

P300 숨긴정보검사에서 t 검증이 부트스트랩 방법보다 덜 민감한가?*

엄진섭
충북대학교 강사

손진훈
충남대학교 교수

박미숙[†]
서울한영대학교 교수

P300 숨긴정보검사에서 조사대상자가 거짓을 말하고 있는지 판단하기 위하여 관련자극의 P300 진폭이 무관련자극의 P300 진폭보다 통계적으로 유의하게 더 큰지를 평가한다. 구체적인 통계적 방법으로 독립표본 t 검증 또는 부트스트랩 방법을 사용할 수 있다. Rosenfeld와 Soskins, Bosh, Ryan(2004)은 “개인 내에서 관련자극과 무관련자극의 P300 평균을 비교하는데 사용하기에는 t 검증이 너무 둔감하다.”면서 부트스트랩 방법을 사용하였다. 본 연구의 목적은 P300 숨긴정보검사에서 t 검증의 검증력이 부트스트랩 방법보다 더 낮은지 평가하는 것이다. 이를 위하여 39명의 실험참가자로부터 측정된 뇌파자료를 이용하여 Monte Carlo 연구를 수행하였다. 연구결과, t 검증과 백분위를 이용한 부트스트랩 방법의 1 종 오류율은 서로 비슷하였으며, 백분위를 이용한 부트스트랩 방법의 검증력이 t 검증의 검증력보다 약간 더 높았다. 두 검증 방법의 1 종 오류율은 모두 유의수준보다 낮은 값을 보였으며, 검증력은 이론적인 t 검증의 검증력보다 약간 낮은 값을 보였다. 반면에 표준오차를 이용한 부트스트랩 방법의 1 종 오류율과 검증력은 이론적인 t 검증의 1 종 오류율 및 검증력과 비슷한 값을 보였으며, t 검증의 검증력보다 실험조건에 따라 .012 ~ .081 더 높았다.

주제어 : P300, 숨긴정보검사, t 검증, 부트스트랩

* 이 논문은 2019년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임
(NRF-2018R1C1B5085034).

† 교신저자 : 박미숙, 서울한영대학교 재활상담심리학과, (08274) 서울특별시 구로구 경인로 290-42
Tel : 02-2669-2468 / E-mail : misook_park1121@hanmail.net

자율신경계 반응을 이용하는 거짓말 탐지 기법을 보완하기 위하여 뇌파를 이용한 거짓말 탐지 방법인 P300 숨긴정보검사(P300-based concealed information test; P300 CIT)가 개발되었으며(Farwell & Donchin, 1991), 최근까지도 P300 CIT에 대한 연구가 활발히 진행되고 있다(Lukács, Grządziel, Kempkes, & Ansoerge, 2019; Rosenfeld, Davydova, Labkovsky, & Ward, 2019). P300 CIT는 범죠행장에 오직 범인만 알 수 있는 정보(예, 범행에 사용된 칼)가 남아있는 경우에 사용할 수 있는 방법으로 뇌파의 P300 성분을 종속측정치로 사용한다. 뇌파를 측정하면서 범죠행 용의자에게 범행에 사용된 칼(관련자극; probe)과 범행에 사용되지 않은 다수의 칼(무관련자극: irrelevant)을 무선적으로 하나씩 제시하고, 각 자극을 알고 있는지 질문한다. 용의자는 관련자극과 무관련자극에 모두 '모른다'는 응답을 하게 된다. 자극은 컴퓨터 모니터로 제시하고 응답은 마우스 버튼이나 키보드로 받는다. 일반적으로 무관련자극의 개수는 관련자극 개수의 4배로 구성하며, 뇌파 측정의 신뢰성을 위하여 관련자극의 제시횟수가 30회 이상이 되도록 전체 자극을 반복 제시한다. 자극제시가 끝난 후, 관련자극과 무관련자극에 대한 사건관련전위(event-related potential; ERP)를 산출하여 P300 진폭을 측정한다. 관련자극에 대한 P300 진폭이 무관련자극에 대한 P300 진폭보다 의미 있게 더 크다면, 용의자가 범행에 사용된 칼을 알고 있으면서 모른다고 거짓말 한 것으로 판단한다.

관련자극의 P300 진폭이 무관련자극의 P300 진폭보다 의미 있게 더 크지를 평가하는 다양한 방법이 있다. 그 중에서 가장 많은 연구자들이 사용하는 방법은 관련자극의 P300 진폭이 무관련자극의 P300 진폭보다 통계적으로

유의하게 더 크지를 평가하는 것이다(Ambach, Bursch, Stark, & Vaitl, 2010; Ellwanger, Rosenfeld, Sweet, & Bhatt, 1996; Kubo & Nitrono, 2009; Lefebvre, Marchand, Smith, & Connolly, 2009; Lukács et al., 2016; Rosenfeld, Angell, Johnson, & Qian, 1991; Verschuere, Rosenfeld, Winograd, Labkovsky, & Wiersema, 2009). 통계적 방법으로는 독립표본 t 검증(independent samples t -test)¹⁾과 부트스트랩 방법(bootstrap method)을 사용할 수 있다.

독립표본 t 검증은 P300 진폭이 가장 큰 100 ms 구간에서 단일 시행 뇌파들(single sweeps)의 평균 전위를 각각 측정하여 수행한다. 그림 1의 좌측에서 볼 수 있는 것처럼 단일 시행 뇌파들에서 값을 추출한다는 점을 제외하고는 일반적인 독립표본 t 검증과 동일하다. 독립표본 t 검증이 이론적인 표집분포로부터 유의확률을 계산하는 반면, 부트스트랩 방법은 표본 자료를 이용하여 경험적으로 구성된 표집분포로부터 확률을 계산한다. 이러한 이유로 부트스트랩 방법은 정규성 가정 없이 유의확률을 계산할 수 있다. 부트스트랩 방법은 그림 1의 우측에 제시된 것과 같이, 관련자극에 대한 N 개의 단일 시행 뇌파들에서 N 개를 무선 복원 추출하고, 무관련자극에서도 동일하게 복원추출한다. 복원추출된 표본에서 P300 진폭을 계산하고, 관련자극에 대한 P300 진폭에서 무관련자극에 대한 P300 진폭을 뺀 차이값을 구한다. 복원추출하고 P300 진폭 차이값을 구하는

1) 집단수준에서 분석할 때 관련자극과 무관련자극의 P300 측정치는 개인별로 쌍을 이루고 있으므로 대응표본 t 검증을 수행하지만, 개인내 수준에서는 관련자극의 측정치와 무관련자극의 측정치가 쌍을 이루고 있지 않으므로 독립표본 t 검증을 수행한다.

