

P300によるGKTの裁決項目と非裁決項目のP300振幅・潜時と 反応時間の比較

平 伸二・三阪梨紗・濱本有希

要約

本研究は、P300振幅、P300潜時と反応時間について裁決刺激と非裁決刺激の違いを比較することを計画した。分析は、173名の有罪参加者による、5つの異なる研究データに基づいている。結果として、裁決刺激の方が非裁決刺激より有意に大きなP300振幅を生起していた。また、裁決刺激の方が非裁決刺激より有意に遅いP300潜時となった。このことは、従属変数としてP300を現場で応用できることを示唆する。しかし、反応時間では裁決刺激と非裁決刺激の間に有意差は認められなかった。

[キーワード：有罪知識検査、P300潜時、反応時間]

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%であった。このように、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。

しかし、高い検出力があることを示している一方で、1割程度の人は正しく判定されず、冤罪や犯罪を見逃してしまうfalse negativeが起こる可能性が考えられる(藤原・小林・古満, 2007)。このような問題点を改善するため、3刺激オッドボール課題で同時に収集できるP300潜時や反応時間(reaction time: RT)の有効性について検討する必要がある。

ユング(Jung, C. G.)は、被疑者に対し、犯罪に関係ある刺激語と、犯罪に関係ない刺激語を巧妙に整列させて呈示し、それに対する反応語の内容、反応語を発するまでのRT、反応状況を検査した。その結果、被疑者が犯罪に関する刺激語に対して反応語を発する

までに、精神的に葛藤を起こしたならば、RTが遅れたり、逆に極端に早まったり、刺激語や反応語を繰り返したり、刺激語と関係ない反応語をあげたり、反応語を発しなかったりといった状況が見られるとした。そして、この連想診断法によって被疑者がウソをついているか否かを判断できるとした。また、この連想診断法を、生理的諸反応の測定と組合せ、ウソ発見の手法として犯罪捜査に適用することを示唆したのが、ミュンスタールベルグ(Münsterberg, H.)であった(倉持, 2000)。

前述のFarwell & Donchin(1991)の研究は、無罪群では裁決項目と非裁決項目のRTに有意差が認められなかった。また、有罪群では裁決項目に対するRTが、非裁決項目に対するRTよりも有意に遅くなることが認められた(統計検定は著者らによる分析)。そして、平・古満(2006)の研究では、模擬窃盗課題を用いて犯罪場面群と大学風景群を設けP300を指標とした3刺激オッドボール課題を行った。その結果、RTは標的刺激、裁決刺激、非裁決刺激の順に遅くなり、刺激間にのみ有意差が認められた。両群とも非裁決刺激よりも裁決刺激で遅くなった。Seymour, Seifert, Shafto, & Mosmann(2000)が、Farwell & Donchin(1991)と同様の方法を用いて行ったRTの研究によれば、RTは裁決項目の方が非裁決項目よりも有意に遅くなることが認められた。

一方、Verschuere, Crombez, & Koster(2004)の研究はProbe分類課題を用いて3種類の実験を行っている。その結果、裁決項目後に呈示されたドットに対するRTと非裁決項目後に呈示されたドットに対するRTの間に有意差は認められなかった。また、藤原他(2007)の研究では、模擬窃盗を行わない無罪群と模擬窃盗を行う有罪群を設けdot-probe探査課題を用いてボタン押しまでのRTを測定した。その結果、財布画像を刺激として呈示した場合のみ、有罪群において無罪群よりも裁決項目のRTが非裁決項目のRTよりも有意に速くなった。

久保・宮谷・入戸野(2007)は、模擬窃盗群と無罪群を設けてP300を指標とした3刺激オッドボール課題を行い、P300振幅、P300潜時、RTの3指標について分析している。その結果、P300振幅では模擬窃盗群でのみ、裁決項目の方が非裁決項目よりも有意に高振幅であった。P300潜時では、模擬窃盗群の方が無罪群よりも潜時が有意に遅くなっており、標的項目と裁決項目に比べて非裁決項目で有意にP300潜時が速くなっていた。RTでは、模擬窃盗群の方が無罪群よりも有意に遅く、標的項目、裁決項目、非裁決項目の順に遅くなっていた。標的項目に対するRTが、裁決項目と非裁決項目に対するRTよりも有意に遅くなった。

