

거짓말은 손가락도 멈추게 한다*

정 호 진¹⁾

최 훈^{1,2)†}

¹⁾한림대학교 심리학과

²⁾한림 응용심리 연구소

범죄자들의 거짓말, 예를 들어 허위진술을 변별하는 문제는 범죄 수사에서 중요한 과업 중 하나이다. 거짓말은 거짓 현실을 생성하고, 거짓이 탄로되지 않도록 감시하는 등 인지부담을 유발하는 것으로 알려져 있다. 거짓말의 인지 부담이 증가한다는 것은 뇌파, fMRI와 같은 신경학적 기법 외에도, 2014년 이후부터 키보드 스트로크(키보드 입력 패턴) 분석법을 통해서도 확인되고 있다. 키보드 스트로크를 이용한 거짓말 탐지는 검사의 제약이 적고, 진술 내용의 질적분석도 가능하다는 이점이 있으나 국내에서는 시도되지 않은 방법이다. 본 연구는 한국인을 대상으로 거짓말을 할 때 어떤 키보드 스트로크 지표가 변화하는지 조사하였다. 온라인 실험에 참여한 참가자(N=38)들을 대상으로 일기 작성 과제를 기록된 키보드 스트로크와, 범행 후 뒷처리를 암시하는 동영상 내용을 회상하는 동안 기록된 키보드 스트로크를 비교했다. 목격 집단(n=22)은 목격한 내용을 솔직하게 작성해달라는 안내문을, 거짓집단(n=16)에는 가해자의 범행이 없었던 것처럼 내용을 꾸며서 작성해달라는 안내문을 제공했다. 분석 결과, 목격집단은 기저선과 자유회상 과제에서 키보드 스트로크 지표의 차이가 유의하지 않았으나 거짓집단은 초당 키입력 횟수, 첫 입력 시간, 정지시간에서 유의한 변화가 관찰됐다. 구체적으로, 거짓말을 해야하는 자유회상 과제에서 초당 키입력 횟수와 첫 입력 시간이 느려지고, 정지 시간이 길어졌다. 이 결과는 허위 진술을 작성할 때 인지부담의 증가로 키보드 입력 행동이 변화한다는 것을 의미하며, 키보드 스트로크가 효과적인 거짓말탐지 지표가 될 수 있음을 시사한다.

주요어 : 거짓말탐지, 키보드 스트로크, 인지 부담

* 이 연구는 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2022S1A5B5A17048989).

† 교신저자: 최 훈, 한림 응용심리 연구소, (24252) 강원도 춘천시 한림대학길 1
E-mail: hoonchoi@hallym.ac.kr

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. Copyright ©2022, The Korean Association of Psychology and Law

범죄 수사에서 목격자들의 진술은 중요한 정보원이며, 특히 물적 증거가 확보되지 않았을 때는 진술 증거의 가치가 더 높아진다. 하지만 때때로 가해자가 수사 방해 혹은 지연을 목적으로 목격자 행세를 하거나 의도적으로 거짓말을 할 수 있다. 이런 상황에서 수사 초기부터 사용할 수 있는 객관적인 거짓말탐지기법이나 평가 도구를 활용할 수 있다면, 사건 초기부터 타당한 진술 자료를 수집하여 효과적이고 신속한 수사를 통해 사법정의 실현에 기여할 것이다.

범죄 가해자가 거짓말을 하는 이유는 범행 사실을 숨김으로써 처벌을 피하려는 동기가 크게 작용한다(Miller & Stiff, 1993). 즉, 거짓말은 처벌을 피하거나 이득을 위해 상대방을 속이는 대처 전략 중 하나이다. 하지만, 대부분의 사법장면에서 거짓말이 탄로나는 경우 더 큰 처벌을 받게 된다. 따라서, 거짓말을 하는 사람은 상대방을 성공적으로 속이기 위해 인지적으로 복잡한 과업을 수행해야 한다. 거짓말의 인지 모델(Vrij, Granhag, Mann, & Leal, 2011)에 따르면, 거짓말은 진실을 말할 때보다 더 큰 인지 부담을 유발한다. 먼저, 거짓말을 해야 한다는 상황 자체가 인지적 부담을 야기할 수 있다. 진실처럼 보이는 대안적 현실을 만들어내는 것도 쉬운 일이 아니다. 그리고 자신이 만들어낸 진술들을 지속적으로 검토하면서 서로 상충되는 부분이 없는지 확인해야 한다. 또한, 거짓말이 탄로날 여지를 남기지 않기 위해서는 진실과 근접한 내용을 진술하지 않도록 억제하는 과정도 필요하다.

이와 같이 거짓말에는 부수적인 인지 과제들이 포함되기 때문에 진실을 말할 때와 거짓말을 할 때의 인지과정이 다를 것이며, 따라서 다른 언어적/행동적 반응이 나타날 것이라

고 예상할 수 있다(Vrij, Hartwig, & Granhag, 2019). 예를 들어, DePaulo 등(2003)은 메타분석을 통해 거짓말을 할 때 변화를 보이는 언어적, 비언어적 지표로 언어행동의 지연이나 목소리 톤, 세부 내용의 양 등이 있음을 확인하였다. 현재 수사기관 및 연구장면에서 진술 신빙성을 평가하거나, 거짓말을 탐지하는 방법도 거짓말과 진실의 인지/정서적 처리 과정이 다를 것이라는 원리에 근거한다(김영운, 2015). 대표적으로, 거짓말탐지기로 알려진 폴리그래프(polygraph)나 뇌파(electroencephalogram)와 같은 신경생리반응을 분석하는 방법이 있다. 이 방법들은 ‘탄로에 대한 두려움’을 경험할 때 발생한다고 알려진 특징적인 생리 반응(예, 심박수, 호흡)의 변화를 조사하거나, 용의자만 특정적으로 알고 있는 범죄관련사항이 있는지 확인하는 숨김정보검사를 실시하는 동안 발생하는 신경학적 반응(예, 뇌파, fMRI)을 분석한다(홍현기 등, 2014; Rosenfeld, Hu, Labkovsky, Meixner, & Winograd, 2013). 메타연구에 따르면 뇌파나 fMRI를 분석했을 때 분류 정확률은 80% 이상으로 폴리그래프보다 높은 편에 속한다(Meijer & Verschuere, 2017). 그러나 이 방법들은 신뢰할 수 있는 신경생리반응의 측정을 위해 통제된 검사 환경이 필수적이고, 수검자의 움직임을 제약하는 등 자율성을 제약하는 요소가 있다는 단점이 있다. 무엇보다도 용의자에게 정서적 각성을 유발하는 범죄 관련 단서가 필요할 때가 있고, 신경생리자극을 유발하는 질문의 선정이나 신경생리 반응의 결과 분석 및 해석에 검사관의 경험과 주관의 반영되어 객관성을 담보하기 어렵다는 한계가 있다. 따라서, 검사 실시 환경에 제약이 적고 간편하며, 객관적인 결과가 산출 가능한 거짓말탐지 방법이 필요할 것이다. 이에

따라, 최근 들어서는 키보드로 언어 자료를 입력할 때 기록되는 키보드 스트로크를 거짓말탐지에 활용하려는 시도가 진행되고 있다 (Banerjee, Feng, Kang, & Choi, 2014; Monaro et al., 2018; Tomas, Tsimperidis, Demarchi, & El Massioui, 2020).

