

3.1	国際式 10-20 電極配置法 [27]	6
3.2	事象関連電位 ([28] をもとに宮本が加筆 [24])	8
3.3	ERP と各成分の解釈のモデル [11]	9
4.1	条件式タスクのデザイン	11
5.1	各被験者の平均化された脳波	15
5.2	$grandAve_{True}$ と $grandAve_{False}$	16
5.3	$indiv_{True}$ と $indiv_{False}$	17

表目次

4.1	各成分の抽出区間	12
5.1	各成分の頂点振幅と頂点潜時	18

1. 序論

ソフトウェア開発者は日々の業務の中でたくさんの意思決定を行う。その中には、変数命名などの比較的小さなものから、アーキテクチャ選定などシステム全体に影響を与える重大なものまで存在する。このような意思決定の多くは開発者の勘や経験に依存し、その判断根拠や経緯を外部から推測することは難しい [1,18]。加えて、ソフトウェア開発者がどんなタイミングで意思決定を行っているかについても、明らかになっていない。開発者がなんらかの気づきや意思決定を行った時点を捉え、そのときにどんな情報を見ていたかが分かれば、開発者が無意識に行っている意思決定の根拠や経緯を定量的に推測できると期待される。近年、視線や脳活動などの生体信号を用いて、ソフトウェア開発者の心理や認知プロセスを定量的に捉える試みがなされている。これらの生体計測の中でも、脳波は計測が容易かつ低コストで、集中度合いなどの被験者の心理状態を反映することが知られている [2,12]。脳波を用いた分析手法は大別して、(1) 周波数成分のトレンドを見るもの、(2) Event-related potential(ERP) を捉えるものが存在する。特に ERP は与えられた刺激に対しての反応性をミリ秒オーダーで評価するため [14]、ソフトウェア開発者の気づきや意思決定の発生タイミングを検出できる可能性がある。

しかし、ソフトウェア工学分野の脳波を用いた既存研究では、ほとんどが周波数成分のトレンドのみに着目しており、ソフトウェア開発における ERP については明らかにされていない。本研究では、プログラム理解時の開発者の ERP を捉えられるか検証することを目的とする。実験では 10 名の被験者が 2 つの数値の大小関係の判定を行うタスクを多数繰り返しているときの脳波を計測する。そして、刺激提示タイミング (stimulus-onset) を基準として脳波を平均化した場合に、ERP のような電位変動が見られるかを調査する。特に ERP に現れる個人差の影響、およびタスク条件の違いによる ERP の形状の差に着目し、それぞれについて分析する。

本研究の主要な貢献は以下の通りである。

1. ソフトウェア工学分野でも ERP が出ることが示せた.
2. 被験者間, 真偽間の ERP の形状は異なることが示せた.

2. 関連研究

近年、視線や脳活動などの生体信号を用いて、ソフトウェア開発者の心理や認知プロセスを定量化しようとする研究が注目を集めている。特にプログラム理解はソフトウェア開発における主要な認知活動のひとつであり、開発者は作業時間の 50% 以上を費やしているため [22]、プログラム理解時のソフトウェア開発者の認知活動を分析する試みがなされている。例えば、Lee らはプログラム理解時の熟練者と初心者の脳波を比較し、熟練者の脳波の β, γ 成分が初心者よりも高くなることを示している [10]。

Frits らは脳波、視線、筋電位の 3 つの生体信号を基に、開発者が取り組んでいるプログラム理解タスクを難しいと感じているか否かを約 70% の精度で推定できることを示した [2]。Müger らは、先の例に心拍を加えた 4 つの生体信号を基に、開発者の心理状態 (Stuck か In Flow, Happy か Frustrated) をそれぞれ 71.4%, 67.7% の精度で識別できることを示した [12]。しかし、脳波を用いたこれらの既存研究は主として周波数成分のトレンドのみに着目しており、プログラム理解時に観測される ERP については、ほとんど明らかにされていない。

ERP は物理的・心理的事象の発生タイミングに時間的に関連して起こる電位変動である [14]。ERP を用いた分析は脳波の時間変化についてミリ秒オーダーで着目するため、被験者の認知処理がいつ発生したかを詳細に分析することができる。

Kutas と Hillyard らの初期の研究は、人間が **grammatical errors** に直面したときに、ERP の N400 反応 (i.e. 基準時点から 400ms 後に観測される Negative な電位) が観測されることを示した [8]。その後、1000 を超える研究が、N400 と言語処理や物体認識などの様々な認知処理を関連づけている [7]。Lau らは、言語処理における N400 成分の観測と、**functional magnetic resonance imaging** を用いて得られた高い空間分解能の証拠を組み合わせることにより、人間の脳に存在する言語ネットワークのモデル化を試みている [9]。ソフトウェア工学分野においても、時間解像度の高い ERP に着目した研究と、fMRI を用いた研究結果を組み合わせることで、人のプログラム理解活動の理解を推進で

きると期待される.

3. 脳波

人間の頭部に2つの電極を貼り付けると、その間にわずかな電位差が生じる。その大きさは数十マイクロボルトであるが、数万倍に増幅するとリズムを持った波が観測できる [28]。これが脳波 (EEG:Electroencephalogram) であり、人の脳活動を非侵襲で測定できるため、様々な分野で用いられている。

3.1 周波数成分

脳波には様々な周波数帯域があり、国際脳波学会によって分類及び名称が定められている。各帯域の名称と周波数は以下の通りである。

- δ 波 (0.5~8Hz)
- θ 波 (4~8Hz)
- α 波 (8~13Hz)
- β 波 (13~20Hz)
- γ 波 (30Hz 以上)

α 波は、静かな暗室で目を閉じ、何も考えないようにしたときに見られる。覚醒が低下すると α 波の振幅が低下し、不連続になる。考えごとをしたり、精神的に興奮すると α 波は減少して、 β 波が出現する。睡眠状態に移行すると δ 波や θ 波や出現し、これらは周波数が低いため徐波と呼ばれる [28-30]。

以上のように、脳波は心理状態の変化、活動状態によって発生する周波数が異なるため、意識や感情・学習などといった心理学に関連する分野で用いられている。また、病的状態の場合には、目覚めている状態でも徐波が観測されたり、周期では分類しにくいような特異な波形が現れることがあり、臨床的にも応用されている [23]。

