



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

차 옥 균 교수 지도
석사학위 청구논문

얼굴 앙상블 판단에서
개별 얼굴 지각의 역할

2024

성신여자대학교 대학원
심리학과
김수진

얼굴 앙상블 판단에서
개별 얼굴 지각의 역할

차 옥 균 교수 지도

이 논문을 석사학위논문으로 제출함

2023년 11월

성신여자대학교 대학원

심 리 학 과

김 수 진

인 준 서

김수진의 석사학위 논문으로 인준함

2023년 11월

심사위원장..... 서 수 연 (서명 또는



심 사 위 원..... 박 혜 경 (서명 또는



심 사 위 원..... 차 옥 균 (서명 또는



성신여자대학교 대학원

논문개요

얼굴에 대한 앙상블 판단(여러 얼굴의 평균 정서에 관한 판단 등)은 얼굴역전 효과와 같은 전문성 효과의 영향을 덜 받는다. 예를 들어 사람들은 역전된 얼굴 하나의 정서는 잘 판단하지 못하지만, 역전된 얼굴 여러 개의 평균 정서는 상대적으로 잘 판단했다(Sun & Chong, 2020). 이러한 선행 연구들은 얼굴 앙상블 판단이 개별 얼굴에 관한 판단과 다를 수 있음을 시사한다. 본 연구에서는 얼굴 앙상블 판단에 개별 얼굴 지각 능력이 얼마나 영향을 줄 수 있는지 알아보았다. 이를 위해, 도형 자극을 이용하여 측정한 앙상블 판단 능력을 통제했을 때도 서로 다른 얼굴-특정적 정보(정서와 정체성)에 대한 앙상블 과제 수행 사이에 공변량이 있는지 알아보았다. 연구참여자는 얼굴 정서 최빈값 판단 과제, 얼굴 정체성 최빈값 판단 과제, 스플래시 도형 최빈값 판단 과제를 차례로 수행했다. 스플래시 도형은 임의의 윤곽선으로 이루어진 도형이었고, 스플래시 도형 최빈값 판단 과제의 수행은 모든 사물 범주에 사용되는 공통 앙상블 능력의 영향을 통계적으로 제거하기 위해 사용되었다. 각각의 최빈값 판단 과제에서는 아홉 개의 얼굴(또는 아홉 개의 스플래시 도형)이 빠르게 제시되었다가 사라지고, 이어서 세 개의 얼굴(또는 세 개의 도형)이 선택지로 제시되었다. 연구참여자는 아홉 개의 얼굴(또는 아홉 개의 스플래시 도형) 중 가장 많은 수로 제시되었던 얼굴(또는 도형)이 무엇인지 응답했다. 응답 후에는 정답 여부만을 알려주는 피드백이 제공되었다. 연구 결과, 스플래시 도형 최빈값 판단 과제의 수행을 통제했을 때도 얼굴 정서 최빈값 판단 과제와 얼굴 정체성 최빈값 판단 과제의 수행 사이에 정적 편상관이 발견되었다. 본 연구 결과는 얼굴-특정적 정보가 얼굴 앙상블 판단에 기여하며, 따라서 개별 얼굴 지각 능력이 얼굴 앙상블 판단에 기여할 수 있다고 시사한다.

주요어 : 앙상블 판단, 얼굴 지각, 얼굴 정서, 얼굴 정체성, 개인차 연구

목 차

논문 개요

I. 서론	1
1. 앙상블 판단 능력	2
2. 개별 얼굴 지각과 얼굴 전문성	5
3. 얼굴 앙상블 판단과 얼굴 전문성	8
4. 얼굴 앙상블 연구에서 개인차 연구 방법의 유용성	12
5. 연구 문제 및 가설	16
II. 연구 방법	17
1. 연구 대상	17
2. 연구 도구	18
3. 연구 자극	18
4. 연구 절차	22
5. 분석 방법	24
III. 연구 결과	26
IV. 논의	29

참 고 문 헌

ABSTRACT

그림 목 차

<그림 1> 세 가지 연구 과제의 연구 자극 화면 예시.....	20
<그림 2> 얼굴 정서 최빈값 판단 과제 절차.....	23
<그림 3> 얼굴 정서 최빈값 판단 과제의 정확도 잔차와 얼굴 정체성 최빈 값 판단 과제의 정확도 잔차의 관계를 나타내는 산포 도.....	28

I. 서론

우리는 집에서, 학교에서, 직장에서, 그리고 길거리에서 다양한 사람들과 상호작용한다. 얼굴, 몸짓, 말투 등에서 상호작용 대상에 대한 여러 가지 정보를 얻을 수 있는데, 그중에서도 얼굴은 특히 중요한 정보를 전달한다. 바로 그 사람이 누구인지에 대한 정보와 그 사람이 어떤 감정을 느끼고 있는지에 대한 정보이다. 얼굴 정보는 큰 노력을 들이지 않고 처리할 수 있는 것처럼 보이지만, 이는 사람들이 일생에 걸쳐서 다른 사람들의 얼굴을 구분하는 경험을 하면서 얼굴에 대한 전문성을 획득했기 때문이다(Gauthier et al., 1999; Tanaka & Curran, 2001). 사람은 태어난 지 얼마 지나지 않아서 얼굴과 비슷한 형태의 시각 자극에 선호를 보이며(Goren et al., 1975; Kelly et al., 2007; Turati et al., 2008), 유아 때부터 양육자에 맞춰서 다른 요구를 하려면 양육자의 얼굴을 구분해야 한다(Jayaraman et al., 2015; Sugden & Moulson, 2019). 그런데 우리가 일상생활에서 하는 상호작용이 항상 한 명의 사람만을 대상으로 하는 것은 아니다. 일대일 면접을 하는 상황과 같이 한 사람의 얼굴 정보를 처리하는 경우가 있다면, 발표하는 장면과 같이 여러 명의 얼굴 정보를 처리하는 경우 또한 존재한다. 여러 명의 얼굴에서 정서나 정체성과 같은 정보를 처리하는 것도 가능할까?

인간은 여러 물체의 시각 정보를 한 번에 처리해야 하는 경우, 물체 하나하나의 정보를 처리하기보다 여러 물체의 정보를 요약한 통계적 속성(예를 들어 평균적인 크기, 평균적인 색)을 처리한다(Whitney & Yamanashi Leib, 2018). 여러 물체의 통계적 속성(평균, 변산 등)에 관한 판단을 앙상블 판단(ensemble judgment)이라고 하며, 얼굴 또한 앙상블 판단의 대상이 될 수 있다(Haberman & Whitney, 2011). 발표 상황에서 청중이 전반적으로 발표를 즐겁게 듣고 있는지 파악하는 것이 그 예시이다. 그런데 얼굴 앙상블 판단을 하기

위해서는 개별 얼굴을 지각한 후 통계적인 연산을 해야 할까? 청중의 평균적인 정서를 파악하기 위해 청중의 얼굴을 하나하나 보는 것이 일반적인 방법은 아닐 것이다. 사람들은 수백 명의 청중을 보면서 몇 초도 안 되는 짧은 시간에 평균적인 정서를 파악하기도 한다. 실제로 사람들은 네 명의 얼굴에서 평균적인 정서를 알아보았음에도 얼굴 하나하나의 정서는 알지 못했다 (Haberman & Whitney, 2007). 본 연구에서는 얼굴 앙상블 판단 능력이 개별 얼굴을 지각하는 능력과 어떤 면에서 공통점이 있고 어떤 면에서 차이점이 있는지 알아보려고 한다.

1. 앙상블 판단 능력

시각 체계는 효율적인 정보 처리를 위하여 눈을 통해 동시에 들어오는 여러 물체 집합(set)의 정보를 요약하여 추출하는 능력을 가지고 있는데(Ariely, 2001; Chong & Treisman, 2003; Whitney & Yamanashi Leib, 2018), 이를 앙상블 판단 능력(ensemble judgment ability)이라고 한다. 과일 가게에서 사과 알이 평균적으로 더 굵은 상자가 무엇인지 비교한다거나, 숲을 덮은 나뭇잎의 전체적인 색 변화를 보고 계절의 변화를 느끼는 것은 모두 앙상블 판단 능력이 사용되었다고 이야기할 수 있다. 앙상블 판단은 시각 체계가 처리하는 물체 정보의 복잡성에 따라 다양한 처리 수준에서 발생할 수 있는데, 주로 하위 수준 시각특질(low-level visual feature)에 대한 앙상블 판단과 상위 수준 시각특질(high-level visual feature)에 대한 앙상블 판단이 논의된다.

먼저, 하위 수준 시각특질(low-level visual feature)에 대한 앙상블 판단이란 선분이나 원과 같은 단순한 물체의 움직임(motion; Watamaniuk et al., 1989), 방위(orientation; Parkes et al., 2001), 그리고 공간적 위치(spatial position; Alvarez & Oliva, 2008) 등 시각특질의 통계적 속성에 관한 판단을 말한다. 우

리가 이런 단순한 시각특질에 대한 상상불 판단 능력을 가지고 있다는 것을 어떻게 알 수 있을까? Chong과 Treisman의 2003년 연구를 살펴보면, 인간은 다양한 원의 평균 크기를 판단할 수 있는 것으로 보인다. 해당 연구에서 연구자들은 화면을 가로로 나누어 왼쪽과 오른쪽에 각각 서로 다른 크기의 원 12개를 제시하여, 총 24개의 원으로 이루어진 연구 자극을 만들었다. 이후 연구 참여자에게 연구 자극을 200ms 동안 보여주고 왼쪽과 오른쪽 중 어떤 원 집합(set)의 평균 크기가 더 컸는지 응답하도록 하였다. 동일한 연구에서 연구 참여자들은 왼쪽 화면과 오른쪽 화면에 하나씩 제시된 원을 보고 크기를 판단하는 과제도 수행하였는데, 두 과제 모두 비슷한 정도로 잘 수행하였다.

일반적으로 단기 기억에 저장할 수 있는 정보의 양이 대략 3개에서 4개인 것을 고려하면, 12개의 원으로 이루어진 두 개의 집합을 한 번에 보고 각 집합의 원 평균 크기를 계산하는 것은 불가능한 일처럼 보일 수 있다. 그러나 Chong과 Treisman(2003)은 연구 결과를 통하여 우리가 보는 물체의 수가 작업기억에서 처리할 수 있는 정보의 용량을 넘어서도 통계적인 요약 처리를 할 수 있음을 보여주었으며, 이런 능력을 상상불 판단 능력이라고 이름을 붙였다. 이와 유사하게 여러 개의 선분의 평균 기울기(이하 방위, orientation) 또한 판단할 수 있는 것으로 나타났다(Kacin et al., 2021; Yörük & Boduroglu, 2020). 여러 개의 점(dots)이 무작위로 움직이는 영상(cinematogram)을 활용한 연구에서는 연구참여자들이 점들이 향하는 평균 방향과 점들이 움직이는 평균 속도를 정확하게 응답했다고 보고한다(Watamaniuk et al., 1989). 시각 특질의 평균뿐 아니라 다른 통계치에 관한 판단도 가능하다. 사람들은 원 크기의 변산(Cha et al., 2022)이나 최솟값, 최댓값(Khayat & Hochstein, 2018)에 관한 판단도 할 수 있었다. 이는 우리가 다수의 물체 정보를 요약할 때, 다양한 시각특질을 요약할 수 있다는 것을 보여준다.