키보드 스트로크 분석은 응답자의 키보드 입력 패턴을 분석하여 개인의 신원을 확인하거나, 심리상태를 파악하는 것을 의미한다. 사람마다 키보드를 누를 때 고유한 입력 패턴을 보인다(Leggett & Williams, 1988). 따라서, 키보드 스트로크는 도입 초기에 행동적 생체인증 수단(behavioral biometric)으로 주목받았다 (Monrose & Rubin, 2000). 그러다가 2010년도 이후부터 정서 상태, 인지 부담 등 개인의 심리상태 측정 방법으로도 연구되기 시작했다 (Nahin, Alam, Mahmud, & Hasan, 2014). 예를 들어, Epp, Lippold와 Mandryk (2011)의 연구에서 참가자들은 화남, 지루함, 스트레스 등 정서와 연관된 문장을 작성하는 동안 정서에 따라 키입력 수, 지속시간, 키입력 간 간격 등의 키보드 스트로크의 특징에서 차이를 보였다. 키보드로 단어나 문장을 작성할 때의 입력 패턴은 언어 행동의 지표로도 사용되었는데, 예를 들어 난이도 높은 문제를 푸는 경우 키보드 입력 시간이 느려지거나 작성 중 멈춤(pause) 시간이 길어졌다(Zulkifli, 2013).

키보드 스트로크와 인지정보 처리과정의 연관성은 거짓말에 대한 연구에서도 발견되고 있다(Banerjee et al., 2014). 이 연구에서 참가자들은 작문과제를 수행했는데, 거짓집단은 식당의 평가나 사회 현상에 대해 자신의 의견과 반대되는 글을 작성하도록 요청받았다. 실험 후 목격집단과 거짓집단의 키보드 스트로크를 분석한 결과, 거짓집단의 명사, 기능어 입력

시간은 목격집단보다 유의하게 느렸다. 신분 위조 식별(Monaro et al., 2018), 허위진술 탐지 (Tomas et al., 2020) 연구에서도 거짓집단은 오반응이 증가하고 응답시간이 지연된 행동 특성을 보였다. 허위신원을 보고하도록 사전에 훈련받은 사람들도 예상치 못한 질문에 대한 답변 작성 시간은 목격집단보다 유의하게 길었다(Monaro et al., 2018). 거짓말의 인지 모델에 따르면, 예상치 못한 질문은 인지부담을 증가시키는 처치 기법이기 때문에, 거짓말 집단이 보여준 예상치 못한 질문에서 응답 시간의 지연은 거짓말의 인지 모델에서 예측하는 것과 일치한다(Vrij, Fisher, & Blank, 2017).

종합해보면, 키보드 스트로크는 거짓말 동안 증가하는 인지부담을 타당하게 측정할 수 있는 언어 행동 지표로 고려되고 있다(Tomas, Dodier, & Demarchi, 2022). 키보드 스트로크는 언어 과정을 매개하여 측정이 이뤄지기 때문에 검사의 자유도가 높다는 강점도 있다. 예를 들어, 폴리그래프와 같이 ‘예’ 혹은 ‘아니오’와 같은 단답형 응답을 분석해서 거짓 여부를 파악하거나(Monaro et al., 2018), 자유회상 형식으로 장문의 언어자료를 수집해서 분석하는 것도 가능하다(Tomas et al., 2020). 이 경우 수집된 언어자료를 활용하여 진술 신빙성을 검증하는 진술분석도 실시할 수 있다. 자유회상 형식의 키보드 스트로크 분석 접근법은 질적 분석이 가능한 진술문을 제공하면서 동시에 양적 분석이 가능한 행동적 지표를 동시에 제공해준다는 점에서 진술신빙성 검증을 다각도로 접근하는 새로운 연구 도구로 활용할 수 있을 것으로 기대한다. 이에 더해, 최근 소셜네트워크 서비스(SNS)나 인터넷 플랫폼에서 증가하고 있는 언택트 범죄(예, “언택트로 진화하는 사기수법 ...전담부서 신설 목소리

커진다”, 2021)의 예방 대책 마련에도 활용할 수 있을 것이다.

앞서 살펴봤듯이 해외에서 수행된 연구들은 키보드 스트로크가 인지부담의 행동적 지표이며, 허위 신원식별과 거짓말탐지에 효과적이라는 것을 보여주었다. 그러나, 국내 법심리학계에서는 키보드 스트로크를 활용한 연구가 전무한 실정이다. 기존의 연구들은 영어권에서 수행되어 한국어의 거짓말 관련 키보드 작성 행동을 적절히 반영하지 못할 수 있기에 한국인 대상 거짓말 관련 키보드 스트로크 지표를 확인하는 기초 연구가 필요하다. 예를 들어, Vrij와 Mann(2001)의 연구에서는 거짓집단이 목격집단에 비해 말을 멈추는 빈도가 높았으나, 국내 대학생 표본에서는 거짓집단과 목격집단의 말 멈춤 빈도에 유의한 차이가 나타나지 않았다는 것(김시업, 전우병, 김경하, 김미영, 전충현, 2005)은 해외 연구 결과가 한국인에게 일반화되지 않을 수 있음을 보여준다. 따라서 본 연구는 한국인이 거짓말을 할 때 키보드 스트로크 특성에서 어떤 변화가 발생하는지 살펴보고자 한다. 거짓말의 인지이론에 기초하면, 직접 경험한 기억으로 구성된 진술문과 거짓 내용을 포함하거나 과장된 진술문의 작성하는 것의 키보드 입력 패턴이 다를 것이라 예상된다. 거짓진술을 작성할 때 인지부담이 증가한다면, 진술문 작성에 소요되는 시간이 증가하거나 초당 키입력 횟수가 감소할 것이다. 이 외에도 키 입력을 잘못하는 빈도가 증가하여 수정을 위한 삭제(delete, backspace) 버튼을 누르는 횟수가 증가할 것이다.