3.2 計測方法

脳波は、頭皮上に装着した電極から計測される。電極をそのまま頭皮に接触させると分極が起こり、電位を導出できないので、頭皮と電極の間には電解質を含んだ電解糊を介在させる必要がある。したがって、電極を装着する際は予め装着する部位をよく脱脂した上で、電解糊をつけた電極を圧着する。電極の配置は1958年の国際脳波学会で標準化された図3.1に示す国際式10-20電極配置法[28]が広く使用されている[30]。国際式10-20電極配置法では、耳のアースを除き19箇所（F₇, F₃, F_z, F₄, F₈, C₃, C_z, C₄, T₃, T_z, T₄, P₃, P_z, P₄, O₁, O₂）の装着位置が指定されているが、実際には、検査や研究の目的によって使用する電極の数が決まる[23,26]。

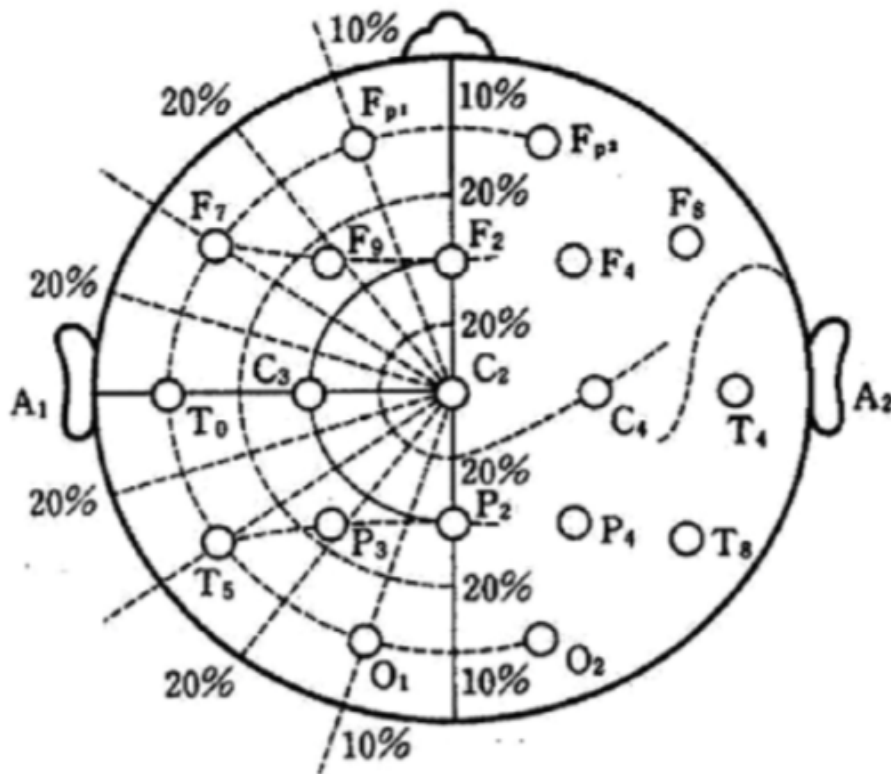


図 3.1 国際式 10-20 電極配置法 [27]

脳波の導出法には主に基準電極導出法と双極導出法の2種類がある。基準電極導出法では、脳電位の電場内に装着した計測用電極と、電場外に装着した基準電極の2つの電極の

電位差として脳電位を測定する。双極導出法は、基準電極を用いず、2つの計測用電極を脳電位の電場内に置いてその電位差を脳電位として測定する。一般に脳波には、脳の限局した領域に発生するものと、比較的広い範囲から同じように記録されるものがある。双極導出の2つの計測用電極の電極間隔が狭い場合には、脳の広い範囲から同じように記録される要素は、両方の電極にほぼ同じように記録されるため、相殺されてほとんど記録されない。したがって、計測用電極の電極間隔が狭い場合に電位差を計測する際は基準電極導出法を、優勢な背景成分を除去して、部位差を強調する目的で計測する際には双極導出法を選択する [26,29]。

3.3 事象関連電位

不自然なことや特異な事象などの外部からの刺激を認識した際の脳波の変動を事象関連電位 (ERP: Event-Related Potential) という。図 3.2 に ERP の例を示す [28]。高さの異なる二つの音をランダムに提示し、より高い音が提示された数を数えるように指示したときの ERP 波形である。横軸は音を提示してからの時間、縦軸は電位（上が負、下が正）を表し、太線が高い音が提示されたときの波形、細線は低い音が提示されたときの波形である。高い音を提示してから約 300ms 後に正に電位が振れていることがわかり、ERP が観測されていることが確認できる。

ERP の計測では、多数の刺激提示と計測から得られた脳波データを刺激の提示時点にそろえて加算平均する方法がよく用いられる。ERP は微小な電位変動であるため、通常の脳波成分に埋没しており、1回の刺激提示によって得られるデータ（単一試行脳波波形）から ERP を抽出することは難しい。しかし、脳波全体の振幅はゼロマイクロボルトを中心とした正規分布であるため、複数回計測し、刺激提示時点を統一して加算平均をとることで、ERP が発生している部分以外の振幅は抑えることができる。一方、ERP は刺激の提示時点から一定の反応時間で生じるため、加算を重ねるごとにその振幅は増大していく。以上のように、加算平均処理によって事象とは時間的に無関係に生じる脳波を相殺することで、ERP 成分のみを得ることができる [23,29]。被験者ごとに加算平均処理された脳波を個人 ERP 波形と呼び、全ての個人 ERP 波形を加算平均処理した波形を総加算平均 ERP 波形と呼ぶ。

ERP を検出する上で、ある事象に対しての被験者の脳波データが多数必要であることに加え、数ミリ秒で起こる小さな電位変動という ERP の特徴に配慮する必要がある。

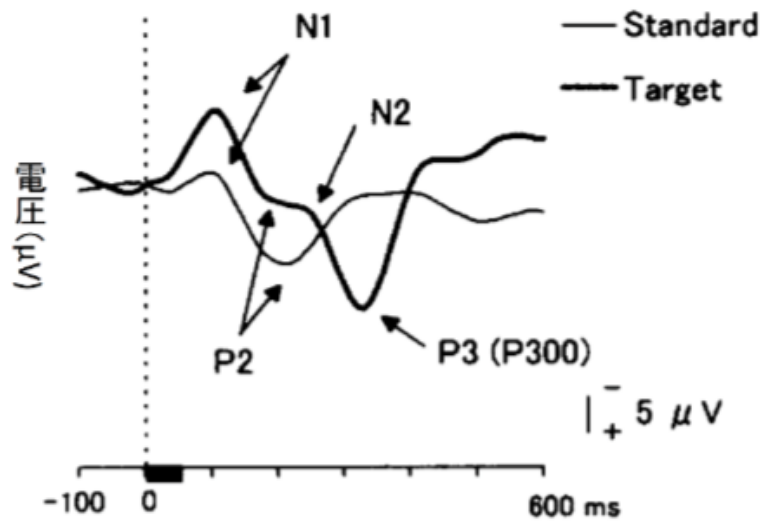


図 3.2 事象関連電位 ([28] をもとに宮本が加筆 [24])