선의 방위, 원의 크기와같이 단순한 물체의 시각 특질뿐 아니라 얼굴의 표

정이나 정체성과같이 복잡한 시각적 특질을 지닌 대상의 정보 또한 요약해서 처리할 수 있는데, 이를 상위 수준 시각특질(high-level visual feature)에 대한 양상블 판단이라고 한다. 단상 위에서 발표하거나 야구 경기를 관람하기 위해 경기장을 방문하는 경우 많은 사람의 얼굴을 한 번에 보는 경험을 할 수 있다. 이때, 한 사람의 얼굴을 하나씩 자세히 살펴보는 것이 아니라 여러 사람의 얼굴을 한 번에 처리하여 다수의 사람이 느끼는 평균적 감정이나 장소의 분위기를 파악할 수 있는데, 이때 사용되는 것도 양상블 판단 능력이다.

Haberman과 Whitney(2007)는 모핑(morphing) 기법을 통해서 중립 얼굴과 정서 얼굴의 중간 얼굴을 만들어 양상블 판단 연구를 수행한 바 있다. 행복(happy)과 슬픔(sad) 표정이라는 두 극단의 정서를 띠고 있는 얼굴을 모핑하여 행복과 슬픔 사이의 다양한 정서가를 갖는 연구 자극을 제작하였고, 모핑한 사진 네 개로 구성된 화면을 제시한 뒤에 하나의 얼굴 사진을 보여주고 앞서 본 네 개의 얼굴들의 평균 정서가 뒤에 본 하나의 얼굴 정서보다 행복하거나 슬픈지 응답하도록 하였다. 동일한 연구에서 얼굴 네 개로 구성된 화면을 본 뒤에 한 개의 얼굴을 보여주고 그 얼굴이 이전에 봤던 네 개의 얼굴 중 있었는지 응답하기도 하였다. 그 결과, 연구참여자들은 평균 정서 과제에서 얼굴이 있었는지 응답하는 과제보다 더 나은 수행을 보였다. 즉, 앞서 본 네 개의 얼굴이 무엇이었는지 정확하게 기억하지 못했지만, 네 개의 얼굴에서 평균 정서 정보를 빠르게 추출할 수 있었다는 것이다.

다른 선행 연구(Haberman & Whitney, 2011)에 따르면, 네 개뿐만 아니라 더 많은 수의 얼굴이 제시되어도 얼굴 정서 판단을 할 수 있는 것으로 보인다. 해당 선행 연구에서는 표정을 띠고 있는 열여섯 개의 얼굴로 이루어진 얼굴 집합을 사용하였다. 연구참여자들은 총 두 개의 자극을 보고 두 번째로 제시된 얼굴들의 평균 정서가 첫 번째로 제시된 얼굴들의 평균 정서보다 긍정적인지 부정적인지 응답해야 했다. 그뿐만 아니라 정확히 어떤 얼굴이 바뀌었는

지 개별 얼굴의 변화에 대해 응답하였다. 그 결과, 연구참여자들은 얼굴들의 평균 정서의 변화는 잘 지각하였지만 정확하게 어떤 얼굴의 정서가 변화하였는지는 잘 응답하지 못했다. 즉, 개별 얼굴을 하나씩 지각하여 전체적인 얼굴의 표정을 계산하는 것이 아니라 여러 얼굴을 전반적으로 파악하여 앙상블 판단이 가능하다는 의미이다. 표정 외에도 새의 종 평균이나(Chang & Gauthier, 2022), 자동차 모델의 다양성(Cha et al., 2021) 판단도 가능하다. 이러한 연구 결과는 우리가 단순한 물체의 시각 특질뿐만 아니라 복잡한 대상에 대한 앙상블 판단을 할 수 있다는 것을 보여준다.

2. 개별 얼굴 지각과 얼굴 전문성

얼굴은 상위 수준 시각특질(high-level visual feature) 앙상블 판단의 대표적인 예시이다(Haberman & Whitney, 2007; Sweeny & Whitney, 2014; Whitney & Yamanashi Leib, 2018). 얼굴을 지각하는 것은 타인과의 상호작용에서 큰 역할을 하는데, 이는 얼굴을 통해 다양한 정보를 얻을 수 있기 때문이다. 대표적으로 얼굴에서는 상대방이 누구인지 알려주는 정체성 정보와 어떤 기분을 느끼고 있는지 알려주는 정서 정보를 얻을 수 있다. 얼굴이 지닌 이러한 정보에 대한 처리는 여러 상황에서 일어나는데, 그 예로 친구와 마주 앉아 이야기를 나누는 상황이나 많은 사람들 앞에서 발표하는 상황이 있다. 즉, 정체성 및 정서 판단은 한 사람을 대상으로 일어나기도 하며 여러 사람을 대상으로 일어나기도 하는 것이다. 다수를 대상으로 정체성과 정서를 판단할 때는 앙상블 판단 능력이 사용된다(Haberman & Whitney, 2007, 2009; Sun & Chong, 2020).

사회적 상호작용 상황에서 상대방이 누구인지 알아보고 그 사람이 어떤 기분인지 표정을 통해서 알아차리는 것은 중요하다. 사람들은 평생 타인의 정체

성과 그들이 표현하는 정서를 구별하며 살아왔고, 이런 경험은 사람들을 얼굴 정보 처리의 전문가로 만들었을 것이다(Gauthier et al., 1999; Tanaka & Curran, 2001; Rossion et al., 2002; Curby et al., 2009; Sunday & Gauthier, 2018). 그런데도 사람들이 어떤 얼굴은 더 잘 알아보고 다른 얼굴은 잘 알아보지 못하는 상황 또한 존재하는데, 그 예시가 바로 타인종 효과(other race effect)이다. 타인종 효과는 자신과 동일한 인종의 사람 얼굴을 잘 구별하는 반면, 다른 인종 사람 얼굴의 구별에 어려움을 겪는 현상이다. 예를 들어, 한국인이 같은 한국인의 얼굴은 잘 구별하지만, 미국인의 얼굴을 구별하는 데 어려움을 겪을 수 있다. 타인종 효과를 설명하는 한 가지 가설에 따르면 자신의 인종 얼굴 정보를 오랫동안 처리한 경험이 다른 인종의 얼굴보다 자신의 인종 얼굴 간의 차이를 지각하는 민감성을 높인다(Malpass & Kravitz, 1969; Tanaka et al., 2004; Walker & Hewstone, 2006). 이는 자주 본 인종 얼굴에 대한 경험이 전문성을 형성하고, 이 전문성이 특정 인종 얼굴을 더 잘 처리하도록 한다는 가설을 지지한다.

현대 사회는 온라인 스트리밍 서비스의 발달로 서양인들이 등장하는 해외 콘텐츠 시청이 가능한 환경이다. 해외 미디어를 많이 소비하는 한국인의 경우에는 서양인 얼굴에 노출되었기 때문에 타인종 효과가 나타나지 않을 수 있다고 생각할 수 있다. 그러나 한국에 거주하는 한국인들은 일상생활에서 서양인의 얼굴을 보고 누구인지 구분하는 일은 드물다. 드라마와 같은 해외 콘텐츠에서는 등장인물의 직업이나 역할, 혹은 자주 입는 의상 등을 통해 해당 인물의 정보를 파악할 수 있기에 얼굴로 인물을 구별하기 위하여 들이는 노력의 크기가 작다. 즉, 미디어에서 보는 얼굴보다는 일상생활에서 상호작용하는 얼굴에 대한 전문성을 더 많이 획득할 것이고 이에 따라 타인종 효과가 나타날 수 있다.

얼굴 전문성 가설이란 인간이 얼굴을 잘 처리하는 이유는 얼굴에 대한 경험

을 많이 했기 때문이라는 주장이다. McGugin과 동료들(2011)은 얼굴과 같이 전문성을 갖는 목표 자극을 찾을 때, 전문성을 가지고 있는 다른 자극이 방해되는지 알아보았다. 전문성이 미치는 영향을 알아보기 위해서 자동차 모델에 대한 구분 능력이 뛰어난 것으로 예측되는 자동차 전문가와 그렇지 않은 일반인을 연구참여자로 모집하였다. 연구참여자의 과제는 다음과 같다. 먼저 제시되는 얼굴 1번과 얼굴 2번을 기억한다. 이후 여러 개의 얼굴 사진과 물체 사진으로 이루어진 연구 자극이 제시되면 화면 속에서 기억한 얼굴을 찾고, 그 얼굴이 1번이었는지 2번이었는지 응답하였다. 한 화면에 여러 개의 얼굴 및 물체 사진이 제시된다는 연구 자극의 특징을 고려하면, 얼굴 사진의 수가 많을수록 목표 얼굴 사진을 찾는 데 오랜 시간이 걸릴 것이라고 예상할 수 있다. 또한 얼굴 사진을 처리하는 것과 같은 신경 기제를 사용하는(McGugin et al., 2014) 자극이 많을수록 응답 시간이 오래 걸릴 것이다. 연구 결과, 자동차 전문가 집단의 연구참여자들은 여러 개의 얼굴과 자동차 사진이 제시된 화면을 볼 때, 여러 개의 얼굴과 소파 사진이 제시된 화면을 볼 때보다 목표 자극을 찾는 데 걸린 시간이 더 길었다. 소파 사진은 얼굴 사진을 찾을 때 큰 영향을 미치지 않았지만, 자동차 사진은 얼굴 사진을 찾는 것을 방해했다고 볼 수 있다. 즉, 전문성이 있는 물체의 존재는 다른 전문성이 있는 물체를 찾는 것을 방해한다는 것이다. 자동차 전문가가 아닌 연구참여자에게서는 이런 전문성 효과가 나타나지 않은 것을 고려하였을 때, 사람들이 얼굴 정보 처리를 할 때 전문성을 사용한다고 생각할 수 있다.

얼굴 전문성 효과의 다른 예로는 얼굴 역전 효과(face inversion effect)가 있다. 얼굴 역전 효과는 얼굴 사진이 뒤집혀서, 즉 역전(inversion)되어 제시될 때, 얼굴 정보 처리를 잘하지 못하는 것을 말한다. 사람들은 얼굴이 역전되어 제시되면 눈이나 입 모양이 자연스럽지 않다는 것을 잘 지각하지 못하고(Yin, 1969; Thompson, 1980), 표정을 인식하는 데 더 오랜 시간이 걸린다(Derntl et

al., 2009). 경험을 통해 전문성을 가지게 된 얼굴의 특징을 그동안 경험했던 것과 다른 형태로 보면서 처리해야 하기 때문이다. 얼굴은 윗부분에 눈이, 가운데에 코가, 그리고 그 아래에 입이 위치하는데, 역전된 얼굴은 윗부분에 입이 위치하고 아랫부분에 눈이 위치한다.

합성 얼굴 효과(composite face effect)도 얼굴 전문성 효과의 일종으로 볼 수 있다. 합성 얼굴이란, 얼굴의 가운데를 기준으로 얼굴 이미지를 위와 아래로 나뉘도록 반으로 잘라 서로 다른 얼굴 이미지의 윗부분과 아랫부분을 합친 연구 자극을 의미한다. 합성 얼굴을 활용한 연구에서 합성 얼굴 자극 두 개를 보여주었는데, 이때 연구참여자는 두 자극의 얼굴 윗부분이 동일한지 판단하는 과제를 수행하였다. 연구 결과, 합성 얼굴 자극의 윗부분이 동일하여도 아랫부분이 다르면 두 얼굴이 다르다고 응답하는 확률이 높아지는 것을 확인하였다(Young et al., 1987). 이는 얼굴의 위와 아래를 동시에 본 풍부한 경험 때문에 얼굴 윗부분을 응시하면 자연스럽게 얼굴의 아랫부분에도 주의가 가는 현상이라고 설명할 수 있다(Richler et al., 2012).