연구 방법

참가자

본 연구는 COVID-19로 인한 대면접촉을 최소화하기 위해 온라인으로 수행되었다. 연구 대상자들은 대학교의 온라인 커뮤니티에서 ‘일상적인 사건과 그렇지 않은 사건의 처리과정을 비교하는 연구’ 모집 공고 글을 읽고 자발적인 참여 의사를 밝힌 사람들이었다. 참가자들에게는 서술형 문제와 계산 문제가 일부 포함되어 있다고 사전 안내했으며, 주어진 질문에 대한 답변을 키보드로 작성해서 제출해달라는 안내문구를 제공하였다. 모든 참가자들은 실험 참여에 대한 보상으로 2천원 상당의 상품권을 제공받았다.

총 45명(남자 6명, 여자 39명)이 참여하였으며, 이들의 평균 연령은 27.40세(SD=8.07)였다. 이들 중 검사 단계에서 1분 이상 키보드 입력이 없었거나 기술적 문제로 키보드 스트로크 기록이 되지 않았던 7명의 자료를 제외하여 최종적으로 38명(남자 6명, 여자 32명)의 자료를 분석했다. 분석 대상자의 평균 연령은 27.97세(SD=7.66)였다.

실험 자극

본 연구는 목격자 기억의 보호 수단을 조사한 선행 연구(함근수, 김예슬, 김기평, 정호진, 2020)에서 사용한 1분 54초 길이의 동영상(<https://youtu.be/TeU5HZLNpFA>)을 후속적으로 수행되는 기억 검사를 위한 부호화 자극으로 사용하였다. 범죄 동영상 부호화 단계 동안 참가자들에게는 ‘동영상을 주의깊게 봐주세요.’라는 안내문을 제공하였다.

이 동영상은 남성과 여성이 말다툼을 하다가 침대에서 남성이 여성을 밀친 후, 남성이

바닥의 피를 닦고 빨래를 돌리고 피가 묻은 칼을 씻는 등 살인사건의 뒷수습을 암시하는 장면으로 구성되어 있다. 하지만, 남성이 직접 여성을 칼로 찌르거나 위협을 가하는 장면은 없었다.

실험 설계

본 연구는 목격집단과 거짓집단으로 구분된 피험자 간 설계를 사용하였다. 목격집단 22명(남자 3명, 여자 19명), 거짓집단 16명(남자 3명, 여자 13명)의 할당은 실험이 시작되면 무선적으로 이뤄졌다. 목격집단(M=29.9, SD=9.00)과 거짓집단(M=26.5, SD=6.47) 응답자들의 나이는 통계적으로 유의한 차이를 보이지 않았다, $t(36)=1.29, p=.206, \text{Cohen's } d=.433$. 두 집단의 참가자들은 기저선 측정, 범죄 동영상 시청, 간섭과제 수행, 진술문 작성의 절차를 순서대로 수행하였다. 기저선 측정은 거짓말을 할 때 키보드 스트로크가 변화하는지 알아보기 위해 이뤄졌다. 이 단계에서 참가자들은 실험 처치 전 자신의 오늘 하루 경험한 일들을 키보드로 작성하도록 요청받았다.

중속 측정치로 사용된 키보드 스트로크는

PsychoPy3(Peirce et al., 2019)를 통해 기록된 자료를 전처리하여 사용했다. 구체적으로, 선행연구(Monaro et al., 2018; Tomas et al., 2020)를 참고하여 기저선(참가자들이 실험 당일 경험한 일을 회상)과 범죄관련내용을 작성할 때 초당 키 입력 횟수, 정지 횟수, 정지 시간, 삭제 버튼(delete, backspace) 빈도, 첫 입력 시간, 총 작성 시간의 자료를 수집하였다. 각 키보드 스트로크 지표의 설명은 표 1에 정리되어 있다.

실험 절차

본 연구의 실험 절차는 선행연구(Tomas et al., 2020)를 참고하여 설계했다(그림 1). 참가자들에게 조용한 환경에서 PsychoPy3(Peirce et al., 2019)로 제작한 실험 프로그램이 구동되는 온라인 링크를 클릭하도록 안내했다.

실험이 시작되면 참가자들에게 안내문이 가장 먼저 제공되었다. 안내문에는 본 연구가 일상적인 사건과 그렇지 않은 사건의 기억 인출과정을 알아보려는 목적으로 수행되며, 약 15분 정도 소요된다는 내용이 포함되어 있었다. 또한, 서술형 문항에 최소 100글자 이상 작성해야 한다는 것과 5분 정도 진행되는 계

표 1. 본 연구에서 분석에 사용한 키보드 스트로크 지표 및 정의

키보드 스트로크 지표(단위)	정의
초당 키입력 횟수(입력 수)	1초에 입력하는 키입력 횟수의 평균. 전체 키입력 횟수/초로 계산.
정지 횟수(빈도)	0.5초 이상 키 입력이 없을 때, 정지로 간주.
정지 시간(초)	정지 후 키보드 입력까지 소요되는 시간.
삭제 버튼 빈도(빈도)	작성한 글을 수정하기 위해 삭제 버튼(delete, backspace)을 누른 빈도.
첫 입력 시간(초)	기억 검사가 시작하고 처음으로 키보드를 입력하는 데까지 소요되는 시간
총 작성 시간(초)	진술문 작성에 소요되는 총 시간.

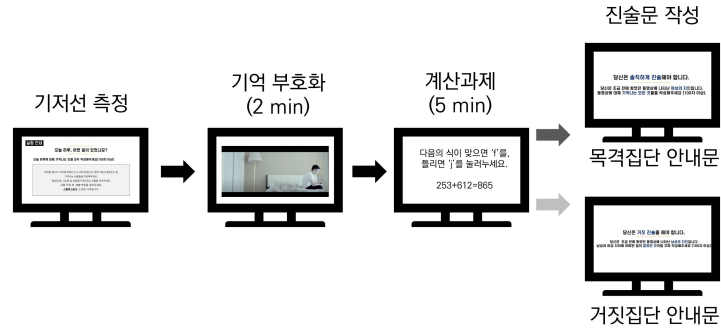


그림 1. 본 연구의 실험 절차

산문제가 있음을 안내하였다. 참가자들은 안내문을 읽고 연구 참여에 동의할 경우 ‘동의합니다.’를 키보드로 작성했으며, 동의하지 않은 경우 ‘esc’ 버튼을 누름으로써 실험을 종료할 수 있었다.

