

사회불안의 사회적 상황에 대한 해석편향: 얼굴표정과 음성의 혼합자극을 활용하여[†]

유 세 은

동국대학교 일산병원
정신건강의학과 수련생

최 지 수

가톨릭대학교 심리학과
박사과정

양 재 원[‡]

가톨릭대학교 심리학과
교수

기존 사회불안 집단의 해석편향 연구들은 주로 시나리오로 제시되는 상황에 대한 해석이나 제시되는 얼굴 자극의 정서가를 판단하는 방식으로 진행되었다. 본 연구에서는 사회적 상황의 시나리오에 더해 얼굴과 음성 자극을 동시에 제시하는 다감각적 정보를 사용하여, 이를 처리하는 과정에서 나타나는 사회불안의 하위 요인 간 해석편향의 양상에 대해 탐색해 보고자 하였다. 89명의 대학(원)생이 연구에 참여하였고, 사회불안의 인지적 및 경험적 측면을 측정하는 척도를 실시하고 실험을 진행하였다. 실험에서는 모호한 상황의 시나리오가 먼저 제시되었고, 이후 얼굴표정 사진과 음성이 동시에 나타났다. 시각과 청각 자극의 정서로는 행복과 중립을 사용하였으며, 총 4가지 조건이었다. 실험 참여자들은 사회적 상황에서 제시되는 시·청각의 자극이 얼마나 긍정적 혹은 부정적으로 느껴졌는지 시각적 아날로그 척도(VAS)로 평정하였다. 연구 결과, 사회불안의 인지적 측면인 부정적 평가에 대한 두려움이 높을수록 중립 음성이 제시되는 조건에서 자극의 정서를 부정적으로 해석하는 편향을 보였다. 반면, 사회불안의 경험적 측면인 사회적 상호작용에 대한 불안이 높을수록 행복 얼굴표정이 제시되는 조건에서 덜 긍정적으로 평가하는 경향을 보였다. 이는 사회불안의 해석편향이 단일한 과정이 아님을 보여준다. 연구 결과를 바탕으로 연구의 의의와 심리치료에 대한 제언을 제시하였다.

주요어: 사회불안, 모호한 상황, 시청각 자극, 다감각 처리, 해석편향

[†] 이 논문의 일부 내용은 2022년 한국심리학회 연차학술대회에서 포스터 발표되었음.

[‡] 교신저자(Corresponding author): 양재원, 가톨릭대학교 심리학과 교수, (14662) 경기도 부천시 원미구 지봉로 43, E-mail : jwyang@catholic.ac.kr

 Copyright ©2025, The Korean Health Psychological Association. 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.

사회적 상호작용은 인간의 삶에서 필수불가결하다. 사람들 간의 소통에는 맥락이 존재하고, 그 맥락 속에서 사람들은 행복, 기쁨, 슬픔, 분노 등 다양한 정서를 느낀다. 일반적으로 사회적 맥락은 모호하고 불확실하기에 개인은 항상 분명한 정서를 드러내는 상황만 마주하지 않는다. 따라서 맥락 속에서 사람들은 각자 다른 정서를 인식하곤 한다.

사회불안(social anxiety)이 높은 사람들은 사회적 상황이나 수행 상황에 노출될 때 극도의 불안과 공포를 느낀다(American Psychiatric Association [APA], 2013). 특히 이들은 정서적 단서나 맥락이 불명확한 상황에서 그 상황을 부정적으로 혹은 위협적으로 해석하는 정보처리적 편향을 보인다(Clark & McManus, 2002; Heinrichs & Hofmann, 2001). 이러한 해석편향은 사회불안장애의 발생과 유지에 매우 중요한 역할을 한다(Clark & Wells, 1995; Rapee & Heimberg, 1997).

사회불안은 타인에게 부정적으로 평가될 것을 염려하는 인지적 측면과 실제 상호작용 자체에서 불안을 느끼는 경험적 측면으로 구분된다(Gomez & Watson, 2017). 이러한 두 측면은 각기 다른 정보 처리 과정의 편향과 관련될 수 있다. 사회불안의 인지적 요소인 부정적 평가에 대한 두려움은 사회적 상황의 의미를 곱씹고 반추하는 사후처리(post-event processing) 과정과 연관되며 (Abbott & Rapee, 2004), 이 과정에서 명확한 단서보다는 해석의 여지가 많은 모호한 정보가 해석편향을 촉발할 가능성이 높다. 반면, 상호작용 자체에서 겪는 불안은 대면 상황에서 실시간으로 유발되는 불편감과 연결된다(Clark & Wells,

1995). 대면 상황에서는 비교적 생생하고 명확한 단서가 제시되지만, 동시에 이러한 단서가 잠재적인 위협 신호로 여겨지는 인지 편향이 발생한다. 하지만 사회불안을 구성하는 요소 간의 차별적인 역할을 구분한 연구들은 적다.

