

評価的観察者が P300 を指標とした虚偽検出検査に及ぼす影響

平 伸二・三阪梨紗

P300 による虚偽検出中の評価的観察者の効果を 22 名の大学生で検討した。模擬犯罪課題の後、すべての参加者は評価的観察者が一緒か、参加者単独かのいずれかで、P300 による秘匿情報検査を受けた。両群ともに、裁決刺激は非裁決刺激に比較して有意に大きな P300 振幅を生起させた。しかしながら、P300 振幅は両群では有意差が認められなかった。

【キーワード 虚偽検出 P300 評価的観察者】

現在、日本の実務では末梢神経系を指標とした虚偽検出検査が行われている。そして、現在新たな指標として中枢神経系を指標とした虚偽検出検査の研究が行われており、なかでも P300 振幅を指標とした虚偽検出検査の研究は数多く検討されている。例えば、Farwell & Donchin(1991)の研究では、P300 を指標として有罪群と無罪群を設定した 3 刺激オッドボール課題を行っている。参加者は、呈示される標的項目と非標的項目(犯罪関連情報である裁決項目と犯罪無関連情報である非裁決項目で構成されている)を、左右のボタン押しをすることで弁別させた。その結果、P300 振幅は、無罪群では裁決項目と非裁決項目の間に有意差が認められなかったのに対し、有罪群では裁決項目の方が非裁決項目よりも最大振幅が有意に大きくなることが認められた。すなわち、P300 振幅が虚偽検出の有効な指標となることが示唆された。そして、平(2005)がまとめた、有罪条件における 10 の研究の P300 振幅による虚偽検出率の平均値は 88.3%(表 1)であり、P300 振幅での虚偽検出の有効性については、数多くの研究が支持している。

表 1 有罪条件における P300 による虚偽検出の正検出率

研 究	正検出率
Allen and Iacono(1997)	86.7%
Farwell and Donchin(1991)	90.0%
Farwell and Smith(2001)	100.0%
Johnson and Rosenfeld(1992)	76.5%
三宅ら(1986)	87.5%
音成ら(1991)	100.0%
Rosenfeld et al.(1987)	90.0%
Rosenfeld et al.(1988)	100.0%
Rosenfeld et al.(1991)	92.3%
佐々木ら(2001)	87.9%
10研究の平均	88.3%

P300による虚偽検出は、シールドルーム内に1人で座った被検査者に対し、視覚刺激をディスプレイで視覚呈示し、脳波の加算平均処理はシールドルーム外のコンピュータで行う。一方、実務の虚偽検出検査は、検査者が被検査者とともに同じ部屋に入り、検査前の面接、質問呈示、生理指標の計測、視察判定などの一連の検査手続きを行っている(小林・吉本・藤原, 2009)。つまり、検査場面に検査者という評価者が同席している状況下にある。

このような評価的観察者が、生理反応に影響を与えるという知見がある。たとえば、長野(2005)の実験では、アドミッタンス法による心拍出量の計測を行い、評価が血行力学的反応に与える影響をより直接的に検討することと、現実場面で想定される多様な評価状況を考慮し、複数の評価者を用い、評価的観察の影響が課題の特質により変化するかを検討した。また、心臓血管系への影響が異なることが知られている2種類の課題、すなわち暗算(参加者は、モニター上に表示された問題を解き、マウスで答えを選択した)と鏡映描写(参加者は、モニター上の図形をマウスでなぞった)を用いることにより、評価的観察の影響が課題の特性により変化するかも検討した。