실험의 첫 단계는 기저선의 키보드 스트로크를 측정하는 것이었다. 기저선 측정 전에 참가자들에게 ‘오늘 하루에 대해 기억나는 모든 것을 작성해주세요.’라는 지시문을 제공했다. 구체적으로 무엇을 했는지, 어디에 있었는지, 누구와 만났는지 등 기억나는 모든 것들을 적도록 안내했다. 참가자들은 지시문을 이해한 후 스페이스바를 누르면 기저선 측정 단계로 넘어갔으며, 이때부터 기저선 키보드 스트로크를 기록하였다. 참가자들은 실험 당일 경험했던 일화적 사건들을 키보드로 입력한 후 마우스로 ‘제출’ 버튼을 누르면 다음 과정으로 진행할 수 있었다. 100자 미만의 서술문을 작성한 경우 제출이 되지 않았으며, ‘글자수가 부족합니다. 조금 더 작성해주세요.’라는 안내 문구를 화면 상단에 표시했다.

기저선 측정 절차를 마치고 나면 범죄 사건의 부호화를 위해 1분 54초 분량의 동영상 재생되었다. 화면 상단에는 동영상을 집중해서 봐야 한다는 문구도 제공되었다. 참가자들

은 동영상이 종료 후 계산 과제를 수행하였다. 계산 과제는 기억 정보의 즉각적인 인출을 차단하는 것 외에도, 참가자들이 사후 기억검사를 예측하고 대비차원에서 동영상의 내용을 의도적으로 시연하는 것을 방지하는 것을 목적으로 포함되었다(Greene, Maloney-Derham, & Mulligan, 2020). 계산 과제는 총 110문제로 구성되었으나, 몇 문제를 풀었는지와는 무관하게 5분이 지나면 종료되었다. 이후 기억 검사 단계로 넘어갔다. 참가자들은 덧셈 혹은 뺄셈 계산식(예, $3+8=11$, $278-116=168$)이 맞는지를 판단하여 버튼을 눌렀으며, 정답 여부에 따른 피드백(‘정답입니다’, ‘오답입니다’)을 제공하였다. 본 연구에서 목적집단과 거짓집단의 계산 과제 정확률은 목적집단($M=.93$, $SD=.10$), 거짓집단($M=.96$, $SD=.05$) 모두 90% 이상이었으며, 통계적으로 유의한 집단 차이는 발견되지 않았다, $t(36)=1.03$, $p=.311$, Cohen’s $d=0.354$.

기억검사 직전에 목적집단과 거짓집단은 서로 다른 안내문을 제공받았다. 두 집단 모두 공통적으로 여성이 남성의 집에 방문한 후 실종되어 수사가 진행 중이며, 경찰이 당신을 참고인으로 소환하여 조사할 것이라고 설명했다. 목적집단의 참가자들은 동영상에서 피해자로 묘사된 여성의 지인으로, 동영상에 대해

기억나는 모든 것을 솔직하게 작성해야 한다는 안내문을 추가적으로 제공하였다. 반면, 거짓집단의 참가자들은 가해자로 묘사된 남성의 지인으로, 남성과 여성 사이에 어떠한 일이 없었던 것처럼 꾸며서 거짓 진술을 해야 한다는 안내문을 제공하였다. 안내문을 읽고 준비가 된 참가자들이 스페이스바를 누르면 기억검사가 시작됐으며, 이때부터 키보드 스트로크가 기록되었다. 기억검사도 기저선 단계와 마찬가지로 100자 이상의 내용을 작성해야 제출이 가능했다. 최종 자료를 제출한 참가자들에게는 디브리핑 문구를 제공하였다. 구체적으로, 본 연구가 거짓말을 할 때 변화하는 키보드 입력 패턴을 알아보기 위해 수행되었다고 안내했다. 또한, 실험과 관련된 질문은 자유롭게 할 수 있도록 연구자의 이메일을 제공하였다.

통계 분석

모든 통계분석은 R 4.1.1(R Core Team, 2021)을 이용하여 수행되었다. 거짓 진술문을 작성할 때 키보드 스트로크가 변화하는지 알아보기 위해 2(집단: 거짓집단, 목격집단)×2(검사 유형: 기저선, 기억검사) 혼합 분산분석(mixed ANOVA)을 실시하였다. 여기서 집단은 피험자 간 변인, 검사유형은 피험자 내 변인으로 포함되었다. 본 연구는 거짓말을 할 때 증가하는 키보드 스트로크 특징을 알아보는데 관심이 있으므로, 집단×검사유형 상호작용 효과를 중점적으로 살펴보았다.

결 과

조작 점검

분석을 실시하기 앞서 참가자들이 작성한 내용을 검토했을 때 기억검사에서 지시와 불일치하거나 불성실한 응답은 관찰되지 않았다. 즉, 목격집단의 참가자들은 동영상의 내용을 작성했으며, 거짓집단의 참가자들은 범행 장면과 일치하지 않은 진술문들을 작성했다. 목격집단과 거짓집단의 참가자가 작성한 문장의 예시는 아래와 같다.

- 목격집단: “남자와 여자가 다뤘다. 남자가 급하게 피로 묻은 방을 정리했다. 세면대와 바닥에는 피로 흥건했다. 마지막에 남자는 검은색 큰 쓰레기봉투 2개를 버리러 가지고 나갔다. 남자가 여자를 살인한 것으로 추정된다.”
- 거짓집단: “집에 도착해서 같이 점심 준비를 해서 점심을 먹고, 영화를 한 편 보다가 상큼하게 먹고 싶어서 과일을 먹었다. 영화가 재미없어서 끄고, 내기를 걸고 게임을 진행했다. 게임을 저서 빠진 여자 친구는 집으로 돌아갔다.”

본 연구에서 기저선 과제는 ‘오늘 하루 있었던 일들’을, 기억 검사는 ‘범죄 동영상’을 회상하는 것이었다. 검사유형에 따른 전반적인 인출 난이도나 정보 인출량에서 차이가 있을 수 있으므로, 키보드 스트로크 분석을 실시하기 전 집단, 검사 유형 별로 정보 인출량(글자, 단어, 문장 수)을 계산하여 살펴보았다(표 2).