そのため実験に用いるタスクには以下の点に注意することが望ましい。

- 1 タスクが簡単であること
- 1 タスクが短時間であること
- 全てのタスクの時間が統一されていること

ERP の波形の特性を説明するために、図 3.3 に ERP 波形をサンプルとして追加で示す。図より、横軸は時間 (ms)、縦軸は電位 (μV) を表している。ERP を用いた議論では、ERP 波形上の成分 (上下に尖った比較的大きなピーク) の振幅 (頂点振幅) や潜時時間 (頂点潜時) について議論することが一般的である。図 3.3 のように各成分には出現順序と極性に基づいて P1 や N1, P2 のようなラベリングすることが一般的である。また図 3.2 のように 300msec 付近に出現した成分を P300 とラベリングすることもある。ERP を示す図において、縦軸の上方向をネガティブにする (e.g. 図 3.3) かポジティブを上方向にする (e.g. 図 3.2) かは研究者によって異なることが多い。

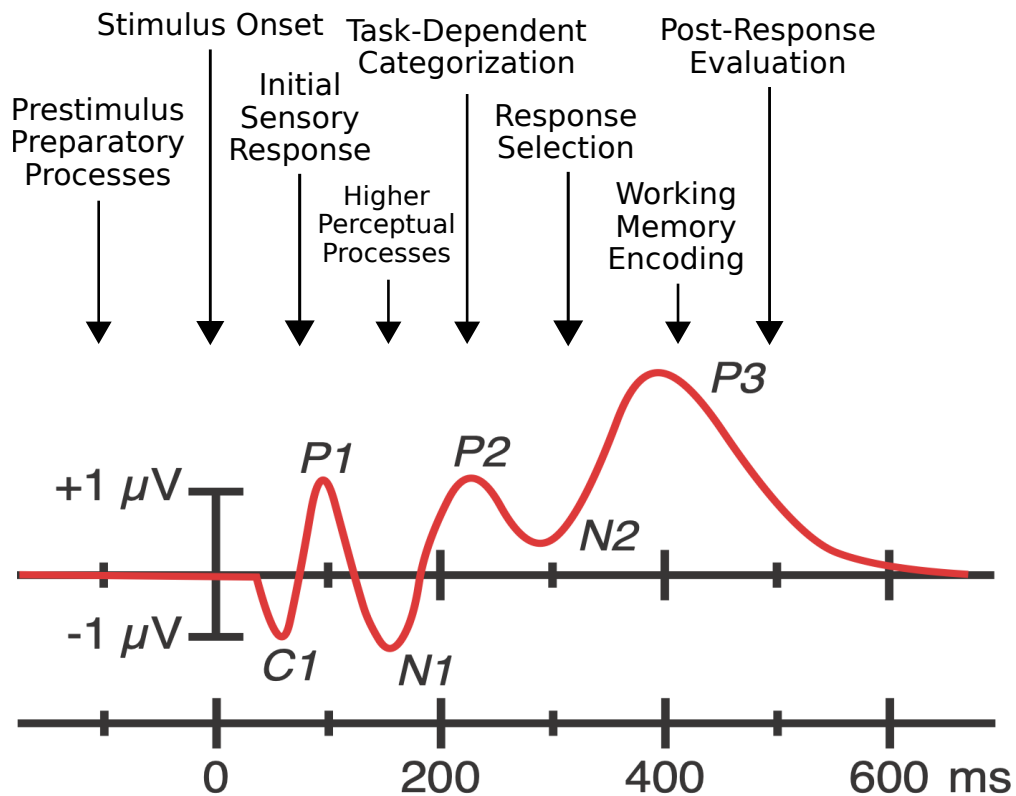


図 3.3 ERP と各成分の解釈のモデル [11]

ERP の波形には以下の特徴がある。実験で得られた波形が ERP であるか否かについては、定義が厳密に定められていないため、現状これらの特徴をもつ加算平均処理された脳波は ERP とされ議論されることが多い。

- 出現が早い成分は比較的先鋭であり、出現が遅い成分ほどなだらかである。
- ERP は極性が交互に出現する成分を持つ。

4. 実験方法

4.1 被験者

本研究の実験では、奈良工業高等専門学校と奈良先端科学技術大学院大学からの計 11 名の被験者（全員男性，19～37 歳）を対象とした。すべての被験者は、正常または正常に矯正された視力を持ち、基本レベルの Java 文法を理解している。被験者の内 1 名について、実験データが正しく計測されなかったため、分析から除外した。

4.2 実験デザイン

ソフトウェア開発者はプログラム理解中に様々な意思決定を行う。本研究では、様々な意思決定の中でも、条件式の分岐判断に着目する。条件式の判定に着目した理由について、何度も繰り返すことができるほど単純なプロセスであり、開発者がこのような判定プロセスに直面することが多いからである。また、プログラミング関連での活動を観察する際の ERP を抽出するために、被験者に提示される条件式は True/False かを判断するだけのシンプルなタスクに設定した（図 4.1）。本実験タスクは ERP を検出する上での制約（3.3 節参照）も満たしている意思決定タスクである。

各試行では、2 つの 1 桁の整数 (0-9) と 1 つの演算子 (!=, ==, >, >=, <, <=) をから構成される条件式を 1.5 秒間被験者に提示した。その後、視線注視マーク（‘+’）を 1.5 秒間表示し、被験者は「True」または「False」に対応するボタンを押すことで回答した。全ての条件式で使用した表記法は Java 言語の文法に従った。条件式の提示と回答の時間を意図的に分けたのは、特に指の動きに関連した運動に起因する潜在的なアーチファクトを軽減するためである。

実験は 10 回のセッションで構成され、各セッションには条件式タスクの 102 回の試行が含まれている（各被験者の合計 1,020 回）。タスクは全体を通してサンプル数のバランス