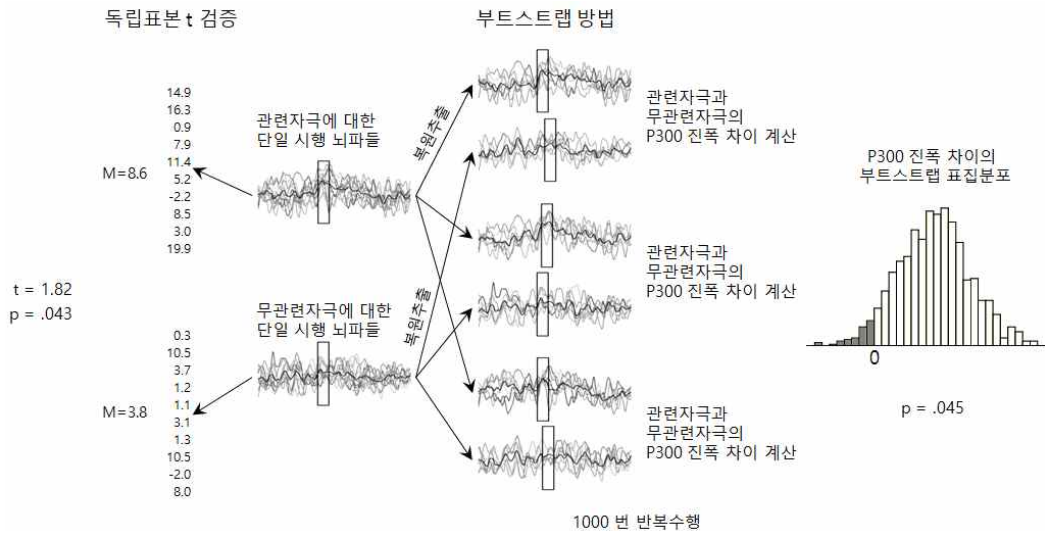


그림 1. P300 진폭 차이에 대한 독립표본 t 검증과 부트스트랩 방법. 굵은 선은 ERP이며 가는 선은 단일 시행 뇌파임.

과정을 1000번 반복하여 P300 진폭차이에 대한 부트스트랩 표집분포를 형성한다.

P300 CIT에서는 부트스트랩 표집분포로부터 유의확률을 계산할 때 두 가지 방법을 주로 사용하고 있다. 한 가지는 부트스트랩 표집분포의 백분위를 이용하는 방법으로(본 논문에서는 ‘백분위를 이용한 부트스트랩 방법’으로 부름) 표집분포에서 0보다 큰 P300 진폭 차이가 95% 이상이면 유의수준 .05에서 관련자극에 대한 P300 진폭이 무관련자극에 대한 P300 진폭보다 통계적으로 유의하게 더 크다고 판단하는 것이다(송인옥, 김혜민, 이경은, 장은희, 김현택, 2018; 전하정, 손진훈, 박광배, 엄진섭, 2017; 정은경, 김영윤, 2013; Rosenfeld & Donchin, 2015; Cutmore, Djakovic, Keibell, & Shum, 2009). 다른 한 가지는 부트스트랩 표집분포의 표준오차를 이용하는 방법으로(본 논문에서는 ‘표준오차를 이용한 부트스트랩 방법’으로 부름) t 검증 공식의 분모에 표본

표준오차를 사용하는 대신에 부트스트랩 표집분포의 표준오차를 사용한다는 점만을 제외하고 t 검증 절차와 동일하다(Bablani & Tripathi, 2018; Matsuda, Nittono, & Ogawa, 2011; Rosenfeld et al., 2004).

통계적 검증의 1종 오류율과 검증력을 확인할 때에는 일반적으로 몬테 카를로 연구(Monte Carlo study)를 사용한다(Ahad, Abdullah, & Lai, 2012; Dwivedi, Mallawaarachchi, & Alvarado, 2017; Konietzschke & Pauly, 2014). P300 CIT가 아닌 일반적인 상황에서 몬테 카를로 연구를 통하여 독립표본 t 검증과 부트스트랩 방법의 결과를 비교한 연구들이 있다. 이 연구들에 의하면 정규분포 가정과 변량의 동질성 가정이 충족된 경우에는 t 검증과 부트스트랩 방법의 검증력이 비슷하며, 모집단 분포가 정규분포로부터 크게 이탈하면서 표본 크기가 20 이하로 작거나 변량의 이질성이 큰 경우에는 부트스트랩 방법이 t 검증보다 더

우수하다고 보고하고 있다(Ahad et al., 2012; Dwivedi et al., 2017; Good, 2005; Konietschke & Pauly, 2014).

P300 CIT에서 단일 시행의 P300 진폭(single trial P300 amplitude)은 정규분포에서 크게 벗어나지 않으며(엄진섭, 전하정, 손진훈, 2018), 관련자극과 무관련자극의 제시 횟수(즉, 표본크기)는 각각 최소 30 이상이다. 단일 시행의 P300 진폭은 순수한 P300 진폭에 배경뇌파가 더해진 것인데(Luck, 2014), 관련자극과 무관련자극에 포함된 배경뇌파는 동질적이므로 개인 내에서 관련자극과 무관련자극에 대한 P300 진폭의 변량은 동질적이다. 결과적으로 P300 CIT에서 측정되는 뇌파는 t 검증을 수행하기 위한 두 가지 조건, 즉 정규성 가정과 변량의 동질성 가정을 충족하고 있다. 따라서 개인 내에서 관련자극과 무관련자극 간의 P300 진폭을 비교하는 통계적 방법으로써 독립표본 t 검증과 부트스트랩 방법의 검증력은 크게 다르지 않을 것으로 추론된다.

P300 CIT에서 독립표본 t 검증과 부트스트랩 방법의 검증력이 비슷할 것으로 예측됨에도 불구하고, P300 CIT 연구자들은 거의 대부분 부트스트랩 방법을 사용하고 있다(송인욱 등, 2018; 전하정 등, 2017; 정은경, 김영운, 2013; Bergström, Anderson, Buda, Simons, & Richardson-Klavehn, 2013; Cutmore et al., 2009; Kubo & Nittono, 2009; Lukács et al., 2019; Marchand, Inglis-Assaff, & Lefebvre, 2013; Matsuda et al., 2011; Meijer, Smulders, Merckelbach, & Wolf, 2007; Verschuere et al., 2009). P300 CIT에서 부트스트랩 방법이 널리 사용된 이유는 “개인 내에서 관련자극과 무관련자극의 평균 P300을 비교하는데 사용하기에는 t 검증이 너무 둔감하다(t tests on single sweeps are too

insensitive to use to compare mean probe and irrelevant P300s within individuals; see Rosenfeld et al., 1991)”는 Rosenfeld 등(2004)의 주장 때문일 것이다.

Rosenfeld 등(2004)의 이러한 주장은 P300 CIT에서 독립표본 t 검증과 부트스트랩 방법의 검증력이 비슷할 것이라는 본 연구자들의 추론과 다른 것이다. Rosenfeld 등(2004)은 자신들의 선행연구(Rosenfeld et al., 1991) 결과를 근거로 부트스트랩 방법의 우월성을 주장하였다. Rosenfeld 등(1991)의 연구에서 8명의 유죄집단을 대상으로 확인한 부트스트랩 방법의 정확판단율은 87.5%였으며, t 검증의 정확판단율은 25.0%였다. 그러나 이 연구에서 사용한 부트스트랩 방법은 일반적인 부트스트랩 방법과 완전히 다른 방법이었다. 부트스트랩 방법에서는 표본들(즉, 단일 시행 뇌파들)을 재표본 추출하여 표집분포를 만들어야만 하지만(Efron & Tibshirani, 1994), Rosenfeld 등(1991)은 표본 평균(즉, ERP의 시계열 자료)만 사용하였다. 따라서 Rosenfeld 등(1991)의 연구가 현재 널리 사용하는 부트스트랩 방법의 우월성을 지지하는 결과라고 보기 어렵다.