3. 얼굴 양상불 판단과 얼굴 전문성

사람들은 얼굴과 같은 복잡한 시각적 특질 자극 또한 양상불 판단을 할 수 있다(Haberman & Whitney, 2007, 2009, 2011; Sun & Chong, 2020). 또한 얼굴 정보 처리에는 얼굴 특징적인 전문성이 사용된다(Gauthier et al., 1999; Tanaka & Curran, 2001; Rossion et al., 2002; Curby et al., 2009; McKeeff et al., 2010; Sunday & Gauthier, 2018). 그렇다면 얼굴 양상불 판단에서도 얼굴 전문성이 영향을 미치는 것일까? 얼굴 전문성과 얼굴 양상불 판단의 관계에 대해 몇몇 선행 연구에서 연구한 바 있다.

Haberman과 Whitney가 수행한 선행 연구들(Haberman & Whitney, 2007,

2009, 2011)은 똑바로 제시된 얼굴(uptight faces) 여러 개가 제시되었을 때 평균 정서 정보를 추출할 수 있다는 것을 보여준다. Haberman과 Whitney(2007)는 행복한 얼굴과 슬픈 얼굴을 모핑하여 점진적으로 정서 강도가 변하는 얼굴 이미지를 만들고, 한 화면에 다양한 정서 강도를 가진 네 개의 얼굴을 제시했다. 네 개의 얼굴이 사라진 뒤 하나의 얼굴을 보여주고 앞에서 본 얼굴들보다 행복한지 슬픈지 응답하도록 하였다. 그 결과 연구참여자들은 뒤에 제시된 한 개의 얼굴이 앞에서 본 네 개의 얼굴의 평균 정서보다 행복하거나 슬픈지 더 잘 응답하였다. 동일한 연구에서 연구자들은 사람들이 연구 자극을 구성하는 네 개의 얼굴을 개별적으로 표상하는지도 확인하였다. 평균 정서를 판단하는 과제에서 사용한 것과 동일한 실험 자극을 사용하였으나 응답 화면에서 하나의 얼굴이 아닌 두 개의 얼굴을 제시하고 자신이 본 얼굴이 무엇인지 선택하도록 했다. 과제 수행 결과, 연구참여자들의 정답률은 우연 수준을 넘지 못하였다(Haberman & Whitney, 2007). 연구자들은 이 결과를 사람들이 네 개의 얼굴을 각각 기억하지는 못하였다는 의미로 해석하였다.

다른 연구(Haberman & Whitney, 2011)에서는 모핑한 얼굴 16개를 4행 4열로 나열하여 만든 연구 자극을 사용하여 여러 개의 얼굴 평균 정서 추출 능력과 개별 얼굴 지각 능력을 살펴보았다. 이 연구에서는 연구참여자들에게 얼굴 16개로 이루어진 연구 자극 화면을 총 두 번 보여주었으며, 과제는 뒤에 제시된 얼굴들의 평균 정서가 먼저 본 얼굴들과 어떻게 다른지 응답하는 것이었다. 연구자들은 먼저 제시되는 16개의 얼굴 중에서 4개의 얼굴을 선택하여 원래 얼굴보다 긍정적인 얼굴로 바꿔 뒤에 제시되는 얼굴 16개를 만들거나 원래 얼굴보다 부정적인 얼굴로 바꿔서 뒤에 제시되는 얼굴 16개를 만들었고, 연구참여자들은 16개의 얼굴 중에서 어떤 얼굴의 표정이 바뀐 것인지에 대해서도 응답해야 했다. 연구 결과, 연구참여자들은 두 번째로 제시된 얼굴 표정들이 평균적으로 더 행복한지 슬픈지에 대해서는 잘 응답하였지만, 4행 4열로 이루

어진 얼굴 이미지 중에서 정확하게 어떤 위치의 얼굴 표정이 변화하였는지는 잘 응답하지 못하였다. 이러한 연구 결과는 앙상블 판단을 할 때 개별 얼굴에 하나씩 주의를 집중하여 얼굴 정보를 처리하는 것이 아니라, 얼굴들의 통계적인 요약 정보를 추출하여 처리한다는 것을 보여준다. 다시 말해, 개별 얼굴 정서 판단 능력과 얼굴 정서 앙상블 판단 능력은 서로 다른 능력일 수 있다는 의미이다. 지금까지 논의된 앙상블 판단 연구들은 사람들이 얼굴 앙상블 판단을 잘한다는 것을 보여주었지만, 그 이유가 얼굴에 대한 전문성 때문인지 확인하지는 않았다. 얼굴을 지각할 때, 얼굴 전문성이 영향을 미치는지 알아보기 위해서는 개별 얼굴 지각에서 나타나는 얼굴 전문성의 효과가 앙상블 판단에서도 나타나는지를 다른 선행연구를 살펴볼 필요가 있다.

Haberman과 Whitney의 2009년 연구에서 뒤집힌 얼굴(inverted faces)과 뒤섞인 얼굴(scrambled faces)을 활용한 바 있다. Haberman과 Whitney의 2007년 연구와 같이 네 개의 모핑 얼굴을 제시하고 이후 제시된 얼굴 한 개의 정서가를 평가하는 과제를 사용하였다. 이전 연구와 다른 점은 위아래를 뒤집은 얼굴 사진과, 얼굴 사진을 3행 4열의 조각으로 나누어 퍼즐처럼 뒤섞은(scrambled) 사진을 연구 자극으로 사용하였다는 점이다. 연구 결과, 뒤집은 얼굴과 뒤섞은 얼굴 모두 똑바르게 제시된 얼굴(upright faces)보다 평균 정서 판단을 잘하지 못하는 것을 확인하였다. 이는 뒤집힌 얼굴이나 뒤섞인 얼굴이 똑바른 얼굴과 동일한 시각적 특질을 가지고 있음에도 불구하고 제시되는 방법이 달라짐에 따라 얼굴 정보를 빠르고 정확하게 처리하는 얼굴 전문성이 손상된다는 것을 알 수 있다. 즉, 얼굴이 뒤집히거나 뒤섞이면 얼굴 앙상블을 더 잘 못할 수 있다는 것이다. 그런데 Haberman과 Whitney(2009)의 연구에는 중요한 한계점이 있다. 연구참여자는 뒤집힌 얼굴 네 개의 평균 정서를 뒤집힌 얼굴 한 개의 정서와 비교하여야 했는데, 이때 뒤집힌 얼굴 네 개의 평균 정서 판단을 잘하여도 비교 대상인 뒤집힌 얼굴 한 개의 정서를 잘 지각하지 못

해 과제 수행이 떨어졌을 가능성이 존재한다.

Sun과 Chong(2020)의 최근 연구에서 선행 연구(Haberman & Whitney, 2009)의 한계점을 극복한 실험을 수행한 결과, 얼굴 반전 효과가 여러 개의 얼굴을 볼 때보다 하나의 얼굴을 볼 때 더 강력하게 나타난다는 것을 보여주었다. 연구자들은 행복과 분노 얼굴을 모핑하여 사용하였으며 연구참여자들에게 한 개의 얼굴이나 여러 개의 얼굴을 한 화면에 제시하고 평균적인 정서 강도에 대해 응답하도록 하였다. 이때, 선행 연구(Haberman & Whitney, 2009)와 다르게, 모든 얼굴이 뒤집혀서 제시되는 시행과 여러 얼굴 중 일부 또는 전체가 뒤집혀서 제시되는 시행이 포함되었으며, 응답 화면에서 비교 대상으로 제시되는 얼굴 한 개는 항상 똑바로 제시되었다. 연구 결과, 뒤집힌 얼굴의 수가 증가하더라도 앙상블 판단 과제에서는 수행의 변화가 없었다. 즉, 여러 개의 얼굴이 뒤집혀서 제시되었음에도 얼굴들의 평균 정서를 판단할 수 있다는 의미이다. 그에 반하여 하나의 얼굴이 뒤집혀서 제시될 때 정서 강도를 판단하기 어려웠다는 결과를 도출하였는데, 이는 얼굴 반전 효과가 여러 개의 얼굴에서 평균 정서를 판단할 때보다 개별 얼굴 정서를 구별할 때 더 강한 영향을 미친다는 것을 의미한다.

얼굴 전문성 효과의 다른 예로 얼굴과 다른 전문성을 갖는 물체를 동시에 볼 때 전문성이 있는 물체를 처리하는 신경 기제 간에 경쟁이 일어나 얼굴을 찾는 데 걸리는 탐색 시간이 늘어나는 효과가 있다(McKeeff et al., 2010; McGugin et al., 2011). 얼굴 앙상블 판단에 있어서 얼굴 전문성 효과가 약해지거나 나타나지 않는다면, 앙상블 과제를 수행할 때 다른 물체에 대한 전문성이 방해 자극으로써 영향을 주지 않을 수도 있다. Sun과 Gauthier(2021)는 얼굴 사진 네 개와 자동차 사진 네 개를 동시에 제시하고 얼굴에 대한 앙상블 판단 과제를 수행하도록 하여 얼굴 전문성과 자동차 전문성 간의 경쟁이 앙상블 판단에서도 나타나는지 확인하였다. 그 결과, 연구참여자들이 자동차 자극

으로 인해 얼굴 앙상블 판단에 방해받는 정도는 자동차 전문성에 따라 차이가 나지 않았는데, 연구자들은 이를 자동차 전문성이 얼굴 앙상블 판단을 방해하지 않았다고 해석하였다. 이러한 연구 결과들은 앙상블 판단 상황에서 특정 대상에 대한 전문성 효과, 즉 얼굴 전문성 효과가 나타나지 않거나 약해짐을 시사한다.

얼굴은 정체성이나 정서와같이 복잡한 상위 수준의 처리가 필요한 시각특질을 가졌으며, 개별 얼굴을 지각할 때 전문성 효과가 나타난다. 그렇지만 여러 개의 얼굴 정보를 요약하여 처리하는 앙상블 판단을 할 때는 전문성 효과가 나타나지 않는 것으로 보인다. 그렇다면 얼굴에 대한 앙상블 판단을 할 때는 개별 얼굴 판단에 사용하는 기제를 하나도 사용하지 않는 것일까? 본 연구에서는 이를 알아보기 위하여 개인차 연구 방법을 사용하였다.

4. 얼굴 앙상블 연구에서 개인차 연구 방법의 유용성

사람은 모두 얼굴 전문가지만, 사람들 사이에는 얼굴 재인 능력에 개인차가 있다. 물리적인 시각적 손상이나 질환을 가지고 있지 않더라도, 사람 얼굴을 잘 기억하지만 표정을 잘 지각하지 못하는 경우도 있고, 얼굴에 나타난 정서를 잘 지각하지만 누구인지는 잘 기억하지 못하는 경우도 있다. 얼굴 지각의 개인차는 경험을 통해서 알 수 있을 뿐만 아니라 실험 과제를 사용하여 측정할 수 있는데, Duchaine과 Nakayama(2006)의 캠브리지 얼굴 기억 검사(Cambridge Face Memory Test; CFMT)는 얼굴 재인 능력의 개인차를 측정하는 대표적 검사이다. CFMT는 먼저 여섯 명의 다른 사람들의 얼굴을 보고 기억하는 학습 단계와 세 개의 얼굴 중 앞에서 기억한 얼굴 하나를 선택하는 문항을 여러 번 수행하는 검사 단계로 구성된다. 검사 단계에서 제시되는 세 개의 얼굴은 정면과 우측 혹은 좌측과 같은 다양한 각도로 촬영되었으며, 사

진 촬영 시 다양한 각도의 조명을 사용하여 얼굴 사진에 진 음영 또한 다양했다.