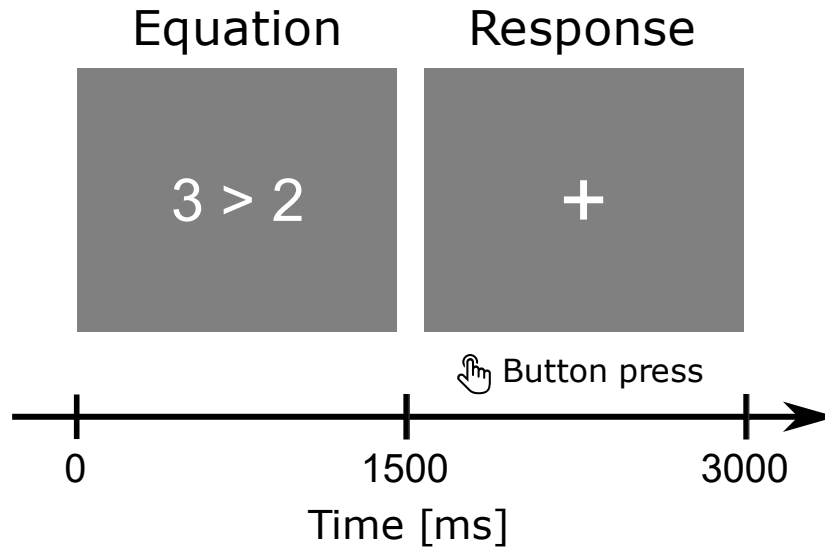


図 4.1 条件式タスクのデザイン

をとるために、510回の試行を True とし、残りの 510回を False となるように構成した。さらに、各演算子を含む試行回数もバランスをとり、各演算子について 170 回の試行となるように構成した。提示方法について、被験者にタスクの提示順から以降に提示されるタスクの回答を予測されることを避けるため、タスク提示順序をあらかじめランダムで設定した。さらに、被験者間で提示順序による回答への影響を一定にするため、全被験者は同じ表示順で実験を行った。刺激提示とすべての行動反応の記録には、PsychoPy (version 1.90.3, [13]) を使用した。

4.3 データ収集と処理方法

NeXus-10 MARK II (Nanotech Image Ltd.) を用いて、被験者がタスクを実施している間、すべての被験者の脳波データを記録した。被験者の脳波を 4 つの異なる電極位置 (10-20 電極法の配置に基づき Fpz, Fz, Cz, Pz) から 256Hz のサンプリング周波数で測定し、40,800 件のデータを得た (1,020 回の試行を 10 人の被験者で 4 チャンネル)。

各試行に対する反応とその回答は、行動データとして全被験者から記録された。分析の前処理として、まず、験者が正しく回答していない (回答が誤り、無回答、回答時間外での回答) 試行を除外する。この前処理により、372 件のデータが除外され、40,428 件の分析対象データが残った。次に、残りの脳波データに無限インパルス応答 (IIR) バンドパス

表 4.1 各成分の抽出区間

成分	抽出区間 (msec)
C1	0 – 200
P1	150 – 300
N1	270 – 400
P2	400 – 550
N2	550 – 900
P3	840 – 1100

フィルタを適用して、タスクに係るアーチファクトを除去した。バンドパスフィルタの低域カットオフ周波数は 2Hz、高域カットオフ周波数は 45Hz であった。この前処理は、各被験者の各試行の脳波データに対して独立して適用した。前処理した脳波データを用いて、その波形が ERP の波形と一致しているかどうかを調べた (Fig.3.3, [11])。

また、脳波から ERP の成分を抽出する。陰極に凸な成分を出現が早い順に C1,N1,N2 とし、陽極に凸な成分も出現が早い順に P1,P2,P3 とする。全被験者の波形の形状を参考に、各成分に対して時間的な区間を著者らが設定し、それぞれ設定する。表 4.1 に各成分の抽出区間を示す。区間内 (Table.4.1) で振幅の絶対値 (以降、ピーク値) が最大な点を求めて、成分の電位と時間とする。区間内にピーク値が複数存在する場合、より早い時間に出現した点を用いる。

4.4 研究課題

私たちは研究目的を達成するために、以下の 2 つ研究課題を用意した。

RQ1 簡単な条件式判断タスク中に計測された脳波データに ERP が現れるか?

RQ2 条件式を True と判断した時と False と判断した時の ERP に違いがあるのか?

RQ1 に回答するために、Fig.3.3 に示すモデルの ERP 波形と前処理した脳波を比較する。開発者の認知プロセスを理解することは、IDE や他のソフトウェア開発システムでの状況に応じた開発者サポートに役立つと期待している。タスク条件によって ERP が異なるかどうかを検証するために RQ2 を設定した。ERP の波形は、タスクの種類や性質によって成分の頂点振幅や頂点潜時が異なることが多くの研究で示されている [8,15]。RQ2

に回答するために、True/False と判断した時のそれぞれの ERP から成分を抽出し、成分間での比較をする。本論文では、結果をできるだけ簡単に示すために、4つのチャンネルの中で最も特徴的な波形を示した Fpz チャンネルの脳波データの結果のみを対象とする。