본 연구의 목적은 P300 CIT에서 독립표본 t 검증과 두 가지 부트스트랩 방법의 검증력을 비교하는 것이다. 정확한 검증력을 산출하기 위해서는 수천에서 수만의 평가자료가 필요하다. 사람을 대상으로 이러한 자료를 모으는 것은 거의 불가능하므로, 통계적 방법의 검증력을 평가할 때에는 대부분 컴퓨터 모의실험을 이용하는 몬테 카를로 연구가 사용된다(Ahad et al., 2012; Dwivedi et al., 2017; Konietschke & Pauly, 2014). 본 연구에서도 몬테 카를로 연구를 통하여 독립표본 t 검증과 두 가지 부트스트랩 방법의 검증력을 비교하

였다. 일반적인 몬테 카를로 연구에서는 특정 변인 Y에 대한 모집단 평균, 표준편차 및 분포의 모양을 정의한 후, 컴퓨터로 구성된 모집단으로부터 무선표본을 추출하여 통계적 방법의 검증력을 확인한다. 그러나 P300 CIT의 분석 단위인 개별 시행 뇌파는 단일한 값이 아닌 1초 내외의 시계열 자료이므로, 컴퓨터로 모집단 분포를 정의하기가 매우 어렵다. 따라서 본 연구에서는 실험을 통하여 측정된 뇌파자료와 ERP를 이용하여 몬테 카를로 연구에 사용할 유한 모집단을 정의한 후, 독립표본 t 검증과 두 가지 부트스트랩 방법의 검증력을 산출하였다. 어떤 통계방법의 검증력이 높게 산출되더라도 1 중 오류율이 유의수준보다 높으면 이 검증력은 과대추정된 것이므로, 각 검증방법의 1 중 오류율을 함께 확인하였다.

방 법

실험조건

P300 CIT에서 독립표본 t 검증과 부트스트랩 방법의 1 중 오류율 및 검증력을 평가하기 위하여, 관련자극과 무관련자극의 순수한 P300 진폭 크기와 유의수준을 정하였다. 무죄조건에서 관련자극과 무관련자극의 순수한 P300 진폭 크기는 동일하게 설정하였다. 유죄조건에서는 Cohen(1988)의 효과크기 분류에 따라 작은 효과크기($d=0.2$)와 중간 효과크기($d=0.5$), 큰 효과크기($d=0.8$)에 해당하는 만큼 관련자극의 P300 진폭을 무관련자극의 P300 진폭보다 더 크게 설정하였다. 유의수준은 선행연구에서 주로 사용하는 .10(Rosenfeld, Shue,

& Singer, 2007; Verschuere et al., 2009)과 .05(Matsuda et al., 2011; Meijer et al., 2007; Merterns & Allen, 2008)를 사용하였다.

모집단 정의

몬테 카를로 연구에 사용할 단일 시행 뇌파의 모집단을 구성하기 위하여 세 단계를 거쳤다. 첫 번째로 순수한 ERP를 만들었다. 두 번째로 P300 CIT 실험을 통하여 측정된 단일 시행 뇌파에서 ERP를 제거하여 배경뇌파를 만들었다. 세 번째로 순수한 ERP에 배경뇌파를 더하여 모집단을 만들었다. 각 단계의 구체적 절차는 다음과 같다.

몬테 카를로 연구에 사용할 순수한 ERP를 만들었다. 순수한 ERP는 일반적인 몬테 카를로 연구에서 모집단의 평균에 해당한다. 본 연구에서는 엄진섭, 음영지, 장은정, 정이내, 손진훈(2015)의 단순 유죄 조건에서 측정된 무관련자극의 ERP를 무관련자극에 대한 순수한 ERP로 사용하였다. 순수한 ERP의 전체 길이는 자극 제시 시점부터 1200 ms였다. 분석방법에서 설명할 정점-정점 방법으로 측정된 P300 진폭은 $5.64 \mu V$ 였다. 무관련자극에 대한 순수한 ERP를 이용하여 네 가지 실험조건에 해당하는 관련자극에 대한 순수한 ERP를 만들었다. 먼저 무죄조건에서는 관련자극에 대한 ERP와 무관련자극에 대한 ERP가 동일하므로, 무관련자극에 대한 순수한 ERP를 그대로 관련자극에 대한 순수한 ERP로 사용하였다. 효과크기 $d=0.2$ 인 유죄조건에서는 무관련자극에 대한 ERP의 P300 진폭을 $d=0.2$ 만큼 더 크게 만들어 구성하였다. 이 경우에 관련자극의 P300 진폭은 $7.70 \mu V$ 가 되도록 구성하였으며, $d=0.5$ 인 경우에 P300 진폭은 $10.79 \mu V$, $d=0.8$

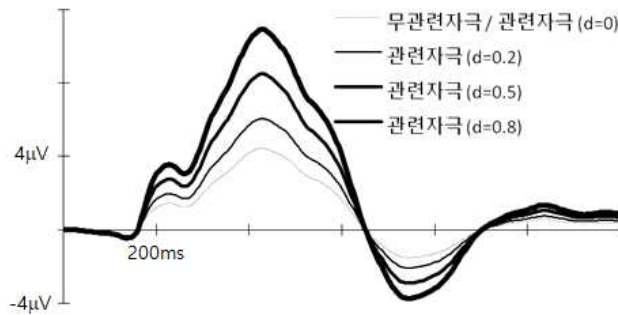


그림 2. 몬테 카를로 연구에 사용한 관련자극과 무관련자극의 순수한 ERP

인 경우에 P300 진폭은 13.88 μV 가 되도록 구성하였다. 그림 2에 몬테 카를로 연구에 사용한 순수한 ERP가 제시되어 있다.

순수한 ERP에 더하여 단일 시행 뇌파를 만드는데 사용할 배경뇌파(즉 노이즈)를 만들었다. 배경뇌파는 일반적인 몬테 카를로 연구에서 모집단의 표준편차를 정의하는 것에 해당한다. 휴식상태의 뇌파와 과제를 수행 중인 뇌파의 구성성분(즉, 알파파와 베타파 등)은 서로 다르다(Fernández et al., 2005; Mathewson et al., 2015). 본 연구에서는 P300 CIT에서 발생하는 배경뇌파와 최대한 유사한 배경뇌파를 만들기 위하여, P300 CIT에서 측정된 단일 시행 뇌파를 이용하여 배경뇌파를 만들었다.