以上のように、P300振幅が虚偽検出の有効な指標となることはすべての研究で一致しているが、P300潜時については久保他(2007)の研究しか検討されておらず、RTに関しては研究間の一致が認められていない。

そこで、著者3名が過去に実施したP300を指標とした3刺激オッドボール課題を用いた5

P 300 による G K T の裁決項目と非裁決項目の P 300 振幅・潜時と反応時間の比較

つの虚偽検出実験のデータを用い($n=173$), P300 振幅, P300 潜時, RT の有効性について検討する。具体的には, 標的刺激, 裁決刺激, 非裁決刺激に対する P300 振幅, P300 潜時, RT の平均値ならびに 173 名のデータ分布を求め, それぞれの指標の有効性について検討する。

方法

分析データ 福山大学で過去に実施した, P300 を指標とした 3 刺激オッドボール課題を用いた, 5 つの虚偽検出実験に参加した 173 名の P300 振幅, P300 潜時, RT を分析対象とした。5 つの虚偽検出実験の参加者, 刺激, 手続きについては付表 1 に示す。

各実験ともに, 実験同意書署名時, 実験の目的と方法(脳波の測定, 模擬窃盗などの記憶課題の方法, 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。刺激は 3 刺激オッドボール課題とした。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。

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, 両耳朶を基準部位として 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。RT は, レベルトリガ変換ユニット付属のソフトで計測した。

結果

173 名の P300 振幅(Pz) の最大値(刺激呈示後 300ms-600ms 間の最大値)を標的刺激, 裁決刺激, 非裁決刺激別に示す(図 1)。

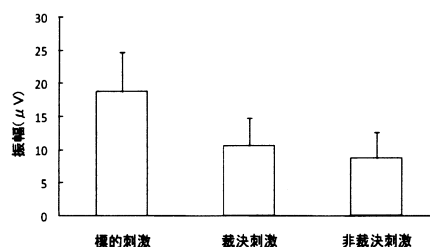


図 1. 173 名の標的刺激, 裁決刺激, 非裁決刺激に対する P300 の平均振幅(Pz)

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。標的刺激は $19.1 \mu V$

($SD=6.0$)、裁決刺激は $10.4 \mu V$ ($SD=4.2$)、非裁決刺激は $8.6 \mu V$ ($SD=4.0$)であり、1 要因分散分析の結果、有意差が認められた ($F(2, 518)=551.56, p<.001$)。ライアン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 ($p<.05$)。

さらに、173 名のデータの分布を比較するため、図 2 に標的刺激、図 3 に裁決刺激、図 4 に非裁決刺激のヒストグラムを示した。

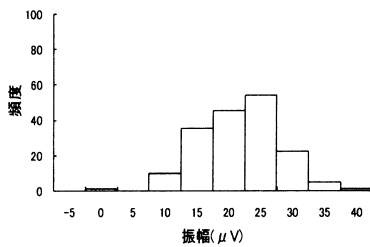


図 2. 173 名の標的刺激に対する P300 の平均振幅 (P_z) のヒストグラム

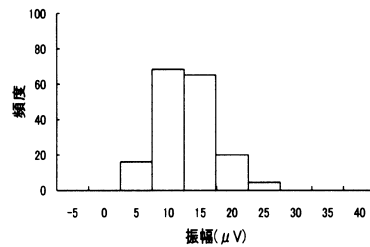


図 3. 173 名の裁決刺激に対する P300 の平均振幅 (P_z) のヒストグラム

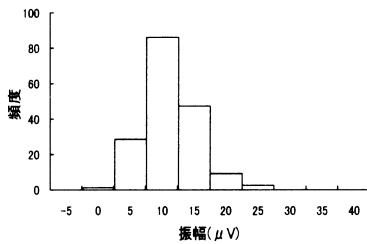


図 4. 173 名の非裁決刺激に対する P300 の平均振幅 (P_z) のヒストグラム

ヒストグラムの形状を見ると、標的刺激は $20 \mu V$ - $25 \mu V$ の区間に最大振幅の頻度が多くなっており、尖度は 0.03、歪度は -0.06 であった。裁決刺激は $5 \mu V$ - $15 \mu V$ の区間に最大振幅の頻度が多くなっており、尖度は 0.53、歪度は 0.46 であった。これに対し、非裁決刺激は $5 \mu V$ - $10 \mu V$ の区間に頻度が多くなっており、尖度は 0.46、歪度は 0.54 であった。