전반적으로 목격집단과 거짓집단 모두 기저선과 기억검사에서 작성한 글의 분량은 비슷한 것으로 나타났다. 목격집단이 거짓집단보

표 2. 각 집단별 검사 유형에 따른 정보 인출량

		글자 수	단어 수	문장 수
목격집단 (n=22)	기저선	234.625(127.786)	56.625(31.623)	5.062(2.999)
	기억검사	225.812(192.651)	55.875(49.777)	5.125(2.986)
거짓집단 (n=16)	기저선	195.045(97.528)	47.818(24.974)	5.455(2.923)
	기억검사	189.955(90.772)	47.273(22.987)	4.364(2.237)

다 많은 양의 정보를 입력한 것으로 보이나, 진술문에서 입력된 글자, 단어, 문장 수에서 통계적으로 유의한 집단차이는 발견되지 않았다, 글자 수: $F(1,36)=1.12, p=.297$; 단어 수: $F(1,36)=0.92, p=.344$; 문장 수: $F(1,36)=0.06, p=.807$. 이 외에 글자 수, 단어 수, 문장 수 모두 통계적으로 유의한 검사 유형의 주효과나 검사 유형×집단 상호작용 효과는 발견되지 않았다($p>.10$). 이는 실험 처치 측면에서는 기저선과 기억검사가 다소 차이가 있을 수 있으나, 과제 유형에 따라 참가자들의 작성량에는 차이가 발생하지 않았다는 것을 보여준다.

키보드 스트로크 분석 결과

조건 별 참가자들의 키보드 스트로크 측정치들의 결과가 그림 2에 제시되어 있다.

이원 혼합 분산분석 결과 요약은 표 3에 정리되어 있다. 통계적으로 유의한 집단×검사유형 상호작용 효과를 보인 키보드 스트로크 측정치를 살펴본 결과, 초당 키보드 입력횟수, 정지 시간, 첫 입력 시간이 집단에 따라 기저선과 회상검사에서의 차이가 다른 것으로 나타났다, 초당 키보드 입력횟수: $F(1,36)=4.54, p=.040, \eta_p^2=.112$, 정지 시간: $F(1,36)=4.52, p=.040, \eta_p^2=.112$, 첫 입력 시간:

$F(1,36)=6.59, p=.015, \eta_p^2=.155$.

초당 키보드 입력횟수의 상호작용 효과를 살펴본 결과, 목격집단은 기저선($M=5.06, SD=1.64$)과 회상검사($M=4.99, SD=1.94$)를 실시할 때 초당 키보드 입력 횟수에서 차이가 없었다, $t(21)=.404, p=.690, Cohens'd=.086$. 그러나, 거짓집단은 기저선($M=5.19, SD=1.37$)보다 거짓말을 해야 하는 회상검사($M=4.48, SD=1.63$)에서 초당 키보드 입력 횟수가 유의하게 느려졌다, $t(15)=2.94, p=.010, Cohens'd=.734$. 정지 시간의 경우 거짓집단은 기저선($M=1.46, SD=0.51$)보다 회상검사($M=1.89, SD=0.73$)에서 유의하게 증가한 반면, $t(15)=-2.30, p=.036, Cohens'd=-.574$, 목격집단은 기저선($M=1.54, SD=0.53$)과 회상검사($M=1.55, SD=0.46$)의 차이가 유의하지 않았다, $t(22)=-0.11, p=.911, Cohens'd=-.024$. 첫 입력 시간의 경우 목격집단은 기저선($M=2.89, SD=3.19$)과 회상검사($M=2.74, SD=2.58$)를 실시할 때 차이가 없었으나, $t(21)=.19, p=.851, Cohens'd=.041$, 거짓집단은 기저선($M=1.97, SD=1.39$)보다 회상검사($M=5.03, SD=3.97$)의 첫 키보드 입력 시간이 유의하게 길었다, $t(15)=-3.16, p=.006, Cohens'd=-.790$.

검사유형을 구분하여 집단 간 차이를 비교했을 때 세 키보드 측정치 모두 기저선과 자유회상 모두 유의한 집단 차이가 발견되지 않

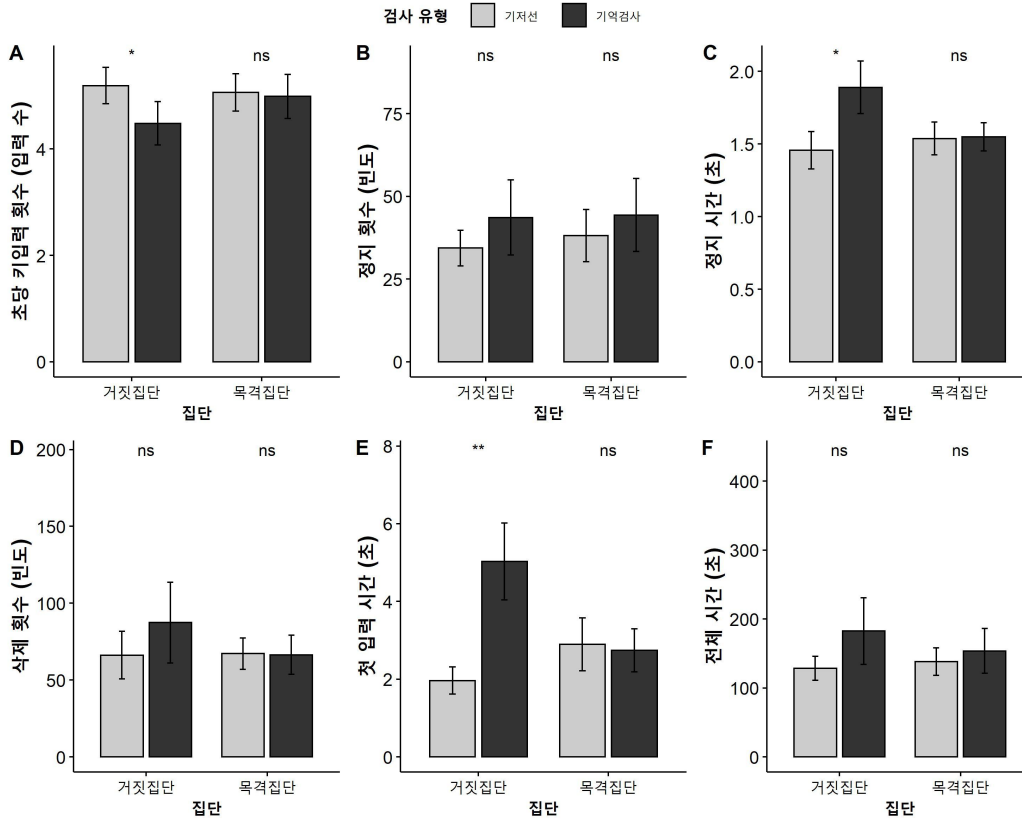


그림 2. 목격집단과 거짓집단의 기저선, 기억검사에서 키보드 스트로크 측정치

(* $p < .05$, ** $p < .01$).