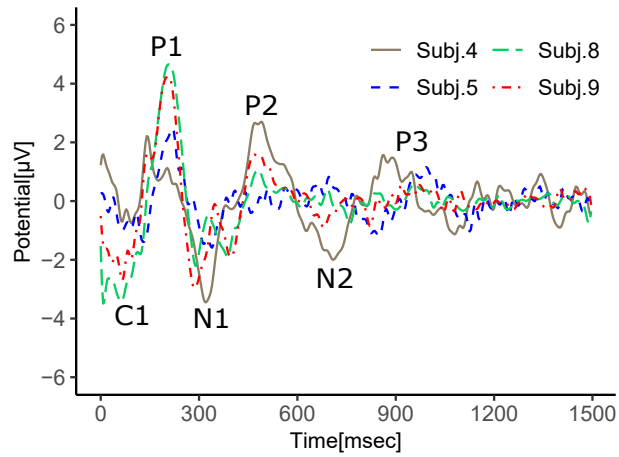
5. 結果と考察

5.1 RQ1: ERP の検出

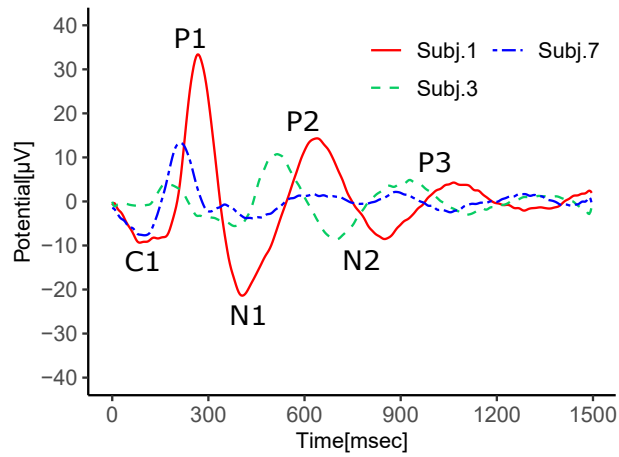
既存研究では、被験者間やタスク間で ERP 波形が異なることが明らかになっている [7,14]. ここでは、まず本実験タスク中に ERP 波形が現れたかどうかを調べるために、個々の被験者の平均化された脳波を調べた. 図 5.1 に全被験者の平均波形を示す. 各波形は各被験者ごとに独立しており、刺激提示タイミングに基づき、すべての脳波データの平均化より得られた. ラベル (e.g. C1,P1) は著者より付与された. 図 5.1(a) より 4 人の被験者は、100-150 msec (C1), 150-250 msec (P1), 300-350 msec (N1) などの成分を持つ明確な ERP 波形を示した. これらの波形を以降 ClearERPs と呼ぶ. 図 5.1(b) より 3 人の被験者は、形状は ERP であるが、40 - 30 μ V と ERP には大きすぎる電位を持つ. これらの波形を以降 PlausibleERPs と呼ぶ. しかしながら、図 5.1(c) に示す残りの 3 人の被験者は、周期的な波形に埋もれた可能性のある異なる波形を示した. これらの波形を NoERPs と呼ぶ.

結果として、明瞭な ERP を 4 人の被験者の脳波データから得ることができた. 図 5.1(a) より、先行の 4 つ成分 (C1,P1,N1,P2) の頂点潜時は、これらの被験者間でほぼ一致していることが示された. 一方、後続の 2 つの成分 (N2 と P3) の頂点潜時と頂点振幅は、他の成分よりもばらつきがあるように見える. このような波形の違いは、内部的な認知要求の違いが影響している可能性がある. 具体的に、先行のいくつかの成分は視覚的な反応に関連しているが、後続の成分はより認知的な要求に関連している可能性がある.

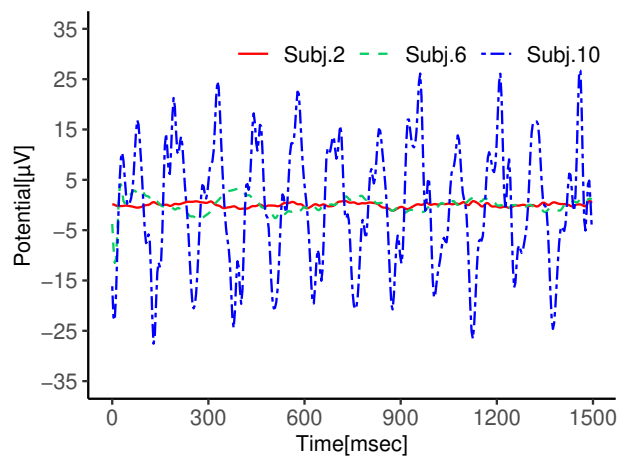
PlausibleERPs (図 5.1(b)) は ClearERPs (図 5.1(a)) より大きな頂点振幅を持つ波形である. 頂点振幅の被験者間の変動は、多数の脳波研究で観測されている [21]. NoERPs (図 5.1(c)) は ERP のような波形を持たない 3 人の平均波形を示す. Subj.10 で観察された周期的な波形は、頭の動きなどの周期的なアーチファクトの影響を受けている可能性がある. また、明確な成分を持たない波形は被験者 (Subj.2 と 6) に有意な ERP がないこ



(a) 4人の被験者より観測されたERP.



(b) 3人の被験者より観測された波形。ERPのような形状ではあるが、振幅が大きすぎる。



(c) 3人の被験者よりERPを伴わない脳波が観測された。

図 5.1 各被験者の平均化された脳波

とを示す可能性がある。

以上の各被験者の波形の結果から、RQ1 の回答は以下のようになる。

RQ1 への回答

10 人中 4 人の被験者の脳波データに ERP 波形が現れた。他の 3 名の被験者は不明瞭ではあるがもっともらしい ERP 波形を示した。

5.2 RQ2: True と False の ERP の違い

本章では、明瞭な ERP 波形を示した 4 人の被験者 (Subj.4,5,8,9) に焦点を当て、True と False の違いを分析した。

図 5.2 は 4 人の被験者による True の総加算平均 ERP 波形 (以降, $grandAve_{True}$) と False の総加算平均波形 (以降, $grandAve_{False}$) を示す。図は、C1, P1, N1, および P2 持つ総加算平均 ERP 波形 ($grandAve_{True}$ と $grandAve_{False}$) である。N2 以降の成分の頂点振幅はやや小さい。加えて、 $grandAve_{True}$ と $grandAve_{False}$ の違いは明瞭でない。 $grandAve_{True}$ の N2 と P3 の頂点潜時は、 $grandAve_{False}$ よりも早い。他の成分 (C1, P1, N1, P2) の頂点振幅と頂点潜時はほぼ同じである。