엄진섭 등(2015)의 단순 유죄 조건에서 측정된 단일 시행 뇌파를 이용하여 배경뇌파를 구성하였다. P300 CIT의 분석방법에 관한 연구에서는 선행연구의 뇌파자료가 있는 경우에 이를 활용하는 것이 일반적이다(엄진섭 등, 2018; 정은경, 김영윤, 2013; Abootalebi, Moradi, & Khalilzadeh, 2009). 엄진섭 등(2015)의 연구에서는 남녀 대학생 39명을 대상으로 자기참조적 정보를 이용한 P300 CIT를 실시하였다. P300 CIT에서 실험참가자의 이름을 관련자극으로 하고, 실험참가자에게 아무런 의미가 없

는 이름 4개를 무관련자극으로 하고, 또 다른 이름 하나를 목표자극으로 하였다. 6개의 자극을 하나씩 무선적으로 제시하였으며, 6개의 자극을 총 40회 반복 제시하였다. 자극은 평균 3000 ms 마다 한번씩 300 ms 동안 제시하였으며, 목표자극이 제시되면 마우스 왼쪽 버튼을 누르고 다른 자극이 제시되면 마우스 오른쪽 버튼을 누르도록 하였다. 뇌파는 Fz, Cz, 및 Pz에서 측정하였으며, 기준전극과 접지전극은 양쪽 귓볼과 이마에 부착하였다. 왼쪽 눈의 위와 아래에 대각선으로 전극을 부착하여 안전도(electrooculogram: EOG)를 기록하였다. 자극은 SuperLab v5.0(Cedrus Corporation, San Pedro, CA, USA)으로 제시하였으며, Grass Model 12 Neurodata Acquisition System(Grass Instruments, Quincy, MA, USA)을 사용하여 뇌파를 20000배 증폭하였고, 0.3~30Hz 대역여과하였다.

본 연구에서는 엄진섭 등(2015)에서 39명을 대상으로 Pz에서 측정된 뇌파를 이용하였다. 먼저 각 참가자별로 안전도를 이용하여 눈 깜빡임에 의한 뇌파의 변형을 교정하였다(Semlitsch, Anderer, Schuster, & Presslich, 1986). 다음으로 각 참가자별로 뇌파를 각 자극 제시 시점부터 자극제시 후 1200 ms까지의 구간으로 나누어 목표자극과 관련자극, 무관련자극

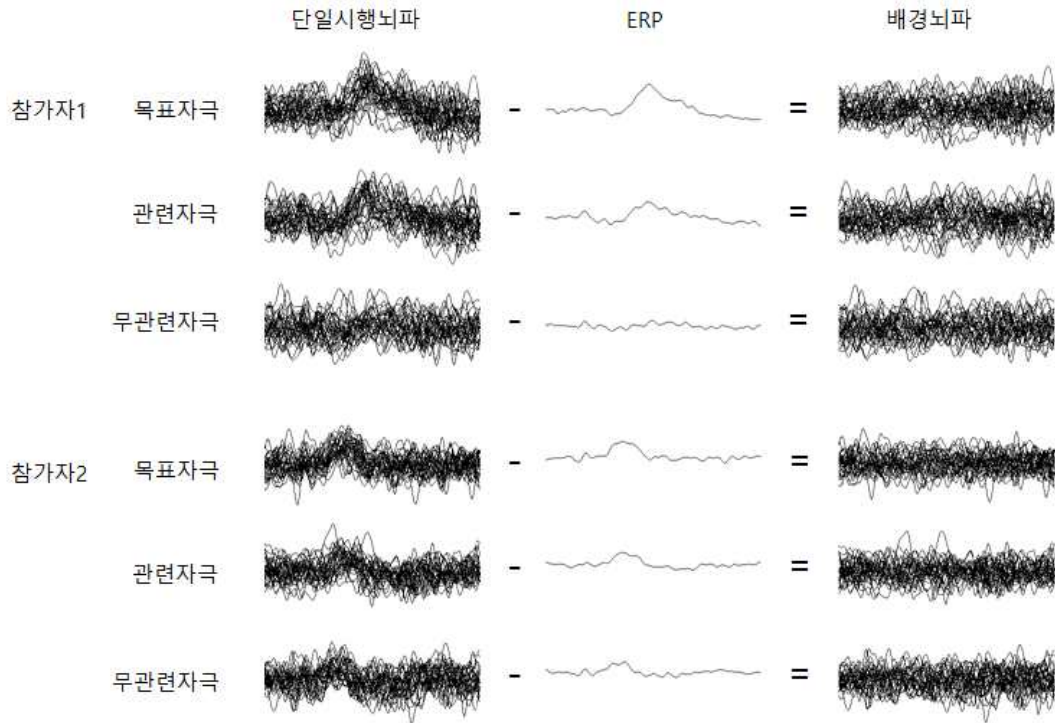


그림 3. 배경뇌파 추출 절차. 단일 시행 뇌파에서 ERP만 제거하여 배경뇌파를 만들었음.

에 대한 단일 시행 뇌파들을 각각 추출하였다. 이때 $\pm 50 \mu V$ 를 넘어서는 값이 있는 시행은 오염된 시행으로 판단하여 분석에서 제외하였다 (Verleger, 1993). 그런 다음 각 참가자별로 단일 시행 뇌파들을 평균하여 목표자극, 관련자극, 무관련자극 각각의 ERP를 산출하였다. 마지막으로 Hanslmayr 등(2006)의 절차에서와 같이 각 참가자별로 단일 시행 뇌파들에서 ERP를 제거하여 배경뇌파를 구성하였다. 최종적으로 1200 ms 길이의 배경뇌파 9091개를 구성하였다. 그림 3에 배경뇌파를 만드는 과정이 제시되어 있다.

배경뇌파에 순수한 ERP를 더하여 몬테 카를로 연구에 사용할 모집단을 만들었다. 무관련자극의 모집단은 9091개의 배경뇌파 각각에

무관련자극에 대한 순수한 ERP를 더하여 9091개의 유한모집단을 구성하였으며, 효과크기 $d=0.0$ 인 관련자극의 모집단은 9091개의 배경뇌파 각각에 $d=0.0$ 인 관련자극에 대한 순수한 ERP를 더하여 9091개의 유한모집단을 만들었다. 효과크기 $d=0.2$, $d=0.5$, $d=0.8$ 인 관련자극의 모집단도 같은 방법으로 만들었다.

무선표본추출

앞서 구성한 단일 시행 뇌파의 모집단에서 무선 표본을 추출하였다. 무죄조건에서는 무관련자극의 모집단에서 단일 시행 뇌파를 30개 무선 표본 추출하고, 효과크기 $d=0.0$ 인 관련자극의 모집단에서 단일 시행 뇌파를 30개

무선 표본 추출하였다. 유죄조건에서는 무관련자극의 모집단에서 단일 시행 뇌파를 30개 무선 표본 추출하고, 효과크기 $d=0.2$ (또는 0.5, 0.8)인 관련자극의 모집단에서 단일 시행 뇌파를 30개 무선 표본 추출하였다.