173 名の P300 の頂点潜時 (P_z) を標的刺激、裁決刺激、非裁決刺激別に示す (図 5)。

P 300 による G K T の裁決項目と非裁決項目の P 300 振幅・潜時と反応時間の比較

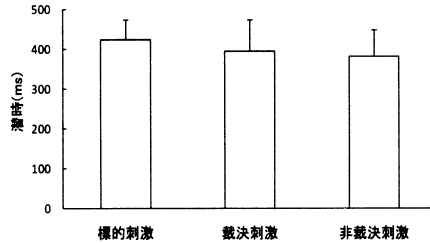


図 5. 173 名の標的刺激, 裁決刺激, 非裁決刺激に対する P300 の平均潜時(Pz)

P300 潜時は, 標的刺激, 裁決刺激, 非裁決刺激の順に遅くなった。標的刺激は 422.5ms ($SD=49.7$), 裁決刺激は 395.7ms ($SD=77.3$), 非裁決刺激は 382.9ms ($SD=65.4$) であった。1 要因分散分析の結果, 有意差が認められた ($F(2, 518)=25.22, p<.001$)。ライアン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 ($p<.05$)。

さらに, 173 名のデータの分布を比較するため, 図 6 に標的刺激, 図 7 に裁決刺激, 図 8 に非裁決刺激のヒストグラムを示した。

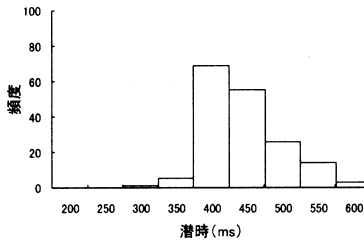


図 6. 173 名の標的刺激に対する P300 の平均潜時(Pz)のヒストグラム

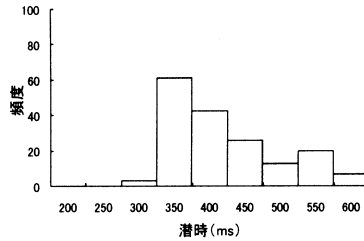


図 7. 173 名の裁決刺激に対する P300 の平均潜時(Pz)のヒストグラム

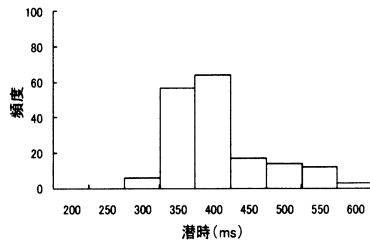


図 8. 173 名の非裁決刺激に対する P300 の平均潜時(Pz)のヒストグラム

ヒストグラムの形状を見ると、標的刺激は 350ms-450ms の区間に頂点潜時の頻度が多くなっており、尖度は 0.27, 歪度は 0.71 であった。裁決刺激は 300ms-400ms の区間に頂点潜時の頻度が多くなっており、尖度は-0.50, 歪度は 0.76 であった。非裁決刺激は 300ms-400ms の区間に頂点潜時の頻度が多くなっており、尖度は 0.62, 歪度は 1.10 であった。標的刺激と裁決刺激の分布が 600ms に向けて漸減しているのに対し、非裁決刺激の分布が 400ms 以降急激に減少しているのが特徴である。