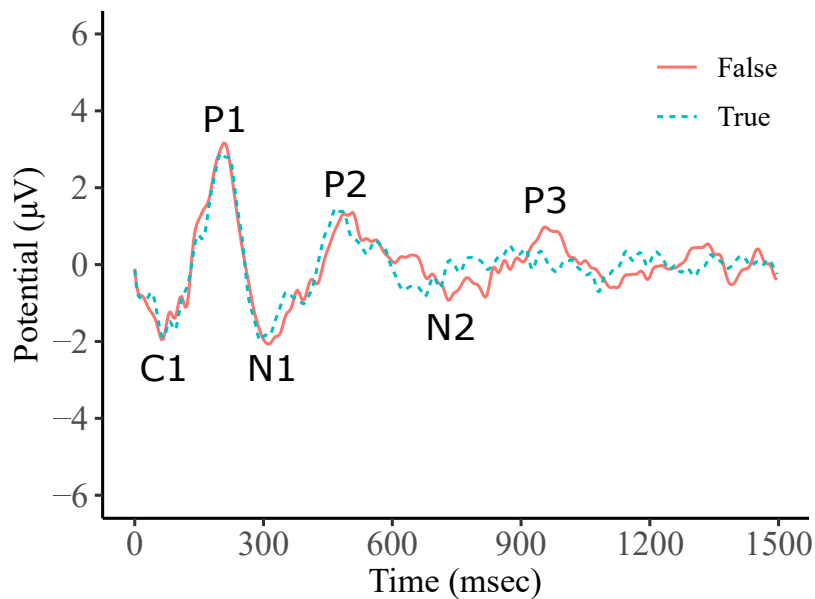
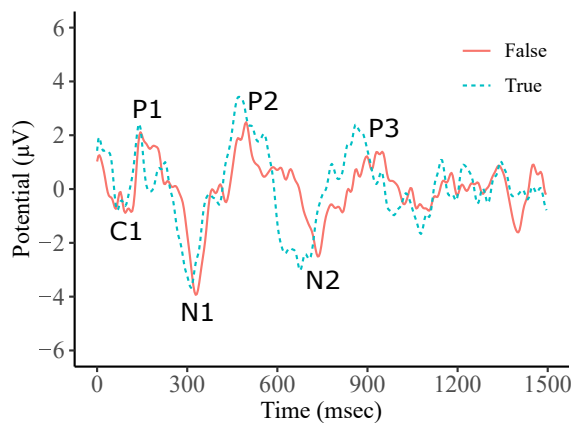
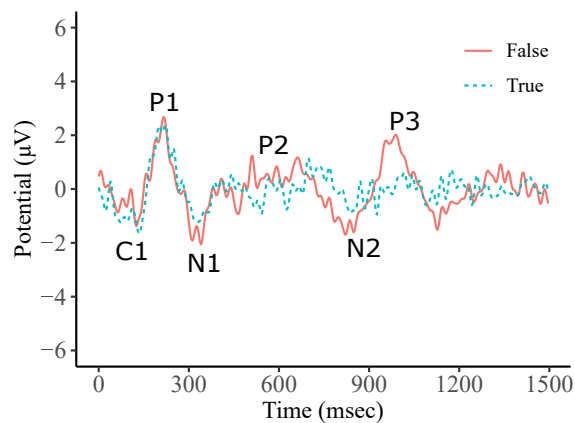


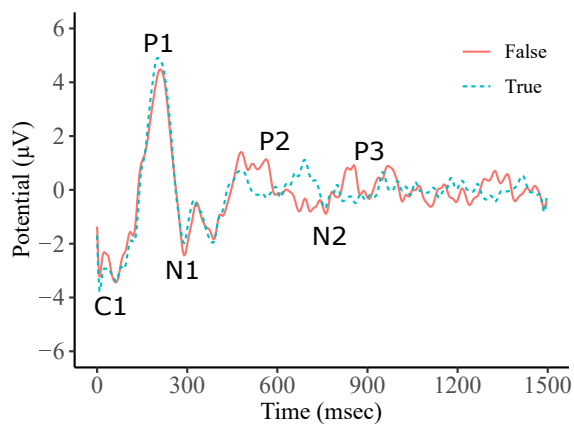
図 5.2 $grandAve_{True}$ と $grandAve_{False}$



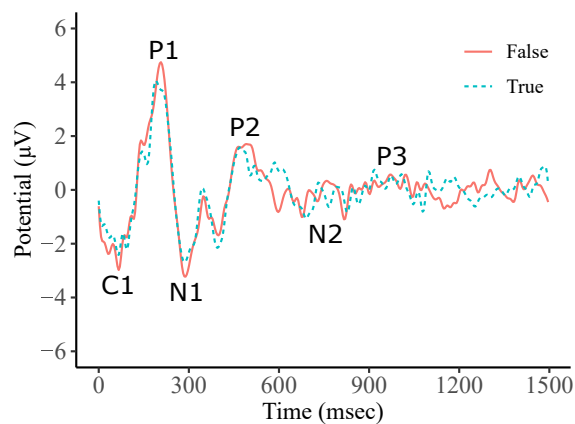
(a) Subj.4



(b) Subj.5



(c) Subj.8



(d) Subj.9

図 5.3 $indiv_{True}$ と $indiv_{False}$

結果が $grandAve_{True}$ と $grandAve_{False}$ の間で異なる特徴を示さなかった理由の 1 つには、個人差があるかもしれない。本実験タスクでは、各条件式を判断するための思考を被験者に要求している。各被験者は条件式の判断に必要な時間が異なるため、各被験者の各成分の頂点潜時が異なる時間帯に現れる可能性がある。図 5.3 は、各被験者が条件式を True と判断した時の ERP 波形 (以降, $indiv_{True}$) と、False と判断した時の ERP 波形 (以降, $indiv_{False}$) を示している。図より、個々の ERP 波形は明確な ERP 成分を持っているが、頂点振幅や頂点潜時は被験者によって異なる。そこで、個々の ERP 波形から各成分を抽出し (4.3 章参照)、その傾向を分析する。

表 5.1 各成分の頂点振幅と頂点潜時

タイプ		頂点振幅 (μV)						頂点潜時 (msec)					
		C1	P1	N1	P2	N2	P3	C1	P1	N1	P2	N2	P3
平均	False	-2.17	3.48	-2.91	1.71	-1.55	1.22	90.8	199.2	315.5	499.0	787.1	945.3
	True	-2.15	3.30	-2.39	1.57	-1.42	1.10	72.3	195.3	308.6	468.8	744.1	940.4
標準偏差	False	1.06	1.16	0.72	0.48	0.63	0.54	25.1	27.2	22.8	11.2	36.5	51.5
	True	1.09	1.23	0.90	1.14	0.94	0.75	44.2	27.2	18.9	13.8	63.2	56.6

表 Table.5.1 は、各成分の頂点振幅と頂点潜時を示している。 $indiv_{False}$ の全ての成分は、 $indiv_{True}$ の成分に比べて、頂点振幅の絶対値が大きく、頂点潜時が遅いことが分かる。特に N1 の頂点振幅は他の成分に比べて大きな差がある。全ての被験者の $indiv_{False}$ において、N1 他の成分より頂点潜時の絶対値が大きい。本論文では、十分なサンプル数がないため統計解析はしていないが、 $indiv_{True}$ と $indiv_{False}$ の比較で示された 2 つの特徴について考察する。

まず、結果は $indiv_{False}$ の方が $indiv_{True}$ の方よりも頂点潜時が遅いこと。 $indiv_{True}$ と $indiv_{False}$ の成分の時間差はタスクの難易度による違いを表している可能性がある。認知心理学の領域において、2 つのタスクを同時に遂行する際にパフォーマンスが低下する (二重課題干渉) 現象は古くから知られている [17,20]。二重課題干渉は二重課題によって、被験者のメンタルワークロード (MWL: 精神的作業負荷) が注意リソースの容量制限を超えるために発生すると説明され、神経科学の領域においても数多く研究されている。

木村と武田は視覚刺激のタスク難易度が ERP のある成分 (DRM) の頂点潜時に影響することを報告している [4]。本研究で被験者に与えたタスクは二重課題ではないものの、 $indiv_{False}$ と $indiv_{True}$ の間に、被験者の MWL に与える影響に差があり、それが頂点潜時として表れている可能性がある。