초기 사회불안 장애의 해석편향 연구는 주로 모호한 사회적·비사회적 상황을 담은 시나리오를 제시하거나 혹은 모호한 사회적 상황에 대한 문장완성과제를 실시하는 방법으로 이루어졌다. 이런 연구들에서 사회불안 집단에서는 부정적 해석편향을 보이는 경향이 확인되었다(Amin et al., 1998; Huppert et al., 2007). 다만 언어적으로 제시되는 시나리오는 간접적인 사회적 자극으로, 사람마다 실제로 표상하는 바가 다르다는 비판이 제기되었다(Benito-Ruiz et al., 2022). 이에 따라 생태학적 타당도가 상대적으로 더 높은 얼굴표정 자극을 이용하여 해석편향 연구가 진행되었다. 그 결과, 사회불안 집단은 비교 집단과는 다르게 모호한 얼굴 표정의 정서를 더 부정적이라고 인식하는 경향이 확인되었다(Yoon & Zinbarg, 2007, 2008). 또한 여러 얼굴들을 순차적으로 혹은 동시에 제시하고 그 얼굴 군집의 평균 정서를 추정하게 하면, 사회불안이 높을수록 더 부정적으로 지각하였다(이대현, 백용매, 2013; 이하영 외, 2020; Bronfman et al., 2018; Yang et al., 2013).

이렇듯 언어 자극이나 시각 자극을 이용한 연구들은 반복해서 사회불안이 높은 사람들이 사회적 상황을 부정적으로 해석하는 편향을 가지고 있음을 보여주었다. 최근의 한 메타분석 연구는 사회불안과 부정적 해석편향 간에는 임상적으로 유의한 수준의 상관관계를 가진다는 것을 밝혔다 (Chen et al., 2020). 그러나 이를 선행 연구들은

여전히 시각이라는 하나의 감각 양식(modality)에만 초점을 두고 있다는 한계를 갖는다. 일상의 사회적 상황에서 얼굴표정과 같은 하나의 시각 자극만 독립적으로 나타나는 경우는 드물다 (Peschard et al., 2014). 많은 사람들은 타인과 상호작용을 하는 동안 그 상황이나 상대방의 말을 해석하기 위해서 다양한 감각 정보를 통합하여 사용하게 된다(Collignon et al., 2008). 따라서 여러 감각기관으로부터 온 정보들을 어떻게 통합하고 해석하는지에 따라 개인이 느끼는 생각과 정서가 달라진다. 그러므로 기존에 이루어졌던 단일 혹은 여러 얼굴만을 사용한 연구들의 결과를 일상생활에 적용하기 어려울 수 있다. 이러한 한계를 보완하고 생태학적 타당도를 높이기 위해서는, 일상에서 경험하는 다양한 감각을 이용한 연구가 필요하다(De Gelder & Bertelson, 2003).

일상적인 상호작용에서 흔히 사용되는 감각 기관은 얼굴표정과 음성으로, 이를 통해 한 개인의 정서 상태가 잘 드러난다. 우리는 타인의 표정을 보는 동시에 목소리를 들으며, 두 가지 이상의 감각 정보를 통합하여 상대방의 정서와 의도를 파악하고 적절한 반응이나 행동을 결정한다 (Campanella & Belin, 2007). 그리고 단일 감각만 있을 때보다 서로 다른 감각 정보가 결합할 때 더 정확하고 빠르게 인지한다(Stein & Stanford, 2008). 하지만 모든 감각 정보가 동일하게 처리되는 것은 아니다. 일반적으로 사람들은 외부 자극을 지각할 때 시각 정보를 청각 정보보다 우선하여 처리하는 경향이 있다(Colavita, 1974). 이런 시각 우세 효과는 정서 정보 처리에도 동일하게 나타난다. 아울러 정서가(emotional valence)의 일치 여부에 따라 정보 처리는 달라진다. 정서가가 일

치하는 정보가 동시에 주어지면, 각 정보가 서로를 상호 보강하면서 의미가 명확해지고 처리 효율이 높아지는 중복 효과(redundancy effect)가 나타난다(Collignon et al., 2008; Giard & Peronnet, 1999). 반면, 서로 상반되는 감정 자극이 제시되면 감각 간의 갈등이 발생하여 의미 해석에 더 많은 인지적 자원이 요구된다(Ethofer et al., 2006). 불일치하는 정보를 통합하는 과정에서 불확실성이 낮아 분명하다고 판단되는 감각에 더 많은 가중치를 부여하여 해석을 내린다(Ernst & Banks, 2002). 그럴 경우에는 일반적인 정서 인식 상황에서는 시각 정보가 음성 정보보다 더 명확하다고 여겨지기 때문에 얼굴 표정을 더 우선하여 처리하는 경향이 나타난다(Collignon et al., 2008; De Gelder & Vroomen, 2000).