표 3. 키보드 스트로크 측정치의 집단×검사유형 혼합 분산분석 결과

종속 측정치	변산원	df	MSS(MSE)	F	p	η_p^2
초당 키 입력 횟수 (keystroke per second)	집단(A)	1	0.66(5.24)	0.13	.725	.003
	검사유형(B)	1	2.85(.41)	6.91	.013	.161
	A×B	1	1.87(.41)	4.54	.040	.112
정지 횟수 (count of pause)	집단(A)	1	96(2746.91)	0.03	.853	<.001
	검사유형(B)	1	1103(659.81)	1.67	.204	.044
	A×B	1	44(659.81)	0.07	.799	.002
정지 시간 (average pause length)	집단(A)	1	0.31(0.43)	0.73	.400	.020
	검사유형(B)	1	0.91(0.18)	5.01	.031	.122
	A×B	1	0.82(0.18)	4.52	.040	.112

표 3. 키보드 스트로크 측정치의 집단×검사유형 혼합 분산분석 결과 (계속)

종속 측정치	변산원	df	MSS(MSE)	F	p	η_p^2
삭제 횟수 (count of remove keystrokes)	집단(A)	1	1856(7429.94)	0.25	.620	.007
	검사유형(B)	1	1963(2197.39)	0.89	.351	.024
	A×B	1	2250(2197.39)	1.02	.318	.028
첫 입력 시간 (first input time)	집단(A)	1	8.55(9.91)	0.86	.359	.023
	검사유형(B)	1	39.27(7.28)	5.40	.026	.130
	A×B	1	47.95(7.28)	6.59	.015	.155
전체 시간 (total time)	집단(A)	1	1689(24994.54)	0.07	.796	.002
	검사유형(B)	1	22486(11154.79)	2.02	.164	.053
	A×B	1	6805(11154.79)	0.61	.440	.017

았다. 다만, 첫 키보드 입력 시간의 경우 회상 검사에서 집단 차이의 경향성이 관찰되었다, $t(36)=2.02, p=.055$.

논 의

본 연구의 목적은 한국인을 대상으로 거짓말을 할 때 변화하는 키보드 입력 패턴 특징을 탐색적으로 알아보는 데 있었다. 이를 위해 목격한 동영상과 상충되는 내용을 작성할 때 키보드 스트로크가 기저선(하루 일화에 대한 기억)과 비교했을 때 차이가 있는지 분석하였다. 본 연구의 가장 중요한 결과는 거짓 집단은 거짓말을 할 때 키보드 입력 속도가 기저선에 비해 유의하게 감소한다는 것이다. 이 외에도 거짓말을 할 때 변화하는 키보드 입력 패턴의 변화로 질문을 보고 처음 답변을 작성할 때까지 소요되는 시간이 길어지거나, 작성 중 정지하는 시간이 길어지는 것이 있었다. 이는 거짓말을 만들어내는 과정에서 발생

한 인지 부담에 따른 변화로 추정되며(Vrij et al., 2017), 키보드 스트로크가 인지 부담의 측정치로 거짓말탐지 수단으로 활용될 수 있음을 시사한다.

거짓말의 인지부담 이론은 거짓말을 할 때 대안적 현실 생성, 지속적인 생각/행동 감시 등 복잡한 인지 과업이 요구된다고 가정한다(Vrij et al., 2011). 이러한 모델에 기초한 선행 연구들은 거짓말을 할 때 반응 속도가 감소하거나(Monaro, Zampieri, Sartori, Pietrini, & Orru, 2021), 키보드 입력 속도가 느려지거나 작성 시간이 길어진다는 것을 보여주었다(Banerjee et al., 2014; Monaro et al., 2018; Tomas et al., 2020). 본 연구의 결과도 거짓말을 할 때 인지 부담이 증가하며, 그 결과 키보드 입력 속도가 감소한다는 예측을 지지하는 것으로 나타났다. 거짓말을 할 때 키보드 입력 속도가 감소하는 것은 본 연구뿐 아니라 기존의 연구에서도 반복적으로 관찰되는 속성이다. 본 연구의 거짓집단에서 관찰된 첫 입력 시간의 지연은 거짓말 준비 시간을 제공받지 못한 상태에

표 4. 본 연구와 기존 연구의 비교

	본 연구(2022)	Monaro 등(2017)	Monaro 등(2018)	Tomas 등(2020)
검사 방법	자유회상	단답형 질문	단답형 질문	자유회상
거짓말 처치	목적 내용과 상충되는 내용을 작성	틀린 신원 정보 입력	사전에 훈련한 허위 신원 정보 입력	범죄 상황을 모사한 글을 보고 실제 목격한 것처럼 작성
실험설계	혼합설계	참가자 내 설계	참가자 간 설계	혼합설계
초당 키 입력 횟수	T>L	-	-	T>L
정지 횟수	n.s.	-	-	n.s.
정지 시간	T<L	-	-	n.s.
삭제 횟수	n.s.	-	-	n.s.
첫 입력 시간	T<L	T<L	T<L	-
전체 시간	n.s.	T<L	T<L	T>L

*T>L: 진실>거짓, T<L: 진실< 거짓, n.s: 유의한 차이 없음, -: 측정하지 않음

서 기억검사를 실시하는데서 오는 당혹감이 영향을 미쳤을 수 있다. 그러나, 사전에 허위 신원을 학습한 거짓집단도 통제집단보다 첫 문자를 입력하는 데 긴 시간이 걸렸다(Monaro et al., 2018). 그리고 첫 입력 시간의 지연은 다양한 실험 설계 환경에서 거짓 집단이 유의하게 느렸던 특징으로 나타났다(표 4 참고). 이는 거짓말을 ‘생성’할 때 발생하는 인지부담이 거짓말에 민감한 인지적 특성임을 시사한다. 이 경우 응답자에게 긴 문장을 생성하는 질문보다, 짧은 질문을 여러 번 제시하는 것이 효과적인 거짓말탐지 방법이 될 수 있을 것이다.