한 연구에서(McGugin et al., 2012) CFMT와 동일한 실험 과제를 사용하여 여러 대상에 대한 지각 능력을 측정하는 바 있다. 연구자들은 얼굴, 올빼미, 비행기, 자동차, 나비, 버섯, 물새, 나뭇잎, 그리고 오토바이 사진을 활용하였는데, 학습 단계에서 여섯 개의 사진을 기억하고 검사 단계에서 기억한 물체를 선택하는 정확도를 측정하였다. 이후 각 사물 범주에 대한 지각 능력 점수 간의 상관을 구하였는데, 얼굴과 자동차를 제외한 나머지 사물들의 지각 능력 간의 높은 정적 상관이 발견되었다. 이 결과는 얼굴을 지각하는 능력이 새나 버섯, 비행기 등을 지각하는 능력과는 구분된다고 해석할 수 있다. 얼굴뿐만 아니라 자동차 지각 능력 또한 다른 대상들의 지각 능력과 상관이 낮았는데, 이는 일상생활에서 얼굴과 자동차에 대한 경험의 개인차가 새나 버섯, 비행기 등에 대한 경험의 개인차보다 크기 때문일 것이다. 실제로 고향의 인구수와 얼굴 지각 능력에 관해서 알아본 연구에 따르면, 고향의 인구수가 적을수록 얼굴을 잘 구별하지 못하는 경향이 나타난다(Balas & Saville, 2017). McGugin과 동료들(2012)의 개인차 연구를 통해 얼굴 지각 능력과 자동차 지각 능력에는 버섯이나 새, 비행기와 같은 사물의 지각 능력만으로는 설명할 수 없는 개인차(예를 들어, 얼마나 다양한 얼굴에 노출되었는지의 차이)가 있다는 것을 알 수 있었다. 이처럼, 개인차 연구 방법은 사람이 가진 인지 능력에 기여하는 요인들, 즉 어떤 인지 과제 수행에 사용되는 능력이나 경험을 파악할 수 있다는 장점이 있다.

최근 양상불 판단 분야에서는 다양한 양상불 판단에 공통으로 사용되는 시각적 통계 능력이 존재하는지에 관한 연구가 개인차 연구 방법을 활용하여 진행되고 있다. 선행 연구에서는 연구참여자들에게 한 화면에 제시된 선분 12개의 평균 길이와 방위를 판단하도록 한 뒤, 길이 양상불 판단 능력과 기울기

양상블 판단 능력 간의 정적 상관성이 있었음을 보여주었다(Kacin et al., 2021). 이 연구는 길이의 평균 계산과 방위의 평균 계산이 서로 다른 시각적 특질을 처리하는 과정이더라도 비슷한 종류의 연산이 필요하면 같은 기제를 사용하는 것이라고 제안한다. Kacin과 동료들(2021)의 연구 방법에는 한계점이 있는데, 본래 측정하고자 하였던 양상블 판단 능력에 공통으로 영향을 미치는 다른 요인이 충분히 고려되지 않았다는 점이다. 예를 들어, 개별 선분의 길이나 방위를 지각하는 능력이 과제 수행에 영향을 미쳤을 수 있다.

개인차 연구에서는 편상관을 이용해서 연구 목적에 따라 통제되어야 하는 개인의 능력을 측정하여 통계적으로 제거할 수 있다. Cha와 동료들(2022)은 평균 비교와 변산 비교에 공통으로 사용되는 양상블 판단 능력이 있는지 알아보고자 편상관을 이용하였다. 이를 위해 연구참여자들의 평균 비교 능력과 변산 비교 능력, 그리고 두 능력에 공통으로 영향을 줄 것이라고 예상되는 원 크기 지각 능력을 측정하였다. 연구참여자들은 세 과제에서 모두 동일하게 두 개의 자극 화면을 보고 비교하였다. 평균 비교 능력을 측정하는 과제에서는 다양한 크기의 원 여섯 개로 이루어진 배열이 두 번 제시된 후 처음 제시된 배열과 두 번째로 제시된 배열 중 원 여섯 개의 평균 크기가 더 큰 것이 어느 배열인지 응답하여야 했다. 변산 비교 능력의 측정을 위한 과제에서도 다양한 크기의 원 여섯 개로 이루어진 배열이 두 번 제시되었는데, 첫 번째 원 배열과 두 번째 원 배열 중 원의 크기가 더 다양했던 배열이 무엇인지 응답하여야 했다. 마지막으로 원 크기 지각 능력 측정 과제의 경우 앞선 두 과제와 다르게 한 개의 원으로 이루어진 자극이 두 번 연달아 제시되었고, 두 원 중에서 크기가 더 큰 원이 무엇인지 응답하여야 했다. 연구 결과, 평균 비교 능력과 변산 비교 능력에서 원 지각 능력을 제거한 뒤에도 두 능력 간의 정적 편상관이 발견되었다. 이는 원의 평균을 비교하는 능력과 변산을 비교하는 능력을 사용할 때 공통으로 사용되는 능력이 있는데, 이 능력에서 양상블 지각과 무

관한 원 지각 능력을 제거하고도 공통된 부분이 일정량 남는다는 의미이다. 연구자들은 원 지각 능력을 제거하고도 남은 공통된 부분이 앙상블 판단과 관련된 능력일 것이라고 해석하였다.

사람은 다양한 얼굴 특성에 관한 앙상블 판단을 할 수 있다(Haberman & Whitney, 2007, 2009, 2011; Whitney & Yamanashi Leib, 2018). 그런데 얼굴 앙상블 판단을 할 때 얼굴 전문성으로 인한 효과가 줄어들거나 나타나지 않는 것처럼 보인다(얼굴 역전 효과: Sun & Chong, 2020; 전문성이 있는 물체의 과제 방해: Sun & Gauthier, 2021). 얼굴 앙상블 과제에서는 정체성, 정서 같은 얼굴 특성의 평균이나 다양성을 판단해야 하므로 얼굴 정보를 필수적으로 처리해야 하는데, 얼굴 전문성 효과가 나타나지 않는 것이 모순적으로 느껴질 수 있다.

본 연구에서는 얼굴 앙상블 판단에서 개별 얼굴에 대한 정보가 사용되는지 알아보기 위하여 선행 연구(Cha et al., 2022)와 같은 접근 방법을 사용하였다. 선행 연구에서는 원 크기 평균 비교 능력과 원 크기 변산 비교 능력 간 공유하는 능력에서 원 지각 능력을 통계적으로 제거하는 방법을 사용하였다. 여러 개의 물체를 보고 통계적인 요약 표상을 만드는 능력을 확인하고자 했던 선행 연구와 다르게, 본 연구에서는 개별 물체, 즉 개별 얼굴 지각 능력에 초점을 두었다. 얼굴 정서 앙상블 판단 능력과 얼굴 정체성 앙상블 판단 능력에서 앙상블 판단 능력을 통계적으로 제거한다면, 얼굴 정서 앙상블 판단과 얼굴 정체성 앙상블 판단에서 공통으로 사용되는 능력 중 개별 얼굴 지각과 관련된 능력만을 남길 수 있을 것이다. 자세한 방법은 다음과 같다. 얼굴 정체성 앙상블 판단 과제와 얼굴 정서 앙상블 판단 과제를 활용하여 두 종류의 서로 다른 얼굴 앙상블 판단 능력을 측정했다. 얼굴과 무관한, 즉 얼굴 전문성을 사용하지 않는 앙상블 판단 능력을 측정하기 위해서는 생소한 도형 자극으로 이루어진 앙상블 판단 과제를 사용했다. 이때 세 과제 모두 동일한 자극 제시 방법

과 응답 방법을 사용한다면, 얼굴 정서 앙상블 판단 과제와 얼굴 정체성 앙상블 판단 과제에서 공통으로 사용되는 능력 중 앙상블 판단 능력을 통계적으로 제거하여 개별 얼굴 지각과 관련된 능력만을 남길 수 있다. 이 분석에서 정적인 편상관이 도출된다면 서로 다른 얼굴 앙상블 판단 과제에 개별 얼굴 지각과 관련된 능력이 공통으로 사용된다고 해석할 수 있다.

5. 연구 문제 및 가설

연구 문제 1. 서로 다른 얼굴 특성의 앙상블 판단에 개별 얼굴 지각과 관련된 능력이 공통으로 사용될 것인가.

연구 가설 1. 얼굴 정서 최빈값 판단 과제의 수행과 얼굴 정체성 최빈값 판단 과제의 수행에서 스플래시 도형 최빈값 판단 과제 수행을 통제하였을 때, 얼굴 정서 최빈값 판단 과제의 수행과 얼굴 정체성 최빈값 판단 과제의 수행 간의 정적인 편상관이 도출될 것이다.

II. 연구 방법

1. 연구 대상

본 연구의 참여자는 미국과 영국에 거주하는 만 18세 이상의 성인으로 한다. 본 연구는 Prolific(<https://www.prolific.co/>)에서 온라인 연구에 참여할 연구참여자를 모집하였다. 연구에 대한 모든 설명이 영어로 이루어졌기 때문에 영어 이해에 문제가 없는 사람을 참여자로 하기 위하여 연구참여자를 미국과 영국 거주자로 제한하며, 이는 온라인 실험 플랫폼에서 접속자의 IP 주소를 확인하여 이루어졌다. 연구참여자는 미국법과 영국법 상 성인인 만 18세 이상으로 제한하였다.

본 연구에서는 새로운 행동 연구 과제를 개발하여 사용하기 때문에 사전 연구를 참고하여 검정력 분석을 할 수 없다. 따라서, 연구참여자 수의 적합성을 보장하는 방법으로 Sequential Bayes Factor(SBF) 설계를 사용하였다 (Schönbrodt et al., 2017). SBF 설계를 사용하는 연구에서 연구자는 먼저 사전에 결정한 최소 연구참여자 수(75명) 이상의 데이터를 수집하고, 그 이후로는 데이터 수집과 Bayesian 통계 분석을 병행한다. 연구자는 Bayesian 통계 분석의 결과가 사전에 결정한 기준($BF > 3$ 또는 $BF < 1/3$)을 만족할 때까지 데이터를 수집할 수 있다. 또한, 사전에 최대 연구참여자 수를 결정해 두었다면, 최대 연구참여자 수를 넘는 데이터를 수집한 후에는 통계 분석의 결과가 기준을 만족하지 못하더라도 데이터 수집을 멈추고 BF 값이 1을 넘는지(영가설과 대립가설 중 어느 쪽이 조금이라도 가능성이 높은지) 방향성을 보고할 수 있다. 본 연구는 이러한 SBF 설계를 사용하여 연구참여자 수의 적합성을 확보하였다.

본 연구는 세 개의 연구 과제를 사용하며, 세 연구 과제 모두에서 연구참여

자의 응답 정확도(accuracy)를 수집하였다. 연구를 끝까지 수행한 연구참여자들은 4.50 GBP의 보상을 받았다. 연구 시작에 앞서 모든 연구참여자에게 온라인으로 연구 참여 동의서가 제공되었고, 동의서 맨 아래의 승낙(accept) 버튼을 누르면 연구 참여에 동의하는 것으로 간주하였다. 본 연구는 성신여자대학교 생명윤리위원회(Institutional Review Board; IRB)의 승인을 받은 연구이다 (SSWUIRB-2022-054).

2. 연구 도구

본 연구는 MATLAB(MathWorks Inc., Natick, MA, USA)을 사용하여 얼굴 사진과 스플래시(splash) 도형을 연구 자극에 활용할 수 있도록 수정하였다. 연구는 JavaScript를 사용하여 연구 자극을 제시하는 HTML 웹페이지로 구성하였다. jsPsych 7.2.1 버전을 사용하여 연구 자극이 제시되는 시점을 통제하였다. 본 연구의 실험 웹페이지는 Cognition(<https://www.cognition.run/>)을 통해 진행되며, 연구 데이터 또한 Cognition에 저장된다.