분석방법

각 실험조건별로 만들어진 뇌파자료에 대하여 t 검증과 두 가지 부트스트랩 방법으로 관련자극의 P300 진폭이 무관련자극의 P300 진폭 보다 통계적으로 유의하게 더 큰지를 검증하였다. P300 진폭은 정점-정점 방법으로 측정하였다. 정점-정점 방법은 P300 정점과 P300 이후에 나타는 부적 정점 간의 진폭 차이를 측정한다(Rosenfeld et al., 2004). 구체적으로 자극 제시 후 300 ms ~ 600 ms 사이에 100 ms 구간의 평균값이 가장 큰 값에서 P300 정점이 나타난 이후부터 1200 ms 사이에 100 ms 구간의 평균값이 가장 작은 값을 뺀 값으로 정의하였다.

독립표본 t 검증은 다음과 같이 진행되었다. 1) 관련자극에 대한 30개의 단일 시행 뇌파들을 평균하여 관련자극에 대한 ERP를 산출하고, 동일한 방법으로 무관련자극에 대한 30개의 단일 시행 뇌파들을 평균하여 무관련자극에 대한 ERP를 산출한다. 2) 관련자극과 무관련자극의 ERP에서 P300 진폭을 측정하고, P300의 정적 정점과 P300 이후 부적 정점이 나타나는 잠재시간을 측정한다. 3) 관련자극에 대한 30개의 단일 시행 뇌파 각각에서 P300 정점 및 부적 정점의 잠재시간에 해당하는 구간의 진폭(즉 단일 시행 P300 진폭)을 계산하고, 무관련자극에 대한 단일 시행 뇌파 각각에 대해서도 동일한 방법으로 단일 시행 P300

진폭을 계산한다. 이 과정을 거치면 관련자극에 대한 30개와 무관련자극에 대한 30개의 단일 시행 P300 진폭 값들이 얻어진다. 4) 관련자극과 무관련자극에 대한 단일 시행 P300 진폭 값들을 이용하여 t 값과 유의확률을 계산한다. 5) 일방검증과 유의수준 .05와 .10에서 영가설 기각여부를 결정한다.

백분위를 이용한 부트스트랩 방법은 다음과 같은 절차로 진행되었다. 1) 관련자극에 대한 30개의 단일 시행 뇌파에서 30개를 복원추출하고, 무관련자극에 대한 30개의 단일 시행 뇌파로부터 30개를 복원추출 한다. 2) 복원추출된 단일 시행 뇌파들로 관련자극과 무관련자극 각각에 대한 ERP를 만들고, ERP에서 P300 진폭을 계산한다. 3) 관련자극에 대한 P300 진폭에서 무관련자극에 대한 P300 진폭을 뺀 차이값을 구한다. 4) 1)에서 3)까지의 과정을 1000번 반복하여, 1000개의 P300 진폭차이 값으로 부트스트랩 표집분포를 구성한다. 5) 일방검증과 유의수준 .10에서는 부트스트랩 표집분포에서 0 보다 작은 값의 비율이 .10 이하이면 유죄 판단을 하며, 유의수준 .05에서는 0 보다 작은 값의 비율이 .05 이하이면 유죄 판단을 한다.

표준오차를 이용한 부트스트랩 방법에서는 백분위를 이용한 부트스트랩 방법의 1)부터 4)까지는 동일하게 진행된다. 5) 부트스트랩 표집분포로부터 표준편차(즉, 표본 평균의 표준오차)를 계산한다. 6) 원래 표본의 관련자극과 무관련자극 간 P300 진폭 차이값을 5)에서 계산한 표준오차로 나누어 t 값을 계산한다. 7) 일방검증과 유의수준 .05와 .10에서 영가설 기각여부를 결정한다.

무죄조건에서 단일 시행 뇌파를 표본추출하고 t 검증과 두 가지 부트스트랩 방법으로

P300 진폭차이검증을 수행하는 과정을 100,000번 반복하여 각 검증방법의 1종 오류율을 산출하였다. 세 가지 수준으로 구성된 유죄조건에서도 동일한 과정을 100,000번 반복하여 t검증과 부트스트랩 방법의 검증력을 산출하였다. 마지막으로 몬테 카를로 연구를 통하여 산출한 검증력의 크기를 평가하기 위하여, G*Power v3.1.9.4를 이용하여 이론적인 t검증의 검증력을 산출한 후 본 연구에서 산출한 검증력과 비교하였다.

결 과

독립표본 t검증과 두 가지 부트스트랩 방법의 1종 오류율, 검증력, 및 이론적인 t검증의 검증력을 표 1에 제시하였다. 무죄조건에서 t검증과 백분위를 사용한 부트스트랩 방법의 1종 오류율은 유의수준 .10에서 .071과 .076으로 나타났고 유의 수준 .05에서 .031과 .034로 나타나, 두 가지 검증방법의 1종 오류

율은 비슷하였지만 유의수준보다 낮은 값을 보였다. 반면에 표준오차를 사용한 부트스트랩 방법의 1종 오류율은 유의수준 .10에서 .099, 유의수준 .05에서 .051로 나타나 유의수준과 거의 비슷한 1종 오류율을 보였다.

유죄조건에서 백분위를 사용한 부트스트랩 방법의 검증력은 t검증의 검증력 보다 실험조건에 따라 .009 ~ .037 더 높은 것으로 나타났다. 그러나 백분위를 사용한 부트스트랩 방법의 검증력은 이론적인 t검증의 검증력보다 실험조건에 따라 .002 ~ .036 더 낮았다. 표준오차를 사용한 부트스트랩 방법의 검증력은 백분위를 사용한 부트스트랩 방법의 검증력보다 실험조건에 따라 .003 ~ .040 더 높았으며, t검증의 검증력보다 실험조건에 따라 .012 ~ .081 더 높았다. 표준오차를 사용한 부트스트랩 방법의 검증력은 이론적인 t검증의 검증력과 매우 유사하였으며, 표준오차를 사용한 부트스트랩 방법의 검증력이 실험조건에 따라 .001 ~ .016 더 높았다.

표 1. 독립표본 t검증과 부트스트랩 방법의 1종 오류율과 검증력, 이론적인 검증력

	Cohen의 <i>d</i>	α	t검증	부트스트랩		이론적 검증력
				백분위 사용	표준오차 사용	
무죄조건	0.0	.10	.071	.076	.099	-
		.05	.031	.034	.051	-
유죄조건	0.2	.10	.249	.268	.308	.304
		.05	.138	.155	.195	.190
	0.5	.10	.696	.723	.747	.739
		.05	.541	.578	.622	.606
0.8	.10	.953	.962	.965	.964	
	.05	.900	.917	.928	.922	

주. 무죄조건 값들은 1종 오류율이며, 유죄조건 값들은 검증력임.

논 의

본 연구에서는 P300 CIT에서 부트스트랩 방법이 독립표본 t 검증보다 검증력이 더 높다는 Rosenfeld 등(2004)의 주장을 평가하기 위하여 뇌파를 이용한 몬테 카를로 연구를 수행하였다.