第二に、 $indiv_{False}$ は、すべての成分において、 $indiv_{True}$ よりも頂点振幅の絶対値が大きく、特に N1 では $indiv_{False}$ の方が、 $indiv_{True}$ よりも頂点振幅の絶対値が大きいことである。N1 の頂点振幅は、一般的に初期の注意選択プロセスと知覚資源の割り当てに関連する指標である [5,6,19]。Solís-Marcos と Kircher は MWL の異なる単一課題、二重課題、三重課題を行う被験者の ERP を計測し、主観的な MWL が低い単一課題において、二重課題、三重課題よりも N1 の頂点振幅が低くなったことを示している [19]。本研究のタスクにおいても、 $indiv_{False}$ と $indiv_{True}$ で被験者に与える MWL が異なり、その結果、

頂点潜時に加え，頂点振幅に影響している可能性がある。

よって， $indiv_{True}$ と $indiv_{False}$ の比較に基づくすべての結果から，RQ2 への回答は次のようになる。

RQ2 への回答

条件式の結果が True と時と False の時で ERP の波形は異なる。平均して $indiv_{False}$ において， $indiv_{True}$ より頂点振幅の絶対値が大きく，また頂点潜時が遅い。

P3 は MWL の計測での生体指標としてよく用いられている。また P3 は注意量や弁別に係る意味的な機能を反映していることが知られている。今回のタスクは単一課題ではあるものの，ERP と MWL を組み合わせることによって，True/False の弁別に関する注意量を定量的に計測できる可能性がある。True/False の判断に関わらず，プログラム理解に関する様々な意思決定の注意量を定量化することが可能になれば，プログラムに対する客観的な困難度の指標として教育支援ツール等による支援や，学習者に適した課題の付与等に役立つと考えている。

6. 問題と改善

本研究の実験結果ではサンプル数は少ないものの、先行の成分が比較的明瞭に得られているが、後続の成分は不明瞭である。一般的に、ERP の初期の成分は低レベルの処理である、刺激の基本的な感覚処理を反映し、後期の成分は高レベルな処理での刺激の知覚・認知処理を反映している [16]。先行する成分は低レベルな処理に対応しているため、各試行で提示されるタスクに対して、ERP 成分が惹起するタイミング差（潜時ジッター）が小さいため、明瞭な形状が得られたと考えられる。一方で、後続の成分は意味的な理解や認知を伴う処理に対応しているため、被験者間およびタスク間で潜時ジッターが大きく、不明瞭な形状の成分となった可能性がある。

特に P3 の振幅に関して、Johnson は三次元モデルを提案している [3]。モデルは、主観的な確率（標的の出現がまれであれば高振幅）と刺激の意味（有意味性が高いほど高振幅）と情報伝達（や不注意によって失われる情報が少ないほど、高振幅）の次元が乗算的に影響することで、P3 の振幅が決定すると主張している。この三次元モデルの主観的な確率について本実験タスクでは、True/False および、各演算子のサンプル数を統制したため、特異な要素（True となるタスクが珍しい等）は存在しない。しかし、意味的な判断を伴う本タスクは、刺激の意味と情報伝達の指標で複雑に関連しているため、明瞭な高い振幅が得られなかった可能性がある。

P3 の振幅に関して他に、覚醒や疲労、摂食などが影響することが知られている [25]。今回の計測実験では 1 人当たりの被験者に課すタスク量が膨大であり、同じ姿勢で長時間でディスプレイを見続ける必要があったため、疲労による影響が P3 の出現を妨げた可能性も考えられる。そのため、今後は被験者への負荷に配慮しつつ、実験方法の再検討をすることが望ましい。P3 は他の成分に比べ信号/ノイズ比が高いため、少ないサンプル数でも検討できることが知られている [25]。よって加算平均処理に必要なサンプル数は減少するが、分析対象サンプルを実験による疲労が少ない計測初期段階でのデータのみで加算平均処理し ERP を観察すると、より明瞭な P3 の検出に有効であると考えている。

今後実験をするときに明瞭な成分が得られなかった場合については、振幅や潜時の定量化の方法を変更することも有効と考えられる。本研究では、頂点振幅をその成分の振幅の代表値として分析をした。しかし、後続の成分になるほど、ノイズ成分が多く成分の形状が不明瞭である。平均振幅や符号付き領域振幅等の時間窓を用いた手法を用いることで、安定性が高くやノイズや潜時ジッターの影響を受けにくい分析ができる可能性がある [31].

7. まとめ

プログラム理解時に、ERPを検出できるのかという目的を達成するために、被験者に簡単な条件式判断をさせ、その時の脳波を測定した。結果として、10人中7人にERPが検出されたため、目的を達成した。7人中、ClearERPは4人、PlausibleERPが3人だった。また、被験者間のERPは頂点振幅や頂点潜時には個人差が見られた。TrueとFalse間を比較すると、両波形は異なることが分かった。平均して、Falseの方がPeak amplitudeは大きく、Peak latencyは遅かった。特にN1については他の成分と比べて振幅差が大きく、全被験者でFalseの絶対値が大きかった。

今後の課題として、なぜ $indiv_{False}$ は $indiv_{True}$ よりも頂点振幅が高く、頂点潜時が長いのかを考察する必要がある。さらに、分析のデータ件数が少ないため統計検定が未実施である。よって追加実験が必要である。また、ERPの抽出方法や成分の選択についても考慮する必要がある。一般的に、ERPの初期の成分は低レベルの処理での刺激の基本的な感覚処理を反映し、後期の成分は高レベルの処理での刺激の知覚・認知処理を反映している [16]。プログラム理解というタスクの性質からは後期の成分の分析は興味深い研究対象である。そのため、P3 (third positive peak) などの後期の成分を計測するための実験設定の構築は重要な課題の1つである。

謝辞

本論文の執筆および研究をすすめるにあたり，多くの方々に協力して頂きました．本研究の初期段階から，手厚く，たくさんのご指導ご鞭撻をしてくださった奈良先端科学技術大学院大学の久保孝富准教授，畑秀明助教，幾谷吉晴さんにはとても感謝しています．