3. 연구 자극

연구참여자들은 자극으로 제시되는 사물 범주와 판단해야 하는 사물의 특성에 따라서 다음의 세 가지 연구 과제를 수행하였다: (1) 얼굴 정서 최빈값 판단 과제, (2) 얼굴 정체성 최빈값 판단 과제, (3) 스플래시 도형 최빈값 판단 과제. 세 개의 과제에서는 각각 다양한 정서를 갖는 얼굴 사진, 다양한 정체성을 갖는 얼굴 사진, 그리고 생소한 모양의 스플래시 도형을 자극으로 사용하였다.

세 가지 연구 과제의 각 시행에서는, 연구 자극 화면과 응답 화면이 연달아

제시되었다. 연구 자극 화면은 144×144 픽셀 크기의 얼굴 아홉 개(그림 1A, 1B)나 스플래시 도형 아홉 개(그림 1C)로 구성되며, 얼굴이나 도형은 3×3의 격자 안에 제시되었다. 이때, 3×3 격자 배열의 각 칸 크기는 168×168 픽셀이며, 각각의 자극은 격자 배열의 한 칸 안에서 무작위 위치에 제시되었다. 즉 가로로 최대 24 픽셀, 세로로 최대 24 픽셀 안에서 무작위로 위치가 결정되었다. 본 연구에서는 연구 자극의 크기와 위치를 시각도(degrees visual angle) 대신 픽셀을 사용하여 보고하였는데, 이는 온라인 연구의 특성상 연구참여자가 사용하는 컴퓨터 모니터의 크기, 해상도 그리고 화면과의 거리를 통제하기 어렵기 때문이다.

얼굴 정서 최빈값 판단 과제에서 한 시행당 제시되는 아홉 개의 얼굴 사진은 한 사람의 네 개의 다른 정서 사진으로 구성되었다. 첫 번째 정서 사진이 다섯 개 제시되어 가장 많은 수로 제시된 정서 사진이 되며, 두 번째 정서 사진이 두 개, 세 번째와 네 번째 정서 사진이 각각 한 개씩 제시되었다. 이해하기 쉽게 예시를 들자면, 얼굴 정서 최빈값 판단 과제의 한 시행에서 분노 사진이 다섯 장, 슬픔 사진이 두 장, 행복 사진이 한 장, 놀람 사진이 한 장 제시될 수 있으며 이때 분노 사진의 수가 가장 많으므로 이 시행의 최빈값은 분노 정서 사진이 된다. 얼굴 정체성 최빈값 판단 과제와 스플래시 도형 최빈값 판단 과제의 배열에서는 가장 많은 수로 제시된 얼굴 이미지 및 스플래시 도형 이미지가 네 개나 다섯 개이며, 두 번째로 많이 제시되는 이미지는 두 개, 나머지 세 개나 두 개의 이미지는 각각 다른 이미지들로 제시되었다. 각 과제의 최빈 얼굴/도형 이미지 개수는 파일럿 실험을 통해 충분한 개인차를 관찰할 수 있도록 난이도를 조절한 결과이다.

응답 화면에는 과제에 따라 얼굴 사진, 또는 스플래시 도형 세 개가 선택지로 제시되었다. 세 개의 선택지 중 하나는 가장 많은 수로 제시된 얼굴/도형이고, 다른 하나는 두 번째로 많이 제시된 얼굴/도형이며, 마지막 하나는 제시

되었던 아홉 개의 이미지 중에 없었던 새로운 얼굴/도형이었다. 응답 화면 속 세 개의 얼굴/도형 아래에는 알파벳 J, K, 그리고 L이 함께 제시되었는데, 이는 연구참여자에게 어떤 버튼을 눌러 응답해야 하는지 알려주기 위함이었다. 모든 자극 화면과 응답 화면은 640×560 픽셀로 이루어진 회색 사각형 안에 제시되었다.

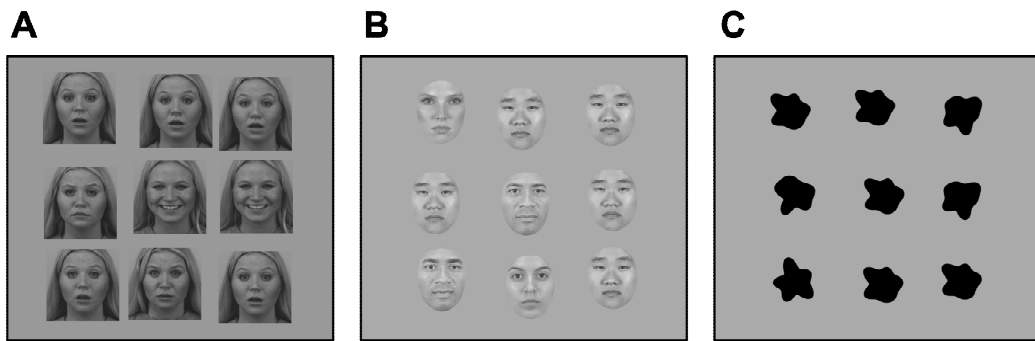


그림 1. 세 가지 연구 과제의 연구 자극 화면 예시. (A) 얼굴 정서 최빈값 판단 과제의 연구 자극 화면 예시. 사용된 얼굴 사진은 Karolinska Directed Emotional Faces (KDEF, Lundqvist et al., 1988)의 AF27AFS, AF27HAS, AF27SUS, AF27SAS 사진 사용. (B) 얼굴 정체성 최빈값 판단 과제의 연구 자극 화면 예시. 사용된 얼굴 사진은 Chicago Face Database (CFD, Ma et al., 2015)에서 선택. (C) 스플래시 도형 최빈값 판단 과제의 연구 자극 화면 예시.

본 연구에서 사용한 얼굴 사진들의 출처는 다음과 같다. 얼굴 정서 최빈값 판단 과제는 Karolinska Directed Emotional Faces(KDEF, Lundqvist et al., 1988)에서 얼굴 사진을 선택하여 사용하였다. KDEF는 20세에서 30세의 백인 여성 35명과 백인 남성 35명의 사진으로 이루어져 있는데, 총 70명의 중립 표정 및 여섯 가지 정서(두려움, 화남, 역겨움, 행복함, 슬픔, 그리고 놀람)의 표정을 다섯 개의 각도에서 촬영한 사진으로 구성된다. 본 연구에서 사용하는 사진을 다음과 같은 절차로 연구에 사용할 수 있도록 수정하였다. 첫 번째로,

KDEF의 모델 70명의 사진 중에서 연구 자극으로 사용할 수 있는 사진을 고르는 작업을 수행하였다. 70명의 다섯 가지 각도의 사진 중에서 과다 노출로 인하여 사진의 전체적인 명암이 다른 사진들과 뚜렷한 차이가 있거나 사람이 아닌 다른 물체가 함께 촬영된 사진을 연구 자극 사진 목록에서 제거하였다. 연구참여자들이 앙상블 과제를 수행할 때, 얼굴과 무관한 시각적 특징으로 앙상블 판단을 할 수 없도록 통제하기 위함이다. 다음으로, 정면을 응시하는 사진 중 여섯 가지 정서의 표정이 하나라도 없으면 해당 모델을 연구 자극 사진 목록에서 제거했는데, 이는 본 연구에서 정면을 응시하는 사진을 사용하기 때문이었다. 본 연구에서 중립 표정은 사용하지 않는데, 이는 각 시행에서 표정이 잘 구별되게 하기 위함이었다. 위의 절차를 거쳐 여성 29명과 남성 22명, 총 51명의 모델 사진이 선택되었다. 모든 사진은 144×144 픽셀 크기로 조절된 후, 회색조로 변환되었다.

얼굴 정체성 최빈값 판단 과제의 얼굴 사진은 Cha와 동료들(2021)이 선행 연구에서 사용한 얼굴 다양성 판단 과제의 얼굴 사진을 사용하였다. Cha와 동료들(2021)은 Chicago Face Database(CFD, Ma et al., 2015)에서 정면을 응시하는 사진을 선택하여 회색조로 수정하고, 이마 위쪽과 얼굴 윤곽선 바깥쪽을 제거하여 머리카락이나 수염과 같이 얼굴과 무관한(non-facial) 특징이 드러나지 않도록 했다. 수정한 얼굴 중 CFD의 얼굴 카테고리(아시아인 여성, 아시아인 남성, 흑인 여성, 흑인 남성, 라틴 아메리카인 여성, 라틴 아메리카인 남성, 백인 여성, 백인 남성)에서 각각 세 명의 모델을 선택했다. 각각의 카테고리에서 모델을 선택할 때는 모든 얼굴 사진의 평균 윤곽선과 가장 윤곽선이 비슷한 모델을 선택했다. 모든 사진의 얼굴 윤곽이 유사했으므로 연구참여자는 오직 얼굴 특징(facial features)을 기반으로 얼굴을 지각해야 한다. 결과적으로 본 연구에서 사용한 얼굴 정체성 과제 사진은 총 24명의 모델 사진이다.

스플래시 도형 최빈값 판단 과제에서 사용한 생소한 도형 네 개를 여섯 세

트 제작하기 위하여 다음의 절차를 따랐다. 첫 번째로, 241개의 1차원 화이트 노이즈 배열을 만들었다. 두 번째로, 각 배열에 로우패스 필터(lowpass filter, cut-off frequency: 6 Hz)와, 노치 필터(notch filter, attenuation frequency: 0 Hz)를 적용하였다. 원에서 필터가 적용된 1차원 배열에 따라 각 부분의 지름을 조절하여 도형 자극을 만들었다. 1차원 배열의 스펙트럼 출력 분포(spectral power distributions)를 입력값으로 하여 k-평균 군집화(k-means clustering)를 수행하여 도형을 여섯 세트로 나누었다. 마지막으로, 각 세트의 중심 도형과 가장 가까운 네 개의 도형을 선택하여 과제에 사용하였다.

4. 연구 절차

연구를 시작하면, 연구참여자는 웹페이지 형태로 제시된 연구 참여 동의서를 읽고 동의서 맨 아래에 있는 “동의합니다” 버튼을 눌러 연구 참여에 동의하였다. 다음 페이지에서 연구참여자는 그들의 나이, 성별, 그리고 인종을 보고하였다. 이후 제시되는 페이지에서 연구참여자가 “시작” 버튼을 누르면, 연구 화면 페이지가 전체 화면으로 전환되었다. 모든 연구참여자는 얼굴 정서 최빈값 판단 과제, 얼굴 정체성 최빈값 판단 과제, 그리고 스플래시 도형 최빈값 판단 과제를 동일한 순서대로 수행하였다. 각 과제가 시작하기 전에 과제에 대한 설명과 연습 시행이 제시되었다.

과제의 절차는 다음과 같았다(그림 2). 얼굴 정서 최빈값 판단 과제의 각 시행은 화면 중앙에 800ms 동안 제시되는 고정점으로 시작하였다. 이어서 700ms 동안 3행 3열 격자 속 얼굴 아홉 개로 이루어진 연구 자극 화면이 제시되었다. 자극 화면이 사라지고 600ms 후에 응답 화면이 제시되었다. 응답 화면에는 세 개의 얼굴 사진과 함께 “어떤 표정의 얼굴이 제일 많이 제시되었는지 선택해 주세요(Which is the face with the most frequent expression?)”