연구결과, 표준오차를 이용한 부트스트랩 방법의 검증력이 가장 높았으며, 백분위를 이용한 부트스트랩 방법의 검증력이 t 검증의 검증력보다 약간 더 높았다. Rosenfeld 등(2004)의 주장처럼 부트스트랩 방법의 검증력이 t 검증의 검증력 보다 더 높게 나타났다. 그러나 표준오차를 이용한 부트스트랩 방법과 t 검증의 검증력 차이는 최대 .08로 Rosenfeld 등(2004)이 주장한 것처럼 t 검증의 민감도(즉 검증력)가 부트스트랩 검증의 민감도(즉 검증력)보다 '너무' 낮지는 않았다.

Rosenfeld 등(2004)이 자신들의 주장에 대한 참고문헌으로 인용한 Rosenfeld 등(1991)의 연구에서는 t 검증의 민감도가 .250이었던 반면 부트스트랩 방법의 민감도는 .875로 둘 간에 .625의 큰 차이가 있었다. 본 연구의 결과가 Rosenfeld 등(1991)의 결과와 다른 이유는 Rosenfeld 등(1991)이 사용한 부트스트랩 방법에 문제가 있었기 때문일 것이다. Rosenfeld 등(1991)이 사용한 방법은 부트스트랩 방법을 고안하고 타당성을 증명한 Efron과 Tibshirani (1986)의 방법과는 완전히 다른 것이었다. Rosenfeld 등(1991)은 관련자극과 무관련자극의 개별 시행 뇌파를 이용하는 대신에, 개별 시행 뇌파들의 평균인 관련자극과 무관련자극의 ERP만 이용하였다. ERP의 468 ~ 1420 ms 구간에 ERP를 구성하는 120개의 시계열 값들이 있었는데, 이 값들을 복원추출하여 P300 진폭

차이에 대한 표집분포를 구성하였다. 그러나 이 방법으로 만들어진 부트스트랩 표집분포는 P300 진폭차이에 대한 표집분포가 아니라, Rosenfeld 등(1991)도 밝혔듯이 P300 잠재시간의 변화에 대한 표집분포이다. 이 방법은 타당성이 증명된 바가 없으며, Rosenfeld 등(1991)에서 단 한 번만 사용되었다. 본 연구의 저자들이 알기로는 각 시행의 측정치 또는 이 측정치들의 변량이 없이 표본 평균(즉, ERP의 P300)의 표준오차를 추정할 수 있는 방법론은 존재하지 않는다. 결과적으로 Rosenfeld 등(1991)이 사용한 부트스트랩 방법은 타당성이 결여되어 있다고 할 수 있다.

P300 CIT에서 현재 사용되고 있는 것과 같은 방식의 표준오차를 이용한 부트스트랩 방법과 t 검증의 결과를 비교한 연구가 두 편 있다(Ellwanger et al., 1996; Rosenfeld, Sweet, Chuang, Ellwanger, & Song, 1996). Rosenfeld 등(1996)의 연구에서 45명의 참가자에게 P300 CIT를 실시한 후 독립표본 t 검증과 부트스트랩 방법으로 P300 진폭 차이검증을 실시한 결과, t 검증의 검증력이 부트스트랩 방법보다 약간 더 좋은 것으로 나타났다. 그러나 41명의 참가자가 포함된 Ellwanger 등(1996)의 연구에서는 부트스트랩 방법의 검증력이 t 검증보다 약간 더 높은 경향이 있어서, Rosenfeld 등(1996)과 다른 결과가 나타났다. 이 연구결과들은 50명 미만의 자료를 대상으로 검증력을 확인하였기 때문에 표본오차에 의한 부정확성이 크긴 하였지만, 두 연구 결과를 종합하면 부트스트랩 방법의 검증력이 t 검증의 검증력보다 매우 높을 것 같지는 않다. 즉, 이 두 편의 연구결과는 Rosenfeld 등(1991)의 결과보다는 본 연구의 결과와 더 유사하다.

본 연구에서 t 검증의 1 종 오류율은 유의

수준보다 낮았으며, 백분위를 사용한 부트스트랩 검증의 1 종 오류율도 유의수준보다 낮았다. 반면에 표준오차를 이용한 부트스트랩 검증의 1 종 오류율은 유의수준과 매우 비슷하였다. 결과적으로 표준오차를 이용한 부트스트랩 검증은 1 종 오류율을 적절하게 통제하는 검증인 반면 t 검증과 백분위를 사용한 부트스트랩 방법은 1 종 오류율을 적절하게 통제하지 못하여 검증력이 약간 낮아진 것으로 보인다. 일반적으로 t 검증과 백분위를 이용한 부트스트랩 방법 모두 1 종 오류율을 적절하게 통제하는 방법임에도 불구하고(Good, 2005), P300 진폭차이검증에서 이러한 결과가 나타난 원인은 P300 진폭을 측정하는 방법 때문인 것으로 추론된다. 본 연구에서 P300 진폭은 300 ~ 600 ms 내에서 100 ms 구간의 평균 전위가 가장 높은 값으로 정의하였다. 이 정의에 따르면 P300이 나타나는 시간(즉, 100 ms 구간)은 표본마다 약간씩 달라질 것이다. 즉 모집단에서는 P300이 나타나는 시간이 고정되어 있지만, 표본에서는 P300이 나타나는 시간이 변하게 된다. 이로 인해서 표본에서 측정된 오차변량은 표본오차에 의한 오차변량에 P300이 발생하는 시간에 대한 오차변량이 더해졌을 것이고, 결과적으로 오차변량이 더 크게 나타났을 것이다. 이로 인해서 검증력이 낮아졌을 가능성이 있다. 그러나 이러한 추론은 표준오차를 이용한 부트스트랩 방법에서 검증력이 낮아지지 않은 이유를 설명하지는 못한다. 추후 연구에서는 P300 CIT에서 표준오차를 이용한 부트스트랩 방법과 백분위를 이용한 부트스트랩 방법의 검증력 차이를 유발하는 원인을 찾을 필요성이 있겠다.

결과적으로 “개인 내에서 관련자극과 무관련자극의 P300 평균을 비교하는데 사용하기에

는 t 검증이 너무 둔감하다”는 Rosenfeld 등(2004)의 주장은 부트스트랩 방법의 검증력을 너무 과대추정한 것이다. 그러나 표준오차를 사용한 부트스트랩 방법의 검증력이 t 검증이나 백분위를 사용한 부트스트랩 방법보다 약간 더 높으므로, P300 CIT에서 표준오차를 사용한 부트스트랩 방법을 사용하는 것이 더 좋을 것이다. 본 연구는 개인 내에서 두 자극의 P300 진폭을 비교할 때 표준오차를 사용한 부트스트랩 방법이 독립표본 t 검증보다 1 종 오류를 더 정확하게 통제하면서 검증력이 더 높다는 것을 증명한 첫 연구라는 가치를 지닌다.