거짓말을 할 때 정지 발생 후 키보드 입력까지 소요되는 시간도 길어졌다. 이러한 특성은 거짓말을 할 때 ‘음’, ‘어’와 같은 말주저가 증가하는 것과 유사하다고 볼 수 있을 것이다(전우병, 김시업, 2006). 즉, 거짓말을 꾸며내다가 막히는 경우 다음 말을 머릿속으로 준비하

는 시간이 길어지면서 나타났을 가능성이 있다. 이는 타당한 거짓말을 생성하는 것이 부담되는 일이라고 주장하는 인지부담 이론과도 일치한다(Vrij et al., 2017). 거짓말 집단에서 첫 입력 시간이 증가하고, 멈춤 시간의 증가를 함께 고려해본다면, 거짓말 집단은 대안적 현실 생성에서 어려움을 겪었을 것으로 추정된다. 본 연구는 자유 회상 과제를 사용하면서 거짓말의 인지과정을 구분하는 것이 어려웠지만, 추후 연구에서 거짓말의 인지과정을 구분하여 질문을 제공했을 때 키보드 스트로크의 변화를 살펴본다면 거짓말의 인지적 속성을 깊이 있게 살펴볼 수 있을 것으로 기대된다.

반면, 정지 횟수, 삭제키 입력 횟수, 전체 입력 시간은 거짓말 여부에 따른 영향을 유의하게 받지 않는 것으로 나타났다. 이는 전체 입력 시간을 제외하고 Tomas 등(2020)의 결과와 일치한다(표 2). 전체입력 시간은 연구마다 상이한 결과를 보고하고 있다. 예를 들어,

Tomas 등(2020)의 연구에서는 거짓 내용을 작성할 때 오히려 시간이 짧았으나, Monaro 등(2018)의 연구에서는 거짓말을 할 때 검토 시간이 길어진다는 것을 발견했다. 이렇게 상충되는 결과는 검사의 속성에 기인할 수 있다. 단답형 질문의 검사를 할 때는 반응 검토가 빈번하게 이뤄지므로 전체 시간이나, 오류 횟수의 증가, 삭제키 등이 증가할 수 있지만(Monaro et al., 2018), 자유회상 형식으로 진술문을 작성할 때는 작성 시간이 길어짐에 따라 실험 외적 요인의 영향력이 커지게 된다. 다만, Tomas 등(2020)의 연구에서 거짓집단은 허위 진술해야 하는 범죄 상황을 묘사하는 글을 사전에 읽었고, 키보드 입력시 글자 수 제한이 없었다. 이로 인해 읽은 내용들에 대해 기억나는 것들을 간략히 작성하고 검사를 마무리해서 전체 시간이 짧게 소요됐을 가능성이 있다. 이와 같이, 전체 입력 시간은 검사 방법이나 거짓말의 내용에 따라 달라지므로 거짓말 탐지 지표로 사용하는 것을 신중하게 고려해야 할 것이다.

마지막으로, 본 연구 결과를 해석하고 활용할 때 염두해야 하는 몇 가지 제한점이 있다. 첫째, 거짓말의 처치가 연구자의 지시에만 의존했다는 것이다. 본 연구와 같이 실험자의 지시나 안내문을 통해 거짓말을 처치하는 연구도 있지만(Monaro et al., 2018; Tomas et al., 2020), 이 방법은 참가자들에게 거짓말을 해야 하는 동기를 적절히 유발하지 못할 수 있다. 추후 연구에서는 참가자가 사실을 은폐했을 때 잠재적인 이득을 얻는 것처럼 현실적인 거짓말 상황을 조작하고 이에 따른 키보드 스트로크의 변화를 조사할 필요가 있다(예, 홍현기 등, 2014). 둘째, 자신의 일화적 사건을 회상하는 기저선과 동영상의 기억을 회상하는 기억

검사가 동일하지 않았을 가능성이 있다. 작성한 글자, 단어, 문장 수에서는 차이가 없었지만 인출해야 하는 정보가 다르다는 것은 기저선과 기억검사를 비교하는 것을 어렵게 만드는 요인이다. 따라서, 추후 연구에서는 동일한 비교를 위해 기저선에서도 동영상에 대한 기억 인출을 할 때 키보드 스트로크를 측정하는 방법을 고려해볼 수 있을 것이다. 셋째, 온라인 실험에 따른 제약이 있다. 온라인 실험을 통해 비대면 연구가 가능했으나, 참가자들의 성비를 체계적으로 통제할 수 없었다. 그렇지만, 거짓말 관련 키보드 스트로크에서 성차가 존재할 가능성이 있다. 예를 들어, 대학생들 대상으로 진행된 연구에서 남성만 거짓말을 할 때 말 주저 빈도가 증가하는 것으로 나타났다(전우병, 김시업, 2006). 키보드 스트로크 관련 선행연구에서 성차가 보고된 것은 없으나 추후 연구에서는 성비를 통제하거나 독립변수화 하는 시도가 필요할 것이다. 추가적으로, 실험 조작과 참가자와의 라포 형성이 어렵다는 단점이 있다는 것도 문제점으로 지적될 수 있다. 폴리그래프 검사에서 라포를 형성하지 않은 집단의 판정 정확성이 저하된다는 점을 고려해본다면(김현지, 조은경, 2021), 라포 형성 절차가 생략된 온라인 실험의 타당성에 대한 지적이 제기될 수 있다. 따라서, 거짓말탐지 연구에서 검사 환경(예, 온라인, 실험실, 수사 관련 연구기관)에 따른 응답자들의 특성도 비교하는 것도 연구 결과의 일반화를 위한 흥미로운 연구 주제가 될 것이다.

이러한 제한점에도 불구하고 본 연구는 국내에서는 사용되지 않았던 키보드 스트로크를 분석하여 거짓말로 인한 인지부담 증가의 언어행동적 지표로 나타난다는 것을 확인했다는 데 의의가 있다. 키보드 스트로크는 내용 분

석에 의존하는 진술분석과 같은 기존 수사 기법에 객관적으로 검증 가능한 수량적 지표들을 제공하고, 뇌파와 같은 신경생리적 자료 수집과도 병행할 수 있다는 장점이 있다. 예를 들어, 거짓 진술을 작성할 때 인지부담의 신경생리적 지표(예, P300)가 동시에 증가하는지 확인한다면 키보드 스트로크 분석법의 타당도를 확인하면서, 검사 기법의 정확성도 비교할 수 있을 것이다. 수사 실무적인 측면에서는 키보드 스트로크 자료는 수사 초기에 간편하게 수집할 수 있으며, 엄격하게 통제된 검사 환경 구성이 어려운 환경(예, 입국심사, 위급상황에서 용의자 신원조회)에서 객관적 진술 자료 수집에 활용할 수 있을 것으로 기대한다.

참고문헌

김시업, 전우병, 김경하, 김미영, 전충현 (2005). 용의자의 거짓말탐지를 위한 비언어적 단서탐색. *한국심리학회지: 사회 및 성격*, 19(1), 151-162.