라는 문장이 회색 사각형 아래에 제시되었다. 연구참여자는 키보드에서 세 개의 사진 아래에 적힌 알파벳 J, K, L 중 하나에 해당하는 키를 눌러 어떤 얼굴 표정이 가장 많이 제시되었는지 응답하였다. 연구참여자가 응답한 뒤에 응답에 대한 피드백이 회색 사각형 아래에 “정답(correct)”이나 “오답(wrong)”이라는 단어로 제시되었다. 6번의 연습 시행 이후 96번의 실험 시행이 진행되었다. 본 과제의 자극 제시 순서는 처음 연구 자극을 제작할 때 무작위로 설정하고 모든 연구참여자에게 동일한 순서를 사용하였다. 이는 연구참여자들이 가진 최빈값 판단 능력의 개인차와 관련이 없는 행동 변산(behavioral variability)의 영향을 최소한으로 하기 위함이다.

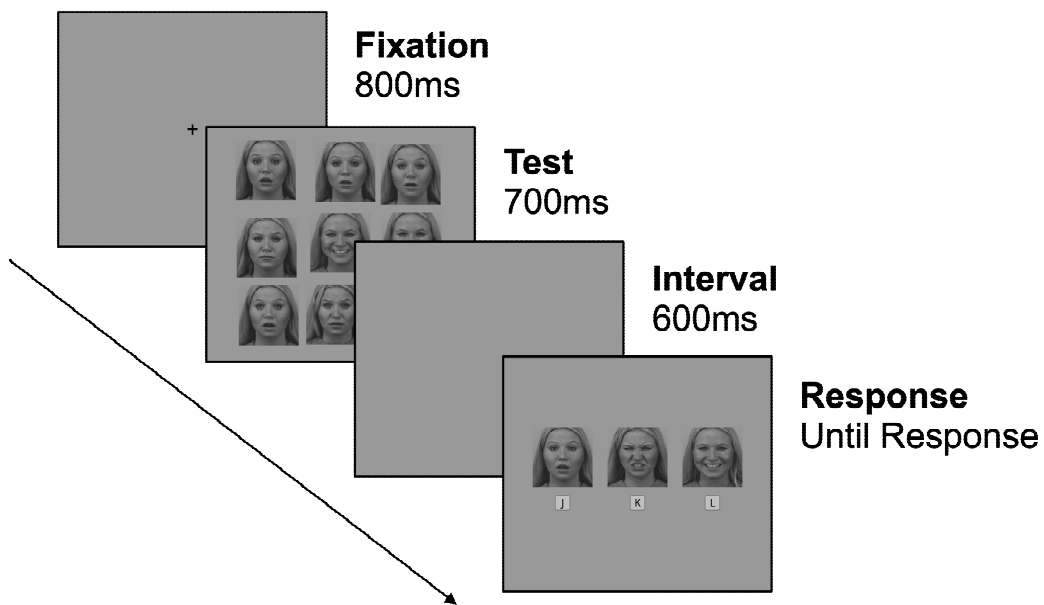


그림 2. 얼굴 정서 최빈값 판단 과제 절차

얼굴 정체성 최빈값 판단 과제의 절차는 얼굴 정서 최빈값 판단 과제의 절차와 두 가지 부분을 제외하고는 모두 동일하였다. 첫 번째로, 연구참여자들이

머리카락이나 수염, 목, 어깨와 같은 특징이 모두 제거된 사진을 보았다. 두 번째로, 연구참여자들이 연구 자극 화면을 보고 제일 많이 제시된 얼굴 정체성이 무엇인지 응답하였다. 스플래시 도형 최빈값 판단 과제의 절차 또한 얼굴 정서 최빈값 판단 과제의 절차와 두 가지 부분을 제외하고는 모두 동일하였다. 첫 번째로, 연구참여자들이 얼굴 사진이 아닌 스플래시 도형을 보았다. 두 번째로, 연구참여자들이 연구 자극 화면에서 제일 많이 제시된 스플래시 도형이 무엇인지 응답하였다.

5. 분석 방법

본 연구에서는 세 가지 과제(얼굴 정서 최빈값 판단 과제, 얼굴 정체성 최빈값 판단 과제, 스플래시 도형 최빈값 판단 과제) 모두를 수행한 연구참여자의 응답 정확도(accuracy)를 수집하여 분석하였다. 결과 분석 과정에서 다음의 절차를 통해 분석에서 제외할 연구참여자를 결정하였다. 먼저 과제 수행 정확도가 이항분포의 단측 90% 신뢰구간보다 낮아 우연 수준(chance level)을 넘기지 못하는 연구참여자의 데이터를 제거하였다. 우연 수준을 넘기지 못한 연구참여자의 수행은 과제를 성실하게 수행하지 않고 무작위로 응답한 결과일 수 있으므로 분석에서 제외하였다. 다음으로, 이렇게 제거된 사람 중, 세 개의 과제 모두에서 각 시행(trial)의 난이도와 연구참여자의 각 시행 정답률의 상관(trial-by-trial correlation)이 .15보다 높은 사람을 분석에 재포함하였다. 이러한 연구참여자의 경우 쉬운 문제를 맞고 어려운 문제를 틀린 것이므로 성실하게 과제를 수행했지만 최빈값 판단 능력이 좋지 못해 수행이 낮은 것으로 볼 수 있다. 여기서 각 시행의 난이도는 우연 수준을 넘긴 연구참여자의 해당 시행 정답률이다.

위의 절차를 통과한 데이터를 기반으로 JASP 통계 프로그램(JASP Team,

2022)을 활용하여 Bayesian 상관분석을 사용하였다. 얼굴 정서 최빈값 판단 과제 정확률과 얼굴 정체성 최빈값 판단 과제 정확률에서 스플래시 도형 최빈값 판단 과제의 정확률을 통제하고, 얼굴 정서 최빈값 판단 과제 정확률과 얼굴 정체성 최빈값 판단 과제 정확률 간의 편상관을 도출하였다. 만약 스플래시 도형 최빈값 판단 과제 수행이 통제된 얼굴 정서 최빈값 판단 과제 수행과 얼굴 정체성 최빈값 판단 과제 수행 간의 정적인 편상관이 도출된다면, 얼굴 정서 판단 능력과 얼굴 정체성 판단 능력 간에 개별 얼굴 지각과 관련된 능력이 공통으로 사용된다고 해석할 수 있다.

Ⅲ. 연구 결과

총 83명의 연구참여자가 세 가지 과제를 모두 수행하였는데, 3명의 데이터가 결과 분석 절차를 넘기지 못하여 분석에서 제외되었다. 최종 결과 분석에는 연구참여자 80명($M = 35.25$, $SD = 12.05$)의 데이터만이 사용되었고, 이 중 40명은 여성 연구참여자, 나머지 40명은 남성 연구참여자였다. 80명의 연구참여자 중 62명(77.5%)은 자신을 백인이라고 보고하였으며, 9명(11.25%)은 흑인이나 아프리카계 미국인으로, 4명(5%)은 아시아인이라고 보고하였다. 나머지 5명 중 4명(5%)은 선택지에 자신의 인종이 없다고 응답하였으며, 1명(1.25%)은 자신의 인종을 보고하고 싶지 않다고 응답하였다.

과제 수행 정확도 점수의 평균과 표준편차는 다음과 같다: 얼굴 정서 최빈값 판단 과제($M = .669$, $SD = .075$), 얼굴 정체성 최빈값 판단 과제($M = .643$, $SD = .085$), 스플래시 도형 최빈값 판단 과제($M = .684$, $SD = .107$). 먼저 Bayesian 상관분석을 이용해서 세 과제 간의 상관관계를 분석하였다. 얼굴 정서 최빈값 판단 정확도는 얼굴 정체성 최빈값 판단 정확도와 정적인 상관이 있었다($r = .539$, 95% C.I. = [0.354, 0.672], $BF_{10} = 65300.503$). Bayesian 상관분석에서의 BF_{10} 은 상관관계의 방향(정적 혹은 부적)과 관계없이 상관이 있는 모델의 가능도(likelihood)를 상관이 없는 모델의 가능도와 비교한 값이다. 즉, 65300.503의 베イズ 팩터는 수집한 데이터를 기반으로 살펴보았을 때 얼굴 정서 최빈값 판단 정확도와 얼굴 정체성 최빈값 판단 정확도 간의 상관이 있는 모델이 맞을 가능성이 두 정확도 간의 상관이 없는 모델에 비해 65300.503배 높다는 의미이다. BF_{10} 가 100이 넘을 경우 극도로 강한 증거라는 의미이며, 30이 넘을 경우 매우 강한 증거, 10을 넘으면 강한 증거, 3을 넘을 경우 보통의 증거임을 의미하며, 1에서 3 사이의 값은 일화적인(anecdotal) 증거로 볼 수 있다(Lee & Wagenmakers, 2014). 얼굴 정서 최빈값 판단 정확도와 얼굴 정

체성 최빈값 판단 정확도 사이의 정적 상관은 두 과제의 수행 사이에 공통적인 변산이 있음을 나타낸다. 얼굴 처리와 관련된 능력을 공통으로 사용하기 때문에 이러한 공통 변산이 나타났을 수도 있고, 과제 특성과 같은 다른 요인으로 인해 이러한 공통 변산이 나타났을 수도 있다.

스플래시 도형 최빈값 판단 과제의 정확도는 얼굴 정서 최빈값 판단 과제의 정확도와 정적 상관이 있었다($r = .492$, 95% C.I. = [0.299, 0.636], $BF_{10} = 5347.862$). 또한 스플래시 도형 최빈값 판단 과제의 정확도는 얼굴 정체성 최빈값 판단 과제의 정확도와도 정적 상관이 있었다($r = .572$, 95% C.I. = [0.395, 0.697], $BF_{10} = 4.911 \times 10^5$).

본 연구에 사용된 세 과제는 아홉 개의 이미지 중에서 최빈값 이미지를 찾아 응답하는 동일한 패러다임을 사용했다. 스플래시 도형 최빈값 판단 과제와 얼굴 이미지의 최빈값 판단 과제들은 자극의 특성을 공유하지 않으므로 제시된 이미지들의 최빈값을 추정하고 세 가지의 선택지 중 그 최빈값이 무엇인지 선택하는 행동이 과제 간의 정적 상관에 크게 기여했을 것이다. 과제들이 공유하는 이러한 특성은 얼굴 정서 최빈값 판단 과제와 얼굴 정체성 최빈값 판단 과제 간의 상관에도 영향을 주었을 것이다.

본 연구에서는 상상불 과제의 특성이 아니라 자극의 특성으로 인한 과제 수행의 변산을 연구하기 위해서, 얼굴 정서 최빈값 판단 수행 점수와 얼굴 정체성 최빈값 판단 수행 점수에서 스플래시 도형 최빈값 판단 과제 수행 점수로 예측할 수 있는 변산을 제거한 후 편상관을 분석하였다. 스플래시 도형의 상상불 판단에 사용된 능력을 통제한 뒤에도 두 얼굴 최빈값 판단 과제 간 정적 편상관이 발견되었다($r = .361$, 95% C.I. = [0.149, 0.531], $BF_{10} = 27.955$). 이러한 정적 편상관은 최빈값 판단 과제의 특성과는 관련이 없는, 두 과제의 공통된 자극 특성(얼굴 정서와 얼굴 정체성)에서 기인한 것으로 볼 수 있다.