指導教員の上野准教授には，研究の知識やアドバイス，発表練習など様々な面でご指導をいただきました．また，山口賢一教授には査読教員として多くの指摘をいただきました．

筆者は実際に体験していない非常に酷な実験タスクであるにも関わらず，大変辛い実験にご協力していただきました被験者の皆様，誠にありがとうございました．

本研究に協力していただいた皆様，ありがとうございました．

参考文献

- [1] T. Dyba, B. A. Kitchenham, and M. Jorgensen, "Evidence-based software engineering for practitioners," *IEEE software*, vol. 22, no. 1, pp. 58–65, 2005.
- [2] T. Fritz, A. Begel, S. C. Müller, S. Yigit-Elliott, and M. Züger, "Using psychophysiological measures to assess task difficulty in software development," in *Proceedings of the 36th international conference on software engineering*. ACM, 2014, pp. 402–413.
- [3] R. J. Jr, "A triarchic model of p300 amplitude," *Psychophysiology*, vol. 23, pp. 367–384, 1986.
- [4] M. Kimura, and Y. Takeda, "Task difficulty affects the predictive process indexed by visual mismatch negativity," *Frontiers in human neuroscience*, vol. 7, no. 267, 2013.
- [5] A. Kok, "Event-related-potential (erp) reflections of mental resources: a review and synthesis," *Biological psychology*, vol. 45, no. 1-3, pp. 19–56, 1997.
- [6] A. F. Kramer, C. D. Wickens, and E. Donchin, "An analysis of the processing requirements of a complex perceptual-motor task," *Human factors*, vol. 25, no. 6, pp. 597–621, 1983.
- [7] M. Kutas, and K. D. Federmeier, "Thirty years and counting: finding meaning in the n400 component of the event-related brain potential (erp)," *Annual review of psychology*, vol. 62, pp. 621–647, 2011.
- [8] M. Kutas, and S. A. Hillyard, "Event-related brain potentials to grammatical errors and semantic anomalies," *Memory & cognition*, vol. 11, no. 5, pp. 539–550, 1983.
- [9] E. F. Lau, C. Phillips, and D. Poeppel, "A cortical network for semantics:(de)

- constructing the n400," *Nature Reviews Neuroscience*, vol. 9, no. 12, p. 920, 2008.
- [10] S. Lee, A. Matteson, D. Hooshyar, S. Kim, J. Jung, G. Nam, and H. Lim, "Comparing programming language comprehension between novice and expert programmers using eeg analysis," in *2016 IEEE 16th International Conference on Bioinformatics and Bioengineering (BIBE)*. IEEE, 2016, pp. 350–355.
- [11] S. Luck, and E. Kappenman, "UC-Davis / SDSU ERP boot camp," <https://erpinfo.org/the-erp-boot-camp>.
- [12] S. C. Müller, and T. Fritz, "Stuck and frustrated or in flow and happy: sensing developers' emotions and progress," in *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*, vol. 1. IEEE, 2015, pp. 688–699.
- [13] J. W. Peirce, "Psychopy-psychophysics software in python," *Journal of neuroscience methods*, vol. 162, no. 1-2, pp. 8–13, 2007.
- [14] T. Picton, S. Bentin, P. Berg, E. Donchin, S. Hillyard, R. Johnson, G. Miller, W. Ritter, D. Ruchkin, M. Rugg *et al.*, "Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria," *Psychophysiology*, vol. 37, no. 2, pp. 127–152, 2000.
- [15] M. A. Pitts, J. L. Nerger, and T. J. R. Davis, "Electrophysiological correlates of perceptual reversals for three different types of multistable images," *Journal of Vision*, vol. 7, no. 6.
- [16] C. Portella, S. Machado, F. Paes, M. Cagy, A. T. Sack, A. Sandoval-Carrillo, J. Salas-Pacheco, A. C. Silva, R. Piedade, P. Ribeiro, A. E. Nardi, and O. Arias-Carrión, "Differences in early and late stages of information processing between slow versus fast participants," *International archives of medicine*, vol. 7, no. 49, 2014.
- [17] D. Redelmeier, and R. Tibshirani, "Association between cellular-telephone calls and motor vehicle collisions," *The New England journal of medicine*, vol. 336, no. 7, pp. 453–458, February 1997. [Online]. Available: <https://doi.org/10.1056/NEJM199702133360701>
- [18] T. Roehm, R. Tiarks, R. Koschke, and W. Maalej, "How do professional developers comprehend software?" in *Proceedings of the 34th International Conference on Software Engineering*. IEEE Press, 2012, pp. 255–265.

- [19] I. Solís-Marcos, and K. Kircher, "Event-related potentials as indices of mental workload while using an in-vehicle information system. cognition," *Cognition, Technology & Work*, vol. 21, no. 1, pp. 55–67, 2019.
- [20] D. L. Strayer, and W. A. Johnston, "Driven to distraction: dual-task studies of simulated driving and conversing on a cellular telephone," *Psychological science*, vol. 12, no. 6, pp. 462–466, 2001.
- [21] R. van Dinteren, M. Arns, M. L. A. Jongsma, and R. P. C. Kessels, "P300 development across the lifespan: A systematic review and meta-analysis," *POLS ONE*, vol. 9, no. 2, 2014.
- [22] X. Xia, L. Bao, D. Lo, Z. Xing, A. E. Hassan, and S. Li, "Measuring program comprehension: A large-scale field study with professionals," *IEEE Transactions on Software Engineering*, vol. 44, no. 10, pp. 951–976, 2017.
- [23] 宮田洋, 藤澤清, 柿木昇治, 山崎勝男, 新生理心理学-生理心理学の基礎- 北大路書房, 1998.
- [24] 宮本靖貴, "視線と脳波の同時計測によるバグ発見時の事象関連電位発生を検証", 奈良工業高等専門学校卒業論文, pp. 5–8, 2016.
- [25] 入野野 宏, "P300 応用 認知科学の立場から", 臨床神経生理学, vol. 41, no. 2, pp. 86–92, 2013.
- [26] 山本愛子, "脳波によるプログラム解法の見当がついた状態の識別", 奈良工業高等専門学校卒業論文, 2016.
- [27] 村上郁也, イラストレクチャー認知神経学 オーム社, 2013.
- [28] 入野野宏, 堀忠雄, "心理学研究における事象関連電位 (ERP) の利用", vol. 26, pp. 15–30, 2000.
- [29] 堀忠雄, 整理心理学-人間の行動を生理指標で測る- 培風館, 2008.
- [30] 満倉靖恵, "脳はウソをつかない-脳波で判るあなたの真実-", 日本耳鼻咽喉科学会会報, vol. 118, no. 4, pp. 461–465, 2015.
- [31] 木村元洋, "事象関連電位のサイズとタイミングの定量化", 基礎心理学研究, vol. 37, no. 1, pp. 99–113, 2018.