その結果、収縮期血圧、拡張期血圧は両課題において上昇したが、課題による差は認められなかった。両課題がもたらす血圧上昇の平均値がほぼ同等であることから、両課題が心臓血管系へ与えた負荷の大きさは同程度であったと考えられる。しかし、暗算課題においては、血圧の上昇は心拍出量増大によるものであるが、鏡映描写においては主として全末梢抵抗の増大により血圧が上昇したことが示された。最も心拍出量を上昇させ、全末梢抵抗の平均変化量が唯一マイナス方向となったのは、暗算課題－評価あり条件であり、暗算課題において観察される、心拍出量増大と全末梢抵抗減少を特徴とした反応は、暗算課題そのものではなく、何らかの対人的刺激を契機として生じている可能性が示唆された。これらのことから、評価的観察者の存在が暗算や鏡映描写課題中の心拍と血圧を増大させる可能性を示唆している。

小川・敦賀・小林・松田・廣田・鈴木(2007)の実験は、実務において予備検査として行われている、カード型検査の **concealed information test** (以下、CITと略す)の手続きを課題として用い、検査場面における評価的観察者の有無が、裁決情報と非裁決情報に対して生じる生理反応の相対的差異に及ぼす影響を検討した。生理指標として皮膚電気活動(**electrodermal activity**: 以下、EDAと略す)に関するものとして **skin conductance response**(以下、SCRと略す)と **skin conductance level**(以下、SCLと略す)、心臓血管系活動を反映するものとして、**heart rate**(以下、HRと略す)、**blood pressure**(以下、BPと略す)、規準化

脈波容積(normalized pulse volume : 以下, NPV と略す), respiration rate(以下, RR と略す)を測定した。また, 心理的覚醒感を測定するために general arousal checklist(以下, GACL と略す)短縮版を用いている。

実験は個別に行われ, 参加者は安静期間を 5 分間測定後, GACL および自意識尺度に回答した。その後, 観察群では実験参加者と同姓で実験前の面識の無い観察者が 1 名, 白衣を着用して入室し, 実験参加者の右斜め前方 1.5m の距離に設置された椅子に座った。そして, カード検査型 CIT を繰り返し実施した。この間, 観察者あり群では, 観察者が CIT と CIT 間でメモをとるなどして, 実験参加者の行動を観察・評価する素振りを行った。なお, 観察者の存在と役割については, 教示において特に言及しなかった。

その結果, CIT 中の裁決情報に対して見られた皮膚電気活動の亢進, 呼吸運動の抑制, 心臓血管系における末梢血管の収縮(NPV の低下)と BP の上昇, それに続く HR の低下という一連の変化は, 長野(2005)の結果と一致した。他者の同席という操作により, 言語報告では主として全般的脱活性(“静かな”, “のどかな”などの項目で構成)の低下がみられ, 生理指標では SCL, SCR 振幅の増大, NPV の低下など, 主として皮膚交感神経系の活動を反映する指標において, 覚醒方向の変化が見られた。これは, 覚醒水準の実験的操作が有効であったことを示唆している。総合的にみると, 少なくとも, 覚醒水準の違いは, CIT 中の裁決・非裁決に対する生理反応の弁別的応答にそれほど影響を及ぼさなかった。これらの結果を実務的観点から見れば, 標準化という処理が覚醒水準の影響を弱め, 裁決と非裁決の差異を明確にする上で有効である可能性を示唆するものと考えられる。

さらに, 大杉・大平(2006)の実験では, 模擬犯罪課題として覚醒度の異なる 2 種類の行為を課し, その後 CIT において, それらの 2 種類の行為に関わる裁決刺激を同じ系列内に配置した。覚醒の操作の妥当性を証明するため, 実験を通して末梢系指標である HR および EDA を測定することで, 中枢(P300)と末梢の両側面から CIT メカニズムを検討している。参加者は模擬犯罪課題実施前に, 無罪の状態での CIT を受けた。その後, 2 種類の模擬犯罪課題が参加者間でランダムに実施された。参加者は刃物名が書かれた 5 枚の封筒の中から 1 つ選び, 指定された刃物でかごの中に入った画像を 5 分間刺し続けた。高覚醒課題の画像は赤ちゃん人形の写真, 低覚醒課題の画像は果物の写真であった。最後に再び CIT を実施した。それぞれの課題間には 5 分間の安静期間を設け, さらに, 各 CIT 終了後, 6 種類の刺激(はさみ, 包丁, カッター, 鎌, アイスピック, のこぎり)に対し, 画像評定が求められ, 感情価と覚