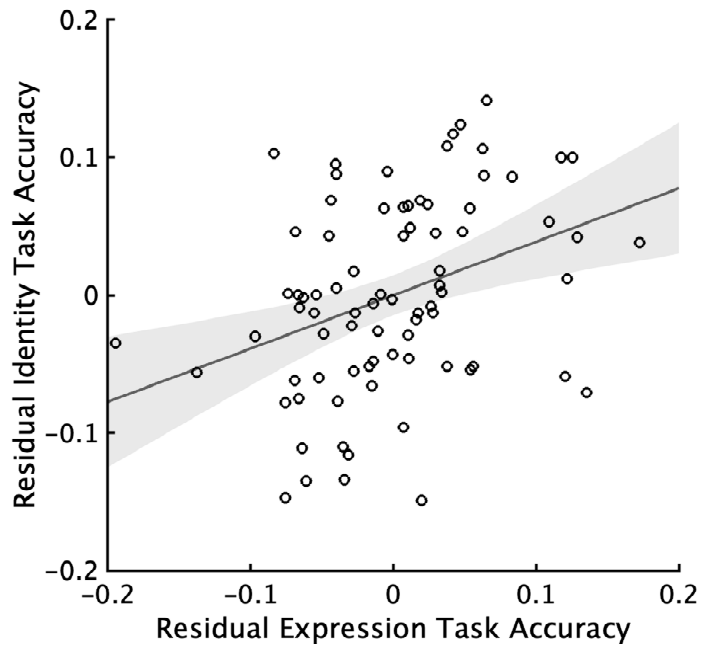


그림 3. 얼굴 정서 최빈값 판단 과제의 정확도 잔차와 얼굴 정체성 최빈값 판단 과제의 정확도 잔차의 관계를 나타내는 산포도. 얼굴 정서 최빈값 판단 과제 수행 점수와 얼굴 정체성 판단 과제 수행 점수에서 스플래시 도형 최빈값 판단 과제 수행 점수를 통제한 잔차를 표시함.

IV. 논의

본 연구는 얼굴 정서 앙상블 판단 능력과 얼굴 정체성 앙상블 판단 능력이 공통된 기제를 사용하는지, 그리고 공통된 기제 중에서 물체의 종류와 상관없이 앙상블 판단에 사용되는 공통 기제를 제거한다면 어떤 기제가 남는지 알아보았다. 앙상블 판단의 대상이 되는 특성 별로 각각 앙상블 판단 능력을 측정하기 위하여 총 세 가지의 과제를 사용하였다: 얼굴 정서 최빈값 판단 과제, 얼굴 정체성 최빈값 판단 과제, 스플래시 도형 최빈값 판단 과제. 얼굴 정서 최빈값 판단 과제와 얼굴 정체성 최빈값 판단 과제 모두에서 공통 앙상블 능력이 사용되므로, 스플래시 도형 최빈값 판단 과제를 사용하여 공통 앙상블 능력을 통제했다. 그 결과, 얼굴 정서 앙상블 판단 수행과 정체성 앙상블 판단 수행 간의 정적인 편상관을 발견하였는데, 이는 개별 얼굴의 특성(정서와 정체성) 지각에 사용되는 능력이 앙상블 판단에도 사용됨을 시사한다.

두 얼굴 앙상블 과제 간의 편상관이 얼굴의 특성 지각을 반영한다는 해석은 세 개의 앙상블 과제에 공통으로 영향을 주는 여러 요인이 모두 통제되었다는 가정에 기반한다. 그 예로, 연구 자극을 보고 아홉 개의 이미지 중에서 최빈값을 찾는 능력이나 과제 수행 동기와 같은 능력이 세 개의 앙상블 과제에 모두 영향을 주는 공통 요인이 된다. 본 연구에서는 스플래시 도형 최빈값 판단 과제의 수행을 이용해서 이런 요인들을 모두 통제했다. 특히 여러 물체의 최빈값을 판단하는 데 사용하는 능력을 통제했으므로, 얼굴 앙상블 판단 과제 수행 사이의 편상관은 얼굴 하나하나의 자극 특성을 처리하는 능력의 영향으로 해석할 수 있다.

그렇다면 얼굴의 서로 다른 특성 판단, 특히 정체성 판단과 정서 판단 사이에 상관관계가 있을까? 선행연구에 따르면, 얼굴 정체성을 잘 판단하는 사람이 얼굴 정서를 더 잘 판단한다(Palermo et al., 2013). Palermo와 동료들

(2013)은 캠브리지 얼굴 기억 검사(Cambridge Face Memory Test; CFMT) 수행과 정서-매칭 과제(emotion-matching task) 수행 간의 관계를 살펴보았다. CFMT는 세 가지 각도에서 촬영한 여섯 명의 얼굴 사진을 익힌 후, 익힌 얼굴을 재인하는(다른 얼굴로부터 구분하는) 능력을 측정한다. 정서-매칭 과제는 연구참여자의 얼굴 정서 지각을 측정하는 과제이며, 세 명의 얼굴 중 두 명의 얼굴 표정과 혼자서 다른 표정을 짓고 있는 얼굴 사진을 선택하는 문항으로 이루어져 있다. Palermo와 동료들(2013)은 CFMT 수행과 정서-매칭 과제 수행 간의 정적 상관관계가 있음을 보여주었다($r = .40$). 이러한 결과는 개별 얼굴에 대한 정서 판단 수행과 정체성 판단 수행이 16%의 변산을 공유하고 있음을 의미하며, 정서와 정체성 지각 능력들 사이에 어느 정도 공유하는 기제가 있음을 보여준다. 본 연구에서는 공통 상상블 능력을 통제된 후 얼굴 정서와 정체성 상상블 판단 능력 간의 공유 변산이 13.03%였고, 이는 Palermo와 동료들(2013)에서 도출한 공유 변산의 크기와 유사하다.

하나의 인지 과제를 수행하기 위해서는 다양한 능력이 필요하다. 예를 들어, 얼굴 정서 최빈값 판단 과제를 성공적으로 수행하기 위해서는 각 얼굴에 나타난 다양한 표정들을 구별할 수 있어야 하며, 여러 정서 중 어떤 정서가 가장 많이 제시되었는지 추정할 수 있어야 한다. 그뿐만 아니라 연구에 참여하고자 하는 동기와 작업 기억과 같은 다양한 요인이 얼굴 정서 최빈값 과제 수행 점수에 영향을 미칠 수 있다. 마찬가지로, 얼굴 정체성 최빈값 판단 과제와 스플래시 도형 최빈값 판단 과제를 수행하기 위해 여러 개의 다른 능력들이 요구된다. 이 능력 중 일부는 세 과제 모두에서 사용되기도 한다. 예를 들어, 여러 개의 이미지가 동시에 제시되었을 때 가장 많이 제시된 이미지가 무엇인지 판단하는 능력이나 많은 시행을 맞추고자 하는 연구 참여 동기는 본 연구의 모든 과제 수행에 영향을 미칠 것이다.

많은 선행 연구는 상상블 과제 수행 사이의 공통 변산에 기여하는 공통 양

상블 능력(혹은 시각적 통계 처리 능력)이 존재하는지에 초점을 맞추어 연구를 수행했다. 이런 선행 연구에서는 서로 다른 양상블 과제 수행 사이의 공유하는 변산을 추정된 뒤에 다른 요인으로 설명될 수 있는 변산을 제거하였다(예, 개별 물체 지각 능력, 기억력, 연구 참여 동기 수준, 등). 본 연구에서는 선행 연구의 관심사였던 공통적인 통계적 처리 능력으로 인한 행동 변산을 통제하고, 선행 연구에서 통제하고자 했던 자극의 특성으로 인한 행동 변산에 초점을 맞추었다. 본 연구에서 사용한 세 개의 과제는 이미지의 종류와 최빈값 판단의 대상이 되는 특성 외에는 모든 면에서 동일했다. 따라서 최빈값을 판단하는 능력, 연구 참여 동기, 기억력과 같이 연구 관심사가 아닌 모든 요인은 편상관 분석 과정에서 제거되었다고 기대할 수 있다. 따라서 두 얼굴 양상블 판단 과제 사이의 편상관은 얼굴이라는 자극의 특성(정서와 정체성)에서 비롯되었다고 해석할 수 있다.

얼굴의 특성이 아닌 연구참여자의 특성이 두 얼굴 양상블 과제의 편상관을 만들었을 가능성은 없을까? 예를 들어 타인종 효과가 두 얼굴 양상블 판단 과제 간의 상관에 기여했을 가능성이 존재하지는 않을까? 얼굴 정서 최빈값 판단 과제는 백인의 얼굴로 구성된 Karolinska Directed Emotional Faces(KDEF, Lundqvist et al., 1988)에서 다양한 표정(두려움, 화남, 역겨움, 행복함, 슬픔, 놀람)의 얼굴을 선택하여 사용하였고, 얼굴 정체성 최빈값 판단 과제에서는 Chicago Face Database(CFD, Ma et al., 2015)에서 중립 표정의 다양한 인종(동양인, 흑인, 라틴아메리카인, 백인) 얼굴을 선택하여 사용하였다. 만약 타인종 효과가 발생하였다면, 얼굴 정서 최빈값 판단 과제의 수행 점수에 영향을 미쳤을 가능성이 있다. 동양권에서 자란 연구참여자의 경우 서양권에서 자란 연구참여자에 비해 백인의 얼굴에서 정서 최빈값을 잘 판단하지 못했을 수도 있기 때문이다. 얼굴 정체성 최빈값 판단 과제의 경우 네 가지 다른 인종으로 이루어진 CFD의 얼굴 이미지를 사용하였기 때문에 동양인과

서양인이 타인종 효과로 인해 받는 영향에 차이가 없었을 것이다. 동양권에서 자란 연구참여자는 서양권에서 자란 연구참여자보다 동양인 얼굴 이미지가 포함된 시행에서는 더 높은 수행을, 백인 얼굴 이미지가 포함된 시행에서는 낮은 수행을 보여 전반적인 수행에서는 차이가 크지 않았을 것으로 예상된다. 타인종 효과가 두 얼굴 양상블 과제 중 얼굴 정체성 최빈값 판단 과제에만 영향을 줄 수 있었기 때문에 얼굴 정서 최빈값 판단 과제와 얼굴 정체성 최빈값 판단 과제 간의 상관에는 기여하지 않았을 것이다.

얼굴 정서 최빈값 판단 과제에서는 최빈값 이미지가 연구 자극 화면에 총 다섯 개 제시되었고, 얼굴 정체성 최빈값 판단 과제와 스플래시 도형 최빈값 판단 과제의 경우 최빈값 이미지가 네 개, 또는 다섯 개 제시되었다. 개인차 연구에서는 실험 조건 간의 점수 차이를 보는 것이 아니기 때문에 같은 연구참여자가 수행한 다른 두 과제 사이의 점수 차이가 아니라, 각각의 과제에서 연구참여자들 간의 상대적 순위를 이용하여 통계분석을 수행한다. 따라서 과제마다 다른 방법으로 자극을 조작하더라도 각 과제 안에서 연구참여자들의 상대적 순위가 유지되면 문제가 없다. 최빈값 이미지가 항상 다섯 개 제시되는 정서 최빈값 과제가 다른 과제보다 쉬웠더라도 모든 연구참여자에게 다 쉬웠을 것이므로 참여자의 상대적 순위에는 영향이 없다. 본 연구에서는 과제에 따라 최빈값 자극의 수를 다르게 조작하였기 때문에 바닥 효과나 천장 효과로 인해 참여자의 상대적 순위가 잘 드러나지 않게 되는 것을 방지할 수 있었다. 각 과제 수행의 평균을 비교해 보면, 얼굴 정서 최빈값 판단 과제의 평균은 .669로 평균이 .643인 얼굴 정체성 최빈값 판단 과제와 평균이 .684인 스플래시 도형 최빈값 판단 과제와 크게 다르지 않았다.