次に、173 名の RT を標的刺激, 裁決刺激, 非裁決刺激別に示す(図 9)。

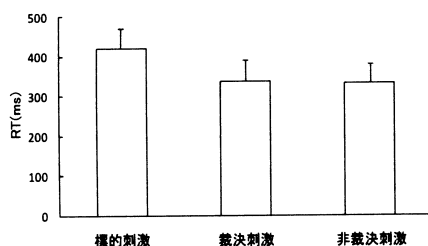


図 9. 173 名の標的刺激, 裁決刺激, 非裁決刺激に対する RT

RT は、標的刺激, 裁決刺激, 非裁決刺激の順に遅くなった。標的刺激は 420.8ms ($SD=49.4$), 裁決刺激は 337.1ms ($SD=51.9$), 非裁決刺激は 332.0ms ($SD=47.0$)であり, 1 要因分散分析の結果, 有意差が認められた ($F(2, 518)=822.52, p<.001$)。ライアン法によって各刺激の多重比較を行った結果, 標的刺激—裁決刺激, 標的刺激—非裁決刺激に有意差が認められた ($p<.05$)。

さらに、173 名のデータの分布を比較するため、図 10 に標的刺激, 図 11 に裁決刺激, 図 12 に非裁決刺激のヒストグラムを示した。

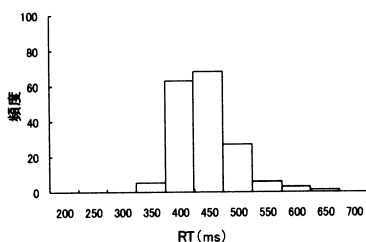


図 10. 173 名の標的刺激に対する RT のヒストグラム

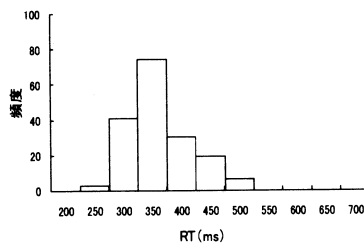


図 11. 173 名の裁決刺激に対する RT のヒストグラム

P 300 による G K T の裁決項目と非裁決項目の P 300 振幅・潜時と反応時間の比較

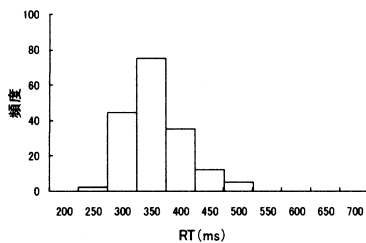


図 12. 173 名の非裁決刺激に対する
RT のヒストグラム

ヒストグラムの形状を見ると、標的刺激は 350ms-450ms の区間に頻度が多くなっており、尖度は 1.81、歪度は 1.07 であった。裁決刺激は 300ms-350ms 区間に最大の頻度が認められ、尖度は 0.32、歪度は 0.75 であった。非裁決刺激も 300ms-350ms 区間に最大の頻度が認められ、尖度は 0.32、歪度は 0.69 であった。つまり、裁決刺激と非裁決刺激の平均値、尖度、歪度はほとんど同じ値を示した。

各指標間の平均値についてピアソンの積率相関係数(r)を算出したところ、P300 潜時と RT の相関は標的刺激では.47、裁決刺激では.34、非裁決刺激では.30 であった($p < .05$)。P300 振幅と潜時の相関は標的刺激では-.22、裁決刺激では.10、非裁決刺激では-.14 であった($p < .05$)。P300 振幅と RT の相関は標的刺激では-.19、裁決刺激では.02、非裁決刺激では-.10 であった($p < .05$)。

参加者毎に裁決刺激の P300 振幅が、非裁決刺激よりも大きい場合を検出成功とした場合、173 名中 131 名 (75.7%) が正しく検出された。P300 潜時では、裁決刺激の方が非裁決刺激よりも P300 潜時が遅い場合を検出成功とした場合、173 名中 87 名 (50.3%) が正しく検出された。RT では、裁決刺激の方が非裁決刺激よりも遅い場合を検出成功とした場合、173 名中 106 名 (61.3%) が正しく検出された。

考察

本研究では、P300 を指標とした 3 刺激オッドボール課題を用いた 5 つの虚偽検出実験から、P300 振幅、P300 潜時、RT の虚偽検出指標としての有効性について検討を行った。