醒度について9段階で評定を行った。

その結果、画像評定については、課題後、高裁決刺激と低裁決刺激に対する覚醒が非裁決刺激より有意に高かった。また、高裁決刺激のみ、課題前と課題後の評定に有意差がみられた。HR および SCL については、課題開始1分間で有意に HR が上昇し、上昇の程度は低覚醒課題より高覚醒課題で大きいことがわかった。Pz において P300 の頂点振幅が、高裁決刺激が有意に低裁決刺激より大きく、またそれらいずれも非裁決刺激より大きいという結果が示された。これは、低覚醒より高覚醒であるほうが虚偽検出に好影響を及ぼすことを示唆している。

本研究は、評価的観察者の影響に関する先行研究において使用された生理指標が、SCR と SCL, HR, BP, NPV, そして RR の末梢系であったことから、中枢系の指標である P300 を用いて評価的観察者の影響を検討する。模擬犯罪課題に関しては、大杉・大平(2006)の手続きを用いる。そして、評価的観察者の影響を明らかにするとともに、裁決刺激に対する P300 振幅が、観察者あり群でも観察者なし群でも非裁決刺激に比較して有意に増大するかを検討することを目的とする。

方法

参加者 実験同意書に署名した福山大学の男子学生 12 名 ($M=21.3$ 歳, $SD=1.20$ 歳), 女子学生 10 名 ($M=20.9$ 歳, $SD=0.99$ 歳), 合計 22 名 ($M=21.1$ 歳, $SD=1.14$ 歳)が本研究に参加した。この 22 名を無作為に観察者あり群と観察者なし群に振り分けて実施した。

装置と指標 脳波と眼球電位の測定には、TEAC 製携帯型多用途生体アンプ (Polymate AP1524)を用いた。脳波は時定数 3s, サンプリング周波数 500Hz で、両耳朵を基準部位として Fz, Cz, Pz から測定した。眼球電位は、左眼眼窩上下縁部から導出し、加算平均の際のアーティファクトチェックとして用いた。模擬犯罪課題で刃物を選択するときを使用したカード、6 種類の刃物(アイスピック, カッター, 鎌, のこぎり, はさみ, 包丁), 模擬犯罪課題で使用した赤ちゃん画像, カッター板を用いた。

刺激 6 種類の刃物の画像を用い、標的的刺激として刃物の画像(はさみ), 裁決刺激として参加者が模擬犯罪課題で使用した刃物の画像(包丁), 非裁決刺激としてそれ以外の刃物の画像(アイスピック, カッター, 鎌, のこぎり)とした。画像の呈示時間は 300ms とし、呈示間隔は 1500ms±300ms とした。1 セッションの呈

示の回数は標的刺激 20 回，裁決刺激 20 回，非裁決刺激 80 回(20×4 種)の計 120 回刺激を無作為な順序で呈示した。参加者は休憩を挟んで 2 セッションの実験を行った。

手続き 参加者に実験同意書署名時，実験の目的と方法(脳波の測定，模擬犯罪課題の方法，虚偽検出検査で嘘をつく必要)に加え，実験中いつでも自由意志で実験を降りることができるとの説明を加えた。まず，参加者には別室にて模擬犯罪課題を行った。刃物が描かれた 6 枚の封筒の中から 1 つ選び，指定された刃物で赤ちゃん人形の写真の画像を 2 分間刺し続けるよう指示した。模擬犯罪課題終了後，ボタン押し課題を用いて脳波による測定を実施した。ボタン押し課題では，標的刺激に対しては利き手に持つボタン，そのほかの刺激に対しては非利き手に持つボタンにより，ディスプレイに刺激が呈示されたらできるだけ速く正確に押すように求めた。また，観察者あり群においては，観察者(女性)は，白衣を着用し，参加者(男性，女性の両方)の右斜め前 1m の位置にある椅子に座り，虚偽検出検査が始まると，観察者は評価用紙に評価を行っているような動作を行った。参加者に対して評価に関する教示は行っていない。