김영윤 (2015). 신경과학을 이용한 거짓말탐지. *교정담론*, 9(3), 87-125.

김현지, 조은경 (2021). 폴리그래프 검사에서 라포 형성에 따른 심리생리적 반응 차이. *한국심리학회지: 법*, 12, 53-73.

박홍용 (2021년 08월 01일). 언택트로 진화하는 사기수법 ...전담부서 신설 목소리 커진다. *서울경제*.
<https://www.sedaily.com/NewsView/22Q0QO6FJR>

전우병, 김시업 (2006). 남녀 대학생들이 거짓말시 나타내는 비언어적 행동단서의 차이. *한국심리학회지: 사회 및 성격*, 20(2), 57-

72.

함근수, 김예슬, 김기평, 정호진 (2020). 시간 지연과 오정보 제시 상황에서 초기 자기기입식 면담(SAI)이 정확 회상과 기억 보호에 미치는 영향. *한국심리학회지: 법*, 11(1), 1-20.

홍현기, 김희송, 지형기, 김기평, 진민진, 홍유나, 현명호 (2014). 숨김 정보 검사에서 유무죄 및 범죄정보 인식에 따른 심리생리 반응 차이. *한국심리학회지: 일반*, 33(3), 627-649.

Banerjee, R., Feng, S., Kang, J. S., & Choi, Y. (2014). *Keystroke patterns as prosody in digital writings: A case study with deceptive reviews and essays*. Paper presented at the Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP).

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129(1), 74-118.

Epp, C., Lippold, M., & Mandryk, R. L. (2011). *Identifying emotional states using keystroke dynamics*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada.

Greene, C. M., Maloney-Derham, R., & Mulligan, K. (2020). Effects of perceptual load on eyewitness memory are moderated by individual differences in cognitive ability. *Memory*, 28(4), 450-460.

Leggett, J., & Williams, G. (1988). Verifying identity via keystroke characteristics. *International Journal of Man-Machine Studies*,

- 28(1), 67-76.
- Meijer, E. H., & Verschuere, B. (2017). Deception detection based on neuroimaging: Better than the polygraph? *Journal of Forensic Radiology and Imaging*, 8, 17-21.
- Miller, G. R., & Stiff, J. B. (1993). *Deceptive communication*. Thousand Oaks, CA, US: Sage Publications, Inc.
- Monaro, M., Spolaor, R., Li, Q., Conti, M., Gamberini, L., & Sartori, G. (2017). *Type Me the Truth!* Paper presented at the Proceedings of the 12th International Conference on Availability, Reliability and Security.
- Monaro, M., Galante, C., Spolaor, R., Li, Q. Q., Gamberini, L., Conti, M., & Sartori, G. (2018). Covert lie detection using keyboard dynamics. *Scientific Report*, 8(1), 1976.
- Monaro, M., Zampieri, I., Sartori, G., Pietrini, P., & Orru, G. (2021). The detection of faked identity using unexpected questions and choice reaction times. *Psychol Res*, 85(6), 2474-2482.
- Monrose, F., & Rubin, A. D. (2000). Keystroke dynamics as a biometric for authentication. *Future Generation Computer Systems*, 16(4), 351-359.
- Nahin, A. F. M. N. H., Alam, J. M., Mahmud, H., & Hasan, K. (2014). Identifying emotion by keystroke dynamics and text pattern analysis. *Behaviour & Information Technology*, 33(9), 987-996.
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Hochenberger, R., Sogo, H., ... Lindelov, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavioral Research Methods*, 51(1), 195-203.
- R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rosenfeld, J. P., Hu, X., Labkovsky, E., Meixner, J., & Winograd, M. R. (2013). Review of recent studies and issues regarding the P300-based complex trial protocol for detection of concealed information. *International Journal of Psychophysiology*, 90(2), 118-134.
- Tomas, F., Dodier, O., & Demarchi, S. (2022). Computational Measures of Deceptive Language: Prospects and Issues. *Frontiers in Communication*, 7.
- Tomas, F., Tsimperidis, I., Demarchi, S., & El Massioui, F. (2020). Keyboard dynamics discrepancies between baseline and deceptive eyewitness narratives. *Applied Cognitive Psychology*, 35(1), 112-122.
- Vrij, A., Fisher, R. P., & Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology*, 22(1), 1-21.
- Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Outsmarting the Liars: Toward a Cognitive Lie Detection Approach. *Current Directions in Psychological Science*, 20(1), 28-32.
- Vrij, A., Hartwig, M., & Granhag, P. A. (2019). Reading Lies: Nonverbal Communication and Deception. *Annual Review of Psychology*, 70, 295-317.
- Vrij, A., & Mann, S. (2001). Telling and detecting lies in a high-stake situation: the case of a convicted murderer. *Applied Cognitive Psychology*, 15(2), 187-203.

Zulkifli, P. A. M. B. (2013). *Applying pause analysis to explore cognitive processes in the copying of sentences by second language users*. (Doctoral). University of Sussex.

1 차원고접수 : 2022. 09. 23.

심사통과접수 : 2022. 10. 31.

최종원고접수 : 2022. 11. 16.

Keyboard dynamics of deceptive eyewitness narratives

Hojin Jeong¹⁾

Hoon Choi^{1,2)}

¹⁾Department of Psychology, Hallym University

²⁾Hallym Institute for Applied Psychological Research

Discriminating the lie of suspect is one of the important tasks in criminal investigations. Lie are known to induce cognitive loads due to additional cognitive task such as creating a false reality or monitoring their statements. The increase of the cognitive load during lying has been observed in studies using neurophysiological measures (e.g., EEG and fMRI), but also observed in keyboard strokes (keyboard input patterns) during writing their statement. Lie detection using keyboard strokes has the advantage of having fewer test restrictions and enabling qualitative analysis of statements, but it is a method that has not been attempted in Korea. This study investigated whether keyboard stroke measures were changed during lying in Korean population. Participants (N=38) typed their diary (baseline) and contents of the video they watched (free recall task). The witness group (n=22) provided a instruction to honestly write what they witnessed, and the lie group (n=16) provided a instruction to write the content as if the suspect had never committed a crime. When comparing the keystroke measures of baseline and free recall, any significant change was not observed in the witness group whereas number of keystrokes per second, the first input time, and the average pause time were significantly changed in the lie group. Specifically, the number of keystrokes per second and the first input time slowed down, and the pause time increased during lying. This result shows that keyboard type behaviors were affected with the increase of cognitive burden when writing deceptive statements, suggesting that keyboard stroke could be an effective indexes for detecting lying.

Key words : lie detection, keyboard stroke, cognitive load