본 연구에서 사용한 P300 CIT는 Rosenfeld 등(2004)이 사용한 P300 CIT와 한 가지 차이점이 있었다. Rosenfeld 등(2004)에서는 관련자극과 무관련자극의 시행수가 각각 30과 120이었지만, 본 연구에서는 관련자극과 무관련자극의 시행수를 모두 30으로 동일하게 하였다. 본 연구에서 시행수를 동일하게 사용한 이유는 두 자극의 시행수가 서로 다른 경우에 두 자극의 P300 진폭을 비교하면 통계적 검증의 1 종 오류가 증가하고(Luck, 2014), 검증력이 과대추정되기 때문이었다. ERP에는 고주파 노이즈가 포함되어 있는데, 시행수가 작을수록 고주파 노이즈가 더 많이 포함된다(Luck, 2014). 고주파 노이즈는 P300 진폭을 과대추정하게 만들기 때문에 시행수가 더 적은 관련자극의 P300 진폭이 시행수가 더 많은 무관련자극의 P300 진폭보다 더 크게 과대추정하게 된다. 결과적으로 통계적 검증에서 두 자극의 순수한 P300 진폭은 동일하다 하더라도 관련자극의 P300 진폭이 더 크게 추정되고, 이로 인해 1 종 오류가 증가하게 된다. 본 연구에서는 독립표본 t 검증과 두 가지 부트스트랩

방법의 1 종 오류율과 검증력을 정확하게 추정하기 위하여 관련자극과 무관련자극의 시행수를 동일하게 설정하였다. 독립표본 t 검증과 부트스트랩 방법에는 비교하는 두 조건의 사례수가 동일하여야 한다는 제약이 없으므로, 표준오차를 사용한 부트스트랩 방법의 검증력이 t 검증의 검증력보다 약간 더 높다는 결과는 두 자극의 시행수가 서로 다른 경우에도 일반화 가능할 것이다. P300 CIT에서는 일반적으로 관련자극 1개 당 무관련자극 4개를 사용하여 관련자극과 무관련자극의 시행수가 30:120이 되지만, 통계적 검증을 실시할 때 관련자극 1개와 무관련자극 중 P300 진폭이 가장 큰 1개의 P300 진폭을 비교하는 경우가 많다(엄진섭 등, 2015; Bowman, Filetti, Alsufyani, Janssen, & Su, 2014; Lefebvre et al., 2009; Lukács et al., 2016, Meixner & Rosenfeld, 2014). 이 경우에는 관련자극과 무관련자극의 시행수가 30:30으로 동일하고, 본 연구의 결과가 그대로 적용될 수 있다.

Rosenfeld와 동료들은 Rosenfeld 등(2008)의 연구부터 변형된 부트스트랩 방법을 주로 사용하고 있으며, 최근 다른 연구자들도 이 방법을 이용한 연구를 수행하였다(송인욱 등, 2018; Lukács et al., 2019). 변형된 부트스트랩 방법은 관련자극에 대한 시행들을 복원추출할 때 관련자극의 시행수 만큼 추출하고, 무관련자극에 대한 시행들을 복원추출할 때에도 관련자극의 시행수 만큼 추출하여 관련자극과 무관련자극의 시행수를 동일하게 만드는 것이다. 그러나 부트스트랩 방법에서는 반드시 원래의 표본수 만큼 복원추출해야만하므로(Efron, 1992; Hernández et al., 2017) 변형된 부트스트랩 방법의 타당성이 의문시된다. 또한 전체 시행들 중에 일부 시행만 추출하여 통계분석

에 활용하는 경우에는 통계적 검증력이 낮아질 것이다(Luck, 2014). 따라서 추후 연구에서는 변형된 부트스트랩 방법의 타당성을 확인할 필요성이 있겠다.

참고문헌

- 송인욱, 김혜민, 이경은, 장은희, 김현택 (2018). P300-CIT 부트스트랩 분석 비교. *한국심리학회지: 법*, 9(2), 75-99.
- 엄진섭, 음영지, 장은정, 정이내, 손진훈 (2015). P300 숨진정보검사와 대응수단. *감성과학*, 18(1), 39-48.
- 엄진섭, 전하정, 손진훈 (2018). P300 숨진정보검사의 오류공정률. *한국심리학회지: 인지 및 생물*, 30(3), 241-259.
- 전하정, 손진훈, 박광배, 엄진섭 (2017). 반응양식이 P300 숨진정보검사의 정확도에 미치는 영향. *감성과학*, 20(3), 109-118.
- 정은경, 김영윤 (2013). P300-기반 거짓지식검사의 탐지정확률 비교. *한국심리학회지: 인지 및 생물*, 25(1), 75-92.
- Abootalebi, V., Moradi, M. H., & Khalilzadeh, M. A. (2009). A new approach for EEG feature extraction in P300-based lie detection. *Computer methods and programs in biomedicine*, 94(1), 48-57.
- Ahad, N. A., Abdullah, S., & Lai, C. H. (2012). Relative power performance of t-test and bootstrap procedure for two-sample. *Pertanika Journal of Science & Technology*, 20(1), 43-52.
- Ambach, W., Bursch, S., Stark, R., & Vaitl, D. (2010). A Concealed Information Test with multimodal measurement. *International Journal*

- of *Psychophysiology*, 75(3), 258-267.
- Bablani, A., & Tripathi, D. (2018). A Review on Methods Applied on P300-Based Lie Detectors. *In Advances in Machine Learning and Data Science* (pp. 251-257). Springer, Singapore.
- Bergström, Z. M., Anderson, M. C., Buda, M., Simons, J. S., & Richardson-Klavehn, A. (2013). Intentional retrieval suppression can conceal guilty knowledge in ERP memory detection tests. *Biological psychology*, 94(1), 1-11.
- Bowman, H., Filetti, M., Alsufyani, A., Janssen, D., & Su, L. (2014). Countering countermeasures: Detecting identity lies by detecting conscious breakthrough. *PloS one*, 9(3), e90595.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.
- Cutmore, T. R., Djakovic, T., Keibell, M. R., & Shum, D. H. (2009). An object cue is more effective than a word in ERP-based detection of deception. *International Journal of Psychophysiology*, 71(3), 185-192.
- Dwivedi, A. K., Mallawaarachchi, I., & Alvarado, L. A. (2017). Analysis of small sample size studies using nonparametric bootstrap test with pooled resampling method. *Statistics in medicine*, 36(14), 2187-2205.
- Efron, B. (1992). Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics* (pp. 569-593). Springer, New York, NY.
- Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science*, 1(1), 54-75.
- Efron, B., & Tibshirani, R. (1994). *An Introduction to the Bootstrap*. CRC Press.
- Ellwanger, J., Rosenfeld, J. P., Sweet, J. J., & Bhatt, M. (1996). Detecting simulated amnesia for autobiographical and recently learned information using the P300 event-related potential. *International Journal of Psychophysiology*, 23(1-2), 9-23.
- Farwell, L. A. & Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event-related Potentials. *Psychophysiology*, 28(5), 531-547.
- Fernández, T., Harmony, T., Rodríguez, M., Bernal, J., Silva, J., Reyes, A., & Marosi, E. (1995). EEG activation patterns during the performance of tasks involving different components of mental calculation. *Electroencephalography and clinical neurophysiology*, 94(3), 175-182.
- Good, P. (2005). *Permutation, parametric and bootstrap tests of hypotheses*. Springer.
- Hanslmayr, S., Klimesch, W., Sauseng, P., Gruber, W., Doppelmayr, M., Freunberger, R., Pecherstrofer, T., & Birbaumer, N. (2006). Alpha phase reset contributes to the generation of ERPs. *Cerebral Cortex*, 17(1), 1-8.
- Hernández, W., Calderón-Córdova, C., González-Posada, V., Parra-Cerrada, Á., Jiménez, J. L., González-García, J. E., ... & Sergiyenko, O. Y. (2017). Bootstrap-based frequency estimation method. *Measurement*, 95, 193-200.
- Konietschke, F., & Pauly, M. (2014). Bootstrapping and permuting paired t-test type statistics. *Statistics and Computing*, 24(3), 283-296.
- Kubo, K., & Nittono, H. (2009). The role