한편, 음성 정보는 그 자체로 복합적인 사회적 단서이기 때문에 시각 정보에 비해 더 모호하게 인식된다. 음성은 억양, 강도, 속도, 리듬 등과 같은 준언어적인(paralinguistic) 정보를 전달한다. 이러한 정보는 미묘하고 다양한 음향 지표로 인해 쉽게 변화한다(Belin et al., 2004). 이에 동일한 내용의 발화라 할지라도 준언어적 요소가 약간 변동되는 것만으로도 의미가 완전히 달라질 수 있다(Cole, 2015). 이러한 음성의 다차원적 특성은 얼굴표정에 비해 더 높은 모호성을 갖게 한다 (Juslin & Laukka, 2003). 얼굴표정은 비교적 명확한 정서 범주로 인식될 수 있지만, 미묘한 음성 톤의 변화는 해석의 여지를 훨씬 많이 남기기 때문이다(Pittam & Scherer, 1993).

뿐만 아니라 외부로부터 들어온 감각 자극들은 맥락과 상호작용하며 처리된다(Barrett & Kensinger, 2010). 발화자가 누구인지, 어떤 상

황인지 등에 따라 동일한 표정이나 목소리도 다양하게 해석될 수 있다. 선행 연구에서 밝혀졌듯이, 얼굴표정을 통해 정서 인식을 할 때 물리적인 형태뿐만 아니라 관련된 맥락에 영향을 받는다(Aviezer et al., 2012). 더불어 맥락 정보는 얼굴 정보를 부호화하는 초기 과정과 해석 및 판단을 하는 후기 과정에 영향을 미친다(Aviezer et al., 2008). 음성 표현도 발화자에 따라 낮은 일관성을 가질 뿐만 아니라 말의 의도에 담긴 중요한 문법과 담화적 맥락에 대한 정보를 전달한다(Cole, 2015). 이처럼 시청각 자극과 맥락은 서로 영향을 주며 사회적 상호작용을 할 수 있도록 돋는다. 그러나 이 요소들이 대인간 상호작용에 매우 중요함에도 불구하고, 최근 얼굴표정 외에 다른 맥락 자극을 이용한 연구가 이루어지고 있지만(예: 최성영, 양재원, 2019; Wieser & Moscovitch, 2015), 사회적 맥락 내에서 얼굴과 음성의 다감각 양식의 자극을 활용하여 진행한 연구는 지금까지 찾아보기 어렵다.

상기한 바와 같이 선행 연구들은 주로 위협과 부정의 과잉해석 경향성, 즉 부정적인 해석편향에 초점을 맞춰왔다(Heimberg et al., 2014; Heinrichs & Hofmann, 2001; Stopa & Clark, 2000). 그러나 최근 연구들은 이러한 부정 편향만으로는 사회불안의 유지 과정을 온전히 설명하기 어렵다고 지적하며, 긍정적 해석편향의 부재 현상에 대해 주목하고 있다(Bean et al., 2023; Romano et al., 2020). 사회불안이 높은 사람들은 사회적 위협에는 민감하게 반응하면서도 긍정적 사회 정보에는 주의를 두지 않으며(Chen et al., 2002; Taylo et al., 2010, 2011), 행복 표정을 인식하는 데 더 많은 시간이 걸리거나(Silvia et al., 2006), 긍정적이

라고 판단하는 기준 자체가 높은 경향을 보인다(Yoon et al., 2014). 또한 타인의 긍정적 표정을 비웃음과 같은 부정적 의도로 해석하기도 한다(임효정, 양재원, 2021). 이처럼 사회불안은 부정적 자극에 대한 과잉 반응뿐만 아니라, 긍정적 자극을 온전히 수용하지 못하는 긍정 편향의 부족과도 깊은 연관이 있다(Alden et al., 2008; Constans et al., 1999; Weeks et al., 2016).