まず、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かを検討したところ、 P_z における P300 振幅の最大値の平均値は、標的刺激、裁決刺激、非裁決刺激の順に大きくなった。さらに、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。この結果により、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 P300 振幅での検出率は 75.7%であり、Ben-Shakhar & Furedy (1990) の 83.9%と平(2005)の 88.3%と比較したところ若干低い検出率となったが、自然確率を上回る高い検出率が得られた。Ben-Shakhar & Furedy (1990) と平(2005)よりも検出率が低かった原因として、先行研究のほとんどは、記憶課題と検査を同一日に実施していたが、本研究で用いた 173 名のうち 104 名(60.1%)は、1 ヶ月以上経過後に脳波測定を行っており、そのことが検出率の低下につながったと考えられる。

次に、裁決刺激に対する P300 潜時が、非裁決刺激と比較して有意に遅くなるかを検討した。Pz における P300 潜時は標的刺激、裁決刺激、非裁決刺激の順に遅くなった。また、すべての刺激間に有意差が認められた。これは、久保他(2007)の研究を支持する結果となったが、検出率が 50.3%と P300 振幅よりも低い検出率を示している。裁決刺激と非裁決刺激の分布(図 7, 図 8)を比較してみると、裁決刺激では分布が 600ms に向けて漸減しているのに対し、非裁決刺激では分布が 400ms 以降急激に減少している。この裁決刺激に対する潜時の遅延は、LPP(late positive potential)が生じた参加者が含まれている可能性を示唆する。Schupp, Cuthbert, Bradley, Cacioppo, Ito, & Lang(2000)は、中性刺激と比較して不快な刺激により大きな LPP が生起することを報告している。本研究では中性刺激は非裁決刺激に、不快な刺激は裁決刺激に相当すると考えられる。その反面、本研究のデータはカード検査と模擬窃盗課題による記憶再認課題で取得しているため、裁決刺激を不快な刺激として認知せず、まれで有意な刺激としてのみ処理した速い潜時の参加者も含まれ、潜時の検出率が 50.3%と低くなったと考えられる。しかし、犯罪捜査における検査では、裁決刺激に不快な情動が付加される(犯行時の情動体験の再認など)ことが予測できるため、実務の検査では LPP を指標とした判定を考慮すべきであろう。

次に、裁決刺激に対する RT が、非裁決刺激と比較して有意に遅くなるかを検討した。各参加者のオッドボール課題の RT の平均値は、標的刺激、裁決刺激、非裁決刺激の順に遅くなった。また、裁決刺激と非裁決刺激の間に有意差は認められなかった。これは Verschuere et al. (2004)の研究を支持する結果となった。個別判定の結果も、RT での検出率は 61.3%と低く、RT での虚偽検出は難しいと考えられる。特に、裁決刺激と非裁決刺激のデータ分布を比較すると(図 11, 12)、分布の形状がほぼ同一で、平均値、尖度、歪度も近似していることがわかった。つまり、RT という行動レベルでは、3 刺激オッドボール課題における参加者の方略は、3 刺激に異なった反応をするのではなく、単に標的刺激かそれ以外の刺激かを弁別して反応しているにすぎない。したがって、3 刺激オッドボール課題の裁決刺激と非裁決刺激に対する RT は、犯罪知識の有無に関して弁別力を持たないと予測される。さらに、RT はカウンタメジャーによって簡単に操作することも可能なため、虚偽検出の有効な指標として使用することは推奨できないと結論できる。

しかしながら、佐々木・平・松田(2001)が、心理的カウンタメジャーを用いた場合、RT が 100ms 遅延し、SD も約 2 倍大きくなることを報告しているように、RT はカウンタメジャーで妨害工作を行っている被検者を検出する指標として期待できる。