結果

観察者あり群(左)，観察者なし群(右)の標的刺激，裁決刺激，非裁決刺激に対する総加算平均波形(Pz)を示す(図 1)。

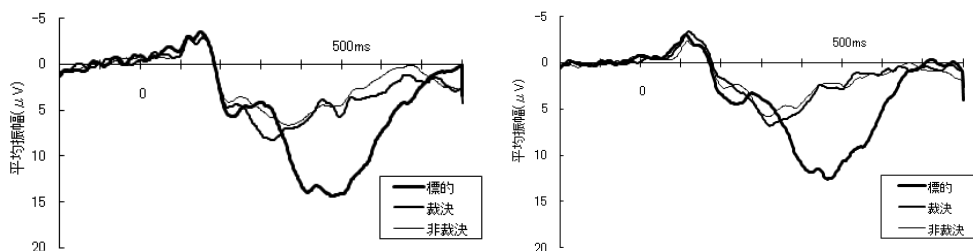


図 1. 観察者あり群(左)，観察者なし群(右)における各刺激に対する総加算平均波形(Pz)

最大振幅は観察者あり群の標的刺激 $15.0\mu\text{V}(SD=4.2)$ ，裁決刺激は $8.3\mu\text{V}(SD=3.4)$ ，非裁決刺激は $6.7\mu\text{V}(SD=3.3)$ ，観察者なし群の標的刺激は $12.6\mu\text{V}(SD=5.1)$ ，裁決刺激は $2.8\mu\text{V}(SD=3.7)$ ，非裁決刺激は $5.8\mu\text{V}(SD=3.7)$ と両群ともに，刺激呈示後 300ms-500ms 間で最大振幅が大きくなった。

次に，Pz における観察者あり群と観察者なし群の P300 の最大振幅の平均値を

示す(図 2)。

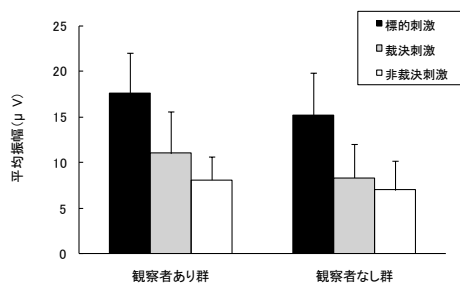


図 2. 観察者あり群と観察者なし群における各参加者の P300 の最大振幅の平均値(Pz)

P300 振幅は標的刺激、裁決刺激、非裁決刺激の順に大きくなった。観察者あり群の標的刺激は $17.6\mu\text{V}$ ($SD=4.5$), 裁決刺激は $11.0\mu\text{V}$ ($SD=4.6$), 非裁決刺激は $8.1\mu\text{V}$ ($SD=2.6$)であった。観察者なし群の標的刺激は $15.2\mu\text{V}$ ($SD=4.6$), 裁決刺激は $8.3\mu\text{V}$ ($SD=3.7$), 非裁決刺激は $7.0\mu\text{V}$ ($SD=3.2$)であった。群(観察者あり群・観察者なし群)×刺激(標的刺激・裁決刺激・非裁決刺激)の 2 要因分散分析の結果、刺激の主効果にのみ有意差が認められた ($F(2,40)= 51.826, p<.001, \epsilon=.776$)。Bonferroni 法によって各刺激の多重比較を行った結果、すべての刺激間に有意差が認められた ($p<.05$)。