종합하면, 선행 연구들은 사회불안을 단일 차원으로 접근하거나 단일 감각 양식에 의존하는 한계가 있었고, 주로 중립적이거나 부정적 상황의 해석편향에 초점을 맞췄다는 한계를 보였다. 따라서 본 연구는 사회적 시나리오에서 얼굴표정(시각)과 음성(청각)의 다감각 정보를 동시에 제시하여, 사회불안의 다차원적 특성이 모호한 사회적 상황에서의 해석편향에 미치는 차별적인 영향을 검증하고자 한다. 즉, 사회불안을 구성하는 타인의 평가에 대한 두려움이라는 인지적 측면과 상호작용 자체에 대한 불안이라는 경험적 측면의 두 하위 요소가 시청각 정서 자극을 마주하였을 때 어떠한 해석편향과 관련되는지를 확인하고자 한다. 이 때 정서의 범주적 판단은 미묘한 정서 강도의 차이나 복합적인 정서 상태를 측정하지 못한다는 한계가 있다(Barrett, 2006). 이에 시각적 아날로그 척도(visual analogue scale [VAS])를 활용하여 모호한 사회적 맥락과 얼굴 및 음성 자극에 대해 평정하도록 하였다.

본 연구의 구체적인 가설은 다음과 같다. 첫째, 선행 연구들과 일관되게 전반적인 사회불안 수준이 높을수록, 언어적으로만 제시된 모호한 사회적 상황 시나리오 과제에서 정서가를 더 부정적으로 평가하는 해석편향을 보일 것이다. 둘째, 전반적인

사회불안 수준이 높을수록, 즉 사회적 상호작용에 대한 불안과 부정적 평가에 대한 두려움 모두가 높을수록 자극의 정서가가 모호한 조건인 중립 표정-중립 음성을 정서가가 분명한 조건인 긍정 표정-긍정 음성보다 더욱 부정적으로 평가할 것이다. 특히 이 과정에서 부정적 해석편향이 얼굴 표정보다는 상대적으로 더 모호한 자극인 음성 단서에서 더욱 강하게 나타날 것이다.셋째, 사회 불안의 해석편향에서 인지적 측면과 경험적 측면 간 차별적인 경향성이 나타날 것이다.

방법

연구 대상

본 연구는 사회불안이 청년층인 대학생에게 가장 흔하게 나타난다는 점을 반영하여(이시형 외, 1994; Jefferies & Ungar, 2020), 일반 대학(원)생들을 대상으로 연구를 진행하였다. G*Power 3.1 프로그램을 이용하여 본 연구에 적합한 대상자 수를 산정하였다(Faul et al., 2007). Cohen(2013)이 제시한 중간 수준의 효과 크기($r=.30$)를 기준으로, 양측 검증에서 유의수준 .05와 검정력 .80을 충족시키는 데 필요한 최소 표본 크기는 84명으로 산출되었다. 연구 탈락자들을 감안하여 본 연구에 국내 대학(원)에 재학 중인 학생 89명(남성 20명, 여성 69명)이 참여하였다. 한국어를 모국어로 사용하여 제시되는 시나리오와 지시문을 이해하는 데 문제가 없으며, 나안 혹은 교정시력이 정상이었으며, 연구 시작 전 예비 음성을 듣고 청력 상에 문제가 없는 사람들이 연구에 참여하였다.

연구 절차

본 연구는 가톨릭대학교 생명윤리심의위원회 (Institutional Review Board [IRB])의 승인을 받은 후 진행되었다(IRB 과제관리번호 1040395-202108-14). 연구 참가자는 연구자가 속한 기관의 홈페이지 내 수업 공지사항을 통해 모집하였다.

모든 참가자들은 본 연구의 목적과 절차에 대해 상세한 설명을 듣고 동의서를 작성했다. 실험 절차를 수행한 이후 인구통계학적 질문과 자기보고식 설문을 구글 폼(google form)을 통해 작성하였다. 참가자들에게는 실험이 종료된 후에 연구에 대해 추후 설명(debriefing)이 제공되었다. 또한 2,000원 상당의 보상을 제공하였으며, 심리학 관련 수업을 듣는 학생의 경우에는 해당 교과목에 대한 추가 점수를 지급하였다. 연구 참여에 소요된 시간은 약 25분이었다.