【注】本論文は平成 20 年度科学研究費補助金（基盤研究(B)）（課題番号 18330143）の研究
成果の一部である。

引用文献

- 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.
- 藤原裕弥・小林一生・古満伊里 (2007). 反応時間を指標とした虚偽検出の試み 総合人間科学, **7**, 39-47.
- 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・古満伊里 (2006). 犯罪場面の事前呈示は P300 による虚偽検出の正確性に影響するか? 福山大学人間文化学部紀要, **6**, 13-20.
- 久保賢太・入戸野 宏・宮谷真人 (2007). 有罪知識質問法における P300 振幅の規定因 生理心理学と精神生理学, **25**, 267-275.
- 倉持 隆 (2000). 科学捜査におけるウソ発見の始まり 平 伸二・中山 誠・桐生正幸・足立 浩平(編著) ウソ発見—犯人と記憶のかけらを探して— 北大路書房 p45.
- 佐々木 実・平 伸二・松田 俊 (2001). 事象関連電位を用いた虚偽検出における心理的カウンタメジャーの効果 心理学研究, **72**, 322-328.
- Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing : The late positive potential is modulated by motivational relevance. *Psychophysiology*, **37**, 257-261.
- Seymour, T. L., Seifert, C. M., Shafto, M. G., & Mosmann, A. L. (2000). Using response time measures to assess “guilty knowledge.” *Journal of Applied Psychology*, **85**, 30-37.
- Verschuere, B., Crombez, G., & Koster, E. H. W. (2004). Orienting to guilty knowledge. *Cognition and Emotion*, **18**, 265-279.

付表 1 P300 振幅・潜時と RT を指標とした 5 つの虚偽検出実験方法

実験	参加者	刺激	手続き
1	犯罪場面群 8 名, 大学風景群 8 名	矢印画像と 5 つの貴金属画像(呈示時間 300ms, 呈示間隔 2500ms ± 25%)	事前に模擬窃盗課題を実施し, 1 ヶ月及び 1 年経過した時点で脳波測定を行った。犯罪場面群は模擬窃盗課題を実施した現場の映像を呈示した。大学風景群は模擬窃盗課題とは無関係である大学キャンパス内の風景を呈示した。ディスプレイに呈示される画像を注視し, 盗んだ品物が検出されないよう教示したうえでボタン押しを課した。
2	外向性 19 名, 内向性 15 名	矢印画像と 1-5 の数字の画像(呈示時間 300ms, 呈示間隔 1500ms ± 20%)	標準的なカード検査を用い, 脳波測定を行った。ディスプレイに呈示される画像を注視し, 選択したカードが検出されないよう教示したうえでボタン押しを課した。
3	高不安 7 名, 低不安 6 名	6 種類のフルーツ画像(呈示時間 300ms, 呈示間隔 1500ms ± 20%)	標準的なカード検査を用い, 脳波測定を行った。ディスプレイに呈示される画像を注視し, 選択したカードが検出されないよう教示したうえでボタン押しを課した。
4	1 ヶ月後犯罪場面群 24 名, 1 ヶ月後大学風景群 20 名, 1 年後犯罪場面群 22 名, 1 年後大学風景群 22 名	中心条件(6 種類の貴金属画像), 周辺条件(6 種類の文房具画像), (呈示時間 300ms, 呈示間隔 1500ms ± 10%)	事前に模擬窃盗課題を実施し, 1 ヶ月及び 1 年経過した時点で脳波測定を行った。犯罪場面群は模擬窃盗課題を実施した現場の映像を呈示した。大学風景群は模擬窃盗課題とは無関係である大学キャンパス内の風景を呈示した。ディスプレイに呈示される画像を注視し, 盗んだ品物が検出されないよう教示したうえでボタン押しを課した。
5	観察者あり群 11 名, 観察者なし群 11 名	6 種類の刃物画像, (呈示時間 300ms, 呈示間隔 1500ms ± 20%)	事前に模擬犯罪課題を実施し, 脳波測定を行った。ディスプレイに呈示される画像を注視し, 模擬犯罪課題で使用した刃物が検出されないよう教示したうえでボタン押しを課した。

P 300 による G K T の裁決項目と非裁決項目の P 300 振幅・潜時と反応時間の比較

Comparison of critical and non-critical items for P300 amplitude,
P300 latency and reaction time on P300-based GKT

Shinji Hira, Risa Misaka and Yuki Hamamoto

The present study was designed to compare differences between critical and non-critical items for P300 amplitude, P300 latency and reaction time. Analyses were based on data from 5 different studies covering 173 guilty participants. Results revealed that critical items elicited significantly larger P300 amplitudes than non-critical items. Critical items also elicited significantly longer P300 latencies than non-critical items. This suggests that field applications may be feasible, with P300 as the dependent variable. However, no significant difference was seen between critical and non-critical items for reaction time.

[KEY WORDS: guilty knowledge test, P300 latency, reaction time]