본 연구에서 사용한 최빈값 판단 과제를 포함하여 많은 양상블 판단 과제들은 작업 기억에 저장할 수 있는 네 개(Luck & Vogel, 1997) 이상의 자극을 이용하는 판단을 요구한다. 그 때문에 참여자들이 소수의 자극만을 이용해서

과제를 수행하는 전략을 사용한다는 주장이 제기될 수 있다. 그 예로, Myczek와 Simons(2008)는 여러 개의 자극이 제시되면 그중 하나의 자극(이미지)만 기억한 후, 응답 화면에서 기억한 자극(이미지)을 이용하여 응답을 선택하는 서브샘플링(subsampling) 전략을 제안했다. 만약 본 연구에서 연구참여자들이 이러한 서브샘플링 전략을 사용했다면, 얼굴 하나를 기억하는 과제를 수행한 셈이므로 Palermo와 동료들(2013)의 연구에서와 같이 정체성과 정서 과제의 상관을 발견하는 것이 당연할 것이다. 이에 본 연구에서는 추가로 많은 연구참여자가 서브샘플링 전략을 사용했을 가능성을 배제하기 위해서 하나의 이미지를 기반으로 최빈값 판단을 내렸을 때 기대할 수 있는 최대 정확도를 분석했다.

먼저 얼굴 정서 최빈값 판단 과제의 경우 다음과 같이 분석할 수 있다. 얼굴 정서 최빈값 판단 과제의 자극 화면은 다섯 개의 동일한 얼굴(응답 화면에서 정답으로 제시되는 선택지)과 두 개의 동일한 얼굴(응답 화면에서 오답으로 제시되는 선택지), 그리고 응답 화면에 제시되지 않는 두 개의 서로 다른 얼굴 이미지로 구성된다. 만약 연구참여자가 하나의 얼굴을 무작위로 선택하여 외운다면, 응답 화면에서 가장 많은 수로 제시된 얼굴을 선택할 확률은 5/9이다. 즉, 정답을 고를 확률이 5/9가 된다. 오답 얼굴을 외워서 해당 시행을 틀릴 확률은 2/9이다. 만약 응답 화면에서 선택지로 제시되지 않는 얼굴을 외운다면, 연구참여자는 제시된 세 가지의 선택지 중 하나를 무작위를 선택하여 응답하여야 하기에 정답을 고를 확률은 1/3이다. 결과적으로 연구참여자가 하나의 이미지를 외워서 최빈값 판단을 했다면 기대할 수 있는 최대의 정확률은 63.0%이고 작업 기억에 저장하지 못하거나 실수할 확률을 고려하면 63.0%보다 낮은 정확률을 보일 것이다. 하지만 실제 연구참여자의 평균 정확률은 66.9%로 하나의 이미지를 외워 응답할 때 기대할 수 있는 최대 정확률보다 높았다. 같은 방법으로 계산하면 얼굴 정체성 최빈값 과제와 스플래시 도형

최빈값 과제에서 하나의 자극만을 이용해서 응답할 때 최대 59.3%의 정확률을 달성할 수 있다. 각 과제에 대한 실제 연구참여자들의 평균 정확률은 각각 64.3%와 68.4%로 서브샘플링 전략을 사용하였을 때의 최대 정확률보다 높았다. 따라서 다수의 연구참여자가 하나의 자극만을 이용해서 응답하는 전략을 사용하지 않았을 것으로 생각할 수 있다.

본 연구를 통해 얼굴 앙상블 판단에도 얼굴 하나의 특성을 볼 때 활용되는 능력이 사용되는 것을 발견하였다. 몇몇 연구는 여러 개의 물체가 동시에 제시될 때, 사람들이 해당 물체들을 텍스처처럼 지각한다고 주장하였다(Dakin & Watt, 1997; Field et al., 1993). 앙상블 지각 연구에서는 여러 개의 물체가 동시에 제시되었을 때 개별 물체를 하나씩 보지 못한다는 결과도 있는데 (Haberman & Whitney, 2011), 이 또한 여러 개의 물체를 텍스처처럼 지각한다는 주장과 동일한 맥락이다. 본 연구에서는 얼굴 앙상블 지각을 하는 동안 개별 얼굴을 지각할 때와같이, 얼굴 정서 지각 능력과 얼굴 정체성 지각 능력 간의 공유하는 변산이 사용된다는 결과를 발견하였다. 이러한 결과는 얼굴 앙상블 지각은 텍스처 지각만으로는 완전히 설명하기 어려우며, 개별 물체 지각에 사용되는 인지 능력이 사용된다는 것을 시사한다.

참 고 문 헌

- Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science*, *19*(4), 392-398.
- Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157-162.
- Balas, B., & Saville, A. (2017). Hometown size affects the processing of naturalistic face variability. *Vision Research*, *141*, 228-236.
- Cha, O., Blake, R., & Gauthier, I. (2021). The role of category- and exemplar-specific experience in ensemble processing of objects. *Attention, Perception, & Psychophysics*, *83*(3), 1080-1093.
- Cha, O., Blake, R., & Gauthier, I. (2022). Contribution of a common ability in average and variability judgments. *Psychonomic Bulletin & Review*, 1-8.
- Chang, T.-Y., & Gauthier, I. (2022). Domain-general ability underlies complex object ensemble processing. *Journal of Experimental Psychology: General*, *151*(4), 966-972.
- Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*(4), 393-404.
- Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 94.
- Dakin, S. C., & Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision Research*, *37*(22), 3181-3192.

- Derntl, B., Seidel, E.-M., Kainz, E., & Carbon, C.-C. (2009). Recognition of emotional expressions is affected by inversion and presentation time. *Perception, 38*(12), 1849-1862.
- Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia, 44*(4), 576-585.
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision Research, 33*(2), 173-193.
- Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience, 2*(6), 568-573.
- Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 56*(4), 544-549.
- Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology, 17*(17), R751-R753.
- Haberman, J., & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance, 35*(3), 718.
- Haberman, J., & Whitney, D. (2011). Efficient summary statistical representation when change localization fails. *Psychonomic Bulletin & Review, 18*, 855-859.

- JASP Team. (2022). JASP (Version 0.16.4) [Computer Software]. Available at <http://jasp-stats.org/>
- Jayaraman, S., Fausey, C. M., & Smith, L. B. (2015). The faces in infant-perspective scenes change over the first year of life. *PLoS One*, *10*(5), e0123780.
- Kacin, M., Gauthier, I., & Cha, O. (2021). Ensemble coding of average length and average orientation are correlated. *Vision Research*, *187*, 94-101.
- Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science*, *18*(12), 1084-1089.
- Khayat, N., & Hochstein, S. (2018). Perceiving set mean and range: Automaticity and precision. *Journal of Vision*, *18*(9), 23-23.
- Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279-281.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122-1135.
- Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and

- other race. *Journal of Personality and Social Psychology*, *13*(4), 330.
- McGugin, R. W., McKeeff, T. J., Tong, F., & Gauthier, I. (2011). Irrelevant objects of expertise compete with faces during visual search. *Attention, Perception, & Psychophysics*, *73*, 309–317.
- McGugin, R. W., Newton, A. T., Gore, J. C., & Gauthier, I. (2014). Robust expertise effects in right FFA. *Neuropsychologia*, *63*, 135–144.
- McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Gauthier, I. (2012). The Vanderbilt Expertise Test reveals domain-general and domain-specific sex effects in object recognition. *Vision Research*, *69*, 10–22.
- McKeeff, T. J., McGugin, R. W., Tong, F., & Gauthier, I. (2010). Expertise increases the functional overlap between face and object perception. *Cognition*, *117*(3), 355–360.
- Myczek, K., & Simons, D. J. (2008). Better than average: Alternatives to statistical summary representations for rapid judgments of average size. *Perception & Psychophysics*, *70*, 772–788.
- Palermo, R., O'Connor, K. B., Davis, J. M., Irons, J., & McKone, E. (2013). New tests to measure individual differences in matching and labelling facial expressions of emotion, and their association with ability to recognise vocal emotions and facial identity. *PloS One*, *8*(6), e68126.
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, *4*(7), 739–744.
- Richler, J. J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology*, *3*, 553.

- Rossion, B., Curran, T., & Gauthier, I. (2002). A defense of the subordinate-level expertise account for the N170 component. *Cognition, 85*(2), 189-196.
- Schönbrodt, F. D., Wagenmakers, E. J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods, 22*(2), 322-339.
- Sugden, N. A., & Moulson, M. C. (2019). These are the people in your neighbourhood: Consistency and persistence in infants' exposure to caregivers', relatives', and strangers' faces across contexts. *Vision Research, 157*, 230-241.
- Sun, J., & Chong, S. C. (2020). Power of averaging: Noise reduction by ensemble coding of multiple faces. *Journal of Experimental Psychology: General, 149*(3), 550.
- Sun, J., & Gauthier, I. (2021). Car expertise does not compete with face expertise during ensemble coding. *Attention, Perception, & Psychophysics, 83*, 1275-1281.
- Sunday, M. A., & Gauthier, I. (2018). Face expertise for unfamiliar faces: A commentary on Young and Burton's "Are we face experts?". *Journal of Expertise, 1*(1), 35-41.
- Sweeny, T. D., & Whitney, D. (2014). Perceiving crowd attention: Ensemble perception of a crowd's gaze. *Psychological Science, 25*(10), 1903-1913.
- Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science, 12*(1), 43-47.
- Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of

- the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, *93*(1), B1-B9.
- Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception*, *9*(4), 483-484.
- Turati, C., Bulf, H., & Simion, F. (2008). Newborns' face recognition over changes in viewpoint. *Cognition*, *106*(3), 1300-1321.
- Valentine, T., Lewis, M. B., & Hills, P. J. (2016). Face-Space: A unifying concept in face recognition research. *Quarterly Journal of Experimental Psychology*, *69*(10), 1996-2019.
- Walker, P. M., & Hewstone, M. (2006). A perceptual discrimination investigation of the own race effect and intergroup experience. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *20*(4), 461-475.
- Watamaniuk, S. N., Sekuler, R., & Williams, D. W. (1989). Direction perception in complex dynamic displays: The integration of direction information. *Vision Research*, *29*(1), 47-59.
- Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology*, *69*, 105-129.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*(1), 141.
- Yörük, H., & Boduroglu, A. (2020). Feature-specificity in visual statistical summary processing. *Attention, Perception, & Psychophysics*, *82*, 852-864.
- Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*(6), 747-759.

ABSTRACT

Face recognition plays a role in ensemble judgments of facial features

Soo Jin Kim
Department of Psychology
Graduate School of
Sungshin University

Ensemble judgments about faces (e.g., making judgments about average facial expression) were less influenced by expertise effects such as the face inversion effect. For instance, participants' mean emotion judgment performance for a crowd of inverted faces was better than emotion judgment performance with single inverted faces (Sun & Chong, 2020). These studies suggest that judgments about an ensemble of faces are different from judgments about single faces. The present study used an individual differences approach to investigate the influence of face recognition ability on the ensemble judgments of facial features. Specifically, the correlation between ensemble judgment abilities for facial expression and facial identity was analyzed, after controlling for ensemble judgment ability for novel objects. Participants completed three tasks: facial expression mode (i.e., the most frequent instance in a group) judgment task, facial identity mode judgment task, and splash shape mode judgment task. Random silhouette shapes were

used in the splash shape mode judgment task. The purpose of the splash shape mode judgment task was to measure and control participants' ensemble judgment ability common to all categories of objects. In each task, nine face/shape images were shown for a short period of time on a virtual 3-by-3 grid, with random jitters in their positions (stimulus display). After a short interval, three face/shape images were presented as multiple-choice options (response display). In the response display, participants selected the most frequently presented image in the stimuli display and they were given non-guided feedback after the response. The performance of the facial expression mode judgment task and the facial identity mode judgment task was positively correlated after removing the influence of the abilities used in the splash shape mode judgment task. This result suggests shared behavioral variance between ensemble judgments for different facial features (expression and identity) that originates not from common ensemble judgment ability, but from abilities related to the processing of facial features.

Keywords : ensemble judgment, face recognition, facial expression, facial identity, individual differences study