個別判定の結果、観察者あり群において非裁決刺激よりも裁決刺激の平均振幅が大きかった参加者は 11 人中 9 人(81.8%)であった。観察者なし群においては、非裁決刺激よりも裁決刺激の平均振幅が大きかった参加者は 11 人中 8 人(72.7%)であった。また、両群合わせた検出率は 77.3%であった。 χ^2 検定を行った結果、観察者あり群と観察者なし群の検出率に有意な差が認められなかった。

次に、Pz における観察者あり群と観察者なし群の P300 の頂点潜時の平均値を示す(図 3)。

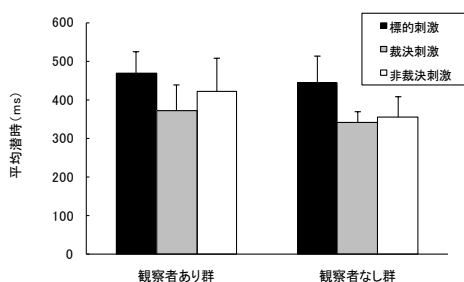


図 3. 観察者あり群と観察者なし群における各参加者の P300 の頂点潜時の平均値(Pz)

P300 潜時は、標的刺激、非裁決刺激、裁決刺激の順に遅くなった。観察者あり群の標的刺激は 469.1ms($SD=54.9$), 裁決刺激は 371.6ms($SD=67.9$), 非裁決刺激は 422.4ms($SD=84.7$)であった。観察者なし群の標的刺激は 443.5ms($SD=70.3$), 裁決刺激は 342.4ms($SD=25.8$), 非裁決刺激は 356.0ms($SD=52.0$)であった。群(観察者あり群・観察者なし群)×刺激(標的刺激・裁決刺激・非裁決刺激)の 2 要因分散分析の結果、刺激の主効果に有意差が認められ($F(2,40)=18.060, p<.001$), 群の主効果に有意な傾向が認められた($F(1,20)=4.175, p<.10$)。Bonferroni 法によって各刺激の多重比較を行った結果、標的刺激と裁決刺激、標的刺激と非裁決刺激の間に有意差が認められた($p<.05$)。

考察

本実験では、中枢指標である P300 を用い、評価的観察者の及ぼす効果、さらには、観察者あり群でも観察者なし群でも裁決刺激に対する P300 振幅が、非裁決刺激に比較して有意に増大するかを検討することを目的としていた。

まず、Pz における観察者あり群と観察者なし群の P300 振幅の最大値の平均値は、観察者あり群の標的刺激、裁決刺激、非裁決刺激の方が、観察者なし群の標的刺激、裁決刺激、非裁決刺激より最大振幅の平均値が大きかった。しかしながら、分散分析の結果、群間に有意差は認められなかった。つまり、観察者の存在が P300 による虚偽検出検査に影響するという結果は得られなかった。一方、裁決刺激に対する P300 振幅は、観察者あり群でも観察者なし群でも非裁決刺激に比較して増大していた。多重比較の結果も標的刺激と裁決刺激、標的刺激と非裁決刺激、裁決刺激と非裁決刺激の間に有意差が認められた。すなわち、裁決刺激に対する P300 振幅が、非裁決刺激と比較して有意に増大することが認められ、虚偽検出の可能性が示唆された。

個別判定を行った結果、本実験の検出率は観察者あり群が 81.8%, 観察者なし群が 72.7%であった。両群合わせると 77.3%であり、Ben-Shakhar & Furedy(1990)の 83.9%と平(2005)の 88.3%に比較して若干低い検出率となった。Ben-Shakhar & Furedy(1990)と平(2005)よりも検出率が低かった原因として、本研究で呈示した刺激の非類似性が指摘できる。足立・山岡(1985)の先行研究によれば、各項目間の非類似性が小さいほど成績は低下し、逆に、非類似性が大きいほど成績は向上するとしている。本実験では、実験の内省報告からも、赤いはさみに集中しようとしたが、赤いカッターもあり赤色に反応してしまったと回答した参加者が多数いた。このことから、標的刺激の赤いはさみと非裁決刺激の赤いカッター