- of intention to conceal in the P300-based concealed information test. *Applied Psychophysiology and Biofeedback*, 34(3), 227-235.
- Lefebvre, C. D., Marchand, Y., Smith, S. M., & Connolly, J. F. (2009). Use of event-related brain potentials (ERPs) to assess eyewitness accuracy and deception. *International Journal of Psychophysiology*, 73(3), 218-225.
- Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.
- Lukács, G., Grządziel, A., Kempkes, M., & Ansorge, U. (2019). Item Roles Explored in a Modified P300-Based CTP Concealed Information Test. *Applied psychophysiology and biofeedback*, 44(3), 195-209.
- Lukács, G., Weiss, B., Dalos, V. D., Kilencz, T., Tudja, S., & Csifcsák, G. (2016). The first independent study on the complex trial protocol version of the P300-based concealed information test: Corroboration of previous findings and highlights on vulnerabilities. *International Journal of Psychophysiology*, 110, 56-65.
- Marchand, Y., Inglis-Assaff, P. C., & Lefebvre, C. D. (2013). Impact of stimulus similarity between the probe and the irrelevant items during a card-playing deception detection task: The “irrelevants” are not irrelevant. *Journal of clinical and experimental neuropsychology*, 35(7), 686-701.
- Mathewson, K. J., Hashemi, A., Sheng, B., Sekuler, A. B., Bennett, P. J., & Schmidt, L. A. (2015). Regional electroencephalogram (EEG) alpha power and asymmetry in older adults: a study of short-term test-retest reliability. *Frontiers in aging neuroscience*, 7, 177.
- Matsuda, I., Nittono, H., & Ogawa, T. (2011). Event related potentials increase the discrimination performance of the autonomic based concealed information test. *Psychophysiology*, 48(12), 1701-1710.
- Meijer, E. H., Smulders, F. T., Merckelbach, H. L., & Wolf, A. G. (2007). The P300 is sensitive to concealed face recognition. *International Journal of Psychophysiology*, 66(3), 231-237.
- Meixner, J. B., & Rosenfeld, J. P. (2014). Detecting knowledge of incidentally acquired, real-world memories using a P300-based concealed-information test. *Psychological science*, 25(11), 1994-2005.
- Mertens, R., & Allen, J. J. (2008). The role of psychophysiology in forensic assessments: Deception detection, ERPs, and virtual reality mock crime scenarios. *Psychophysiology*, 45(2), 286-298.
- Rosenfeld, J. P., & Donchin, E. (2015). Resampling (bootstrapping) the mean: A definite do. *Psychophysiology*, 52(7), 969-972.
- Rosenfeld, J. P., Angell, A., Johnson, M., & Qian, J. (1991). An ERP-based control-question lie detector analog: Algorithms for discriminating effects within individuals's average waveforms. *Psychophysiology*, 28(3), 319-335.
- Rosenfeld, J. P., Davydova, E., Labkovsky, E., & Ward, A. C. (2019). Financial Incentive does not affect P300 in the Complex Trial Protocol (CTP) version of the Concealed Information Test (CIT) in malingering detection. II. Uninstructed subjects. *Frontiers in psychiatry*,

- 10, 189.
- Rosenfeld, J. P., Labkovsky, E., Winograd, M., Lui, M. A., Vandenboom, C., & Chedid, E. (2008). The Complex Trial Protocol (CTP): A new, countermeasure resistant, accurate, P300 based method for detection of concealed information. *Psychophysiology*, 45(6), 906-919.
- Rosenfeld, J. P., Shue, E., & Singer, E. (2007). Single versus multiple probe blocks of P300-based concealed information tests for self-referring versus incidentally obtained information. *Biological psychology*, 74(3), 396-404.
- Rosenfeld, J. P., Soskins, M., Bosh, G., & Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology*, 41(2), 205-219.
- Rosenfeld, J. P., Sweet, J. J., Chuang, J., Ellwanger, J., & Song, L. (1996). Detection of simulated malingering using forced choice recognition enhanced with event-related potential recording. *The Clinical Neuropsychologist*, 10(2), 163-179.
- Semlitsch, H. V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology*, 23(6), 695-703.
- Verleger, R. (1993). Valid identification of blink artefacts: are they larger than 50 μ V in EEG records?. *Electroencephalography and clinical Neurophysiology*, 87(6), 354-363.
- Verschuere, B., Rosenfeld, J. P., Winograd, M. R., Labkovsky, E., & Wiersema, R. (2009). The role of deception in P300 memory detection. *Legal and Criminological Psychology*, 14(2), 253-262.

1 차원고접수 : 2019. 10. 06.

심사통과접수 : 2020. 03. 11.

최종원고접수 : 2020. 03. 12.

Is the *t*-test insensitive than the bootstrap method in the P300-based concealed information test?

Jin-sup Eom¹⁾

Jin-Hun Sohn²⁾

Mi-Sook Park³⁾

1)Department of Psychology, Chungbuk National University

2)Department of Psychology, Chungnam National University

3)Department of Rehabilitation Counseling Psychology, Seoul Hanyoung University

In P300-based concealed information test (P300 CIT), it evaluates whether the P300 amplitude for the probe is significantly greater than that of the irrelevant to determine if the suspect is telling a lie. An independent sample *t*-test or a bootstrap method can be used as a statistical test to make that decision. Rosenfeld et al. (2004) used the bootstrap method, claiming that “*t* tests on single sweeps are too insensitive to use to compare mean probe and irrelevant P300s within individuals” and their method has been accepted to date. The purpose of the study is to evaluate whether the power of *t*-test is lower than that of the bootstrap method in the P300 CIT. The Monte Carlo study was conducted by using EEG collected from 39 participants. The results showed that the type I error rates of the *t*-test and the percentile bootstrap method were similar and the power of the percentile bootstrap method was slightly higher than that of the *t*-test. The type I error rates of the *t*-test and the percentile bootstrap method were slightly lower than the significance level and the powers of the two tests were also slightly lower than that of the theoretical *t*-test. On the other hand, the type I error rate and power of the standard error Bootstrap method were the same as those of the theoretical *t*-test and its power was .012 ~ .081 higher than that of *t*-test depending on experimental conditions.

Key words : P300, concealed information test, *t*-test, bootstrap