一の識別性の低下，すなわち，刺激の非類似性が小さくなってしまい，検出率が若干低下したと考えられる。

しかし，両群の検出率に有意な差が認められなかったことから，評価的観察者の存在する状況下においても刺激の弁別が可能であり，検査者の正判定を回避するためにカウンターメジャーが用いられる実務においても，P300 振幅による虚偽検出の有効性が期待できる。平(2009)は，脳機能研究の中では，現時点で実務導入に最も可能性が高い指標は P300 であると指摘している。本実験もそれを支持する結果であった。

P300 潜時では，標的刺激，非裁決刺激，裁決刺激の順に遅くなった。群の主効果に有意な傾向がみられたが，裁決刺激と非裁決刺激の間に有意な差は認められなかった。観察者の存在により潜時が遅くなる傾向はみられたものの，両群の裁決刺激と非裁決刺激に有意な差は認められず，P300 潜時が P300 振幅のように虚偽検出の有効な指標になることが期待できるといった結果は得られなかった。

今回の研究では，小川(2007)で用いられた畑山・Antonides・松岡・丸山(1994)の GACL を使用して参加者の心理的覚醒水準を測定していなかったため，参加者の虚偽検出検査時の覚醒水準が推定できない。加えて，標的刺激と非裁決刺激の非類似性が小さくなってしまったことなどの問題点があげられる。また，観察者あり群，観察者なし群ともに，有罪群だったため，無罪群との比較ができなかった。今後は，本研究であげられた問題点を改善し，有罪群と無罪群を設定して評価的観察者が虚偽検出検査に与える影響についての検討が望まれる。

引用文献

- 足立浩平・山岡一信 (1985). 質問項目間の非類似性が情報の再認および虚偽検出に及ぼす効果 科学警察研究所報告法科学編, 38, 126-131.
- Ben-Shakhar, G., & Furedy, J.J. (1990). *Theories and applications in the detection of deception: A psychophysiological and international perspective*. New York: Springer-Verlag.
- Farwell, L.A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event-related brain potentials. *Psychophysiology*, 28, 531-547.
- 畑山俊輝・Antonides, G.・松岡和生・丸山欣哉 (1994). アラウザルチェックリスト(GACL)から見た顔のマッサージの心理的緊張低減効果 応用心理学研

- 究, 19, 11-19.
- 平 伸二 (2005). 虚偽検出に対する心理学の貢献と課題 心理学評論, 48, 384-399.
- 平 伸二 (2009). 脳機能研究による concealed information test の動向 生理心理学と精神生理学, 27, 57-70.
- 小林孝寛・吉本かおり・藤原修治 (2009). 実務ポリグラフ検査の現状 生理心理学と精神生理学, 27, 5-15.
- 長野祐一郎 (2005). 評価的観察が精神課題遂行中の心臓血管反応に与える影響 心理学研究, 76, 252-259.
- 小川時洋・敦賀麻理子・小林孝寛・松田いづみ・廣田昭久・鈴木直人 (2007). 覚醒水準が隠匿情報検査時の生理反応に与える影響 心理学研究, 78, 407-415.
- 大杉朱美・大平英樹 (2006). 行動時の感情喚起が CIT に及ぼす影響—ERP および末梢指標の同時計測によるアプローチ— 生理心理学と精神心理学, 26, 194.

The effect of evaluative observer's existence during detection of deception with P300

Shinji Hira and Risa Misaka

The effect of evaluative observer's existence during detection of deception with P300 was tested in 22 undergraduate students. After a mock crime task, all participants were examined with the P300-based concealed information test, either with or without an evaluative observer. In both groups, the critical stimuli elicited significantly larger P300 amplitudes than the non-critical stimuli. However, mean P300 amplitude did not differ between two groups.

【Key words: detection of deception, P300, evaluative observer】