측정 도구

사회적 상호작용 불안 척도. 실제 사회적 상호작용 상황 자체에서 느끼는 불안과 불편감을 측정하기 위해 Mattick과 Clarke(1998)이 개발하고 김향숙(2001)이 우리말로 번안 및 타당화한 사회적 상호작용 불안 척도(Social Interaction Anxiety Scale [SIAS])를 활용하였다. 본 연구에서는 김향숙(2001)이 번안한 19문항에 최병희가 번안한 1문항을 덧붙인 20문항의 척도를 사용하였다. 이 척도는 “사회적 상황에서 무엇을 말해야 할지 몰라 걱정하곤 한다”, “여러 사람들과 어울릴 때 긴장한다”와 같이 사회적 상호작용을 요구하는 다양한 상황에서의 인지적, 정서적, 행동적

반응에 대한 20개의 문항에 대해 응답하는 자기 보고식 척도이다. 척도는 5점 리커트(Likert) 척도(0=전혀 그렇지 않다~4=매우 그렇다)로 응답하도록 되어 있다. 김향숙(2001)의 연구와 본 연구의 내적 일치도 계수(Cronbach's α)는 모두 .92였다. 본 연구에서는 SIAS를 참가자의 사회적 상호작용에 대한 불안이라는 사회불안의 경험적 측면을 측정하는 주요 변인으로 사용하였다.

단축형 부정적 평가에 대한 두려움 척도.
사회불안의 인지적 특성인 타인에게 부정적인 평가를 받는 것에 대한 두려움을 측정하기 위해 Leary(1993)가 제작한 단축형 부정적 평가에 대한 두려움 척도(Brief Fear of Negative Evaluation [BFNE])를 사용하였다. 이 척도는 Watson과 Friend(1969)가 30문항으로 제작한 원척도에서 FNE 전체 점수와 .50 이상의 상관이 있는 문항들만 선별하여 단축형의 12문항으로 구성한 것이다. 이정윤과 최정훈(1997)이 우리말로 번안 및 타당화하였다. 한국판 BFNE는 각 문항에 대해 5점 리커트 척도(1점=전혀 그렇지 않다~5점=매우 그렇다)로 응답하도록 구성되었다. 구체적인 문항의

예로는 “사람들이 나의 결점을 알아차릴까봐 자주 걱정된다”, “누군가와 얘기할 때 그가 나에 대해 어떤 생각을 하고 있을지 염려된다” 등이다. 이정윤과 최정훈(1997)의 연구에서 보고한 내적 일치도 계수(Cronbach's α)가 .90이었으며, 본 연구에서의 내적 일치도 계수는 .93이었다. 본 연구에서는 BFNE를 타인의 부정적 평가에 대한 두려움이라는 사회불안의 인지적 측면을 측정하는 주요 변인으로 사용하였다.

실험 방법

실험 장치

PsychoPy v3.0을 사용하여 실험을 제작하였다 (Peirce, 2007). 실험은 개인별로 자연광과 외부 소음이 최대한 차단된 대학 내 심리학 실험실에서 진행되었다.

실험 자극

사회적 상황 시나리오는 사회불안이 해석편향에 미치는 영향을 알기 위해서 개발되었다. 사회공포증 척도(Social Phobia Scale, 김향숙, 2001;

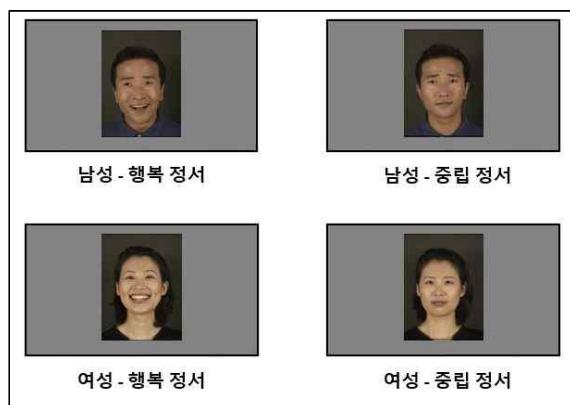


그림 1. 실험에 사용된 얼굴 표정의 예

Mattick & Clarke, 1998)와 SIAS를 참고하여 사회적 상황 17개를 개발하였다. 시나리오는 “당신이 모임에 가기 위해서 새로운 옷을 샀습니다. 새로 산 옷을 입고 갔을 때 친구들이 말했습니다.”, “당신이 학교 전공 수업에서 발표를 했습니다. 발표를 마친 후 교수님이 말씀하셨습니다.”와 같이 대학생들이 일상에서 흔히 접하는 사회적 상황이었다. 그리고, 각 상황에서 보이는 상대방의 반응은 단어 그 자체로는 긍정적이지만, 발화자의 어조에 따라서 긍정적으로도 혹은 부정적으로도 인식될 수 있는 단어로 구성하였다. 예를 들어, 앞서 언급한 전자의 시나리오에 대한 반응은 “괜찮다”였고, 후자의 경우에는 “잘했다”였다.

상기 자극은 사전에 15명의 심리학 전공 대학원생들에 의해 타당화되었다. 이들에게 총 42개의 사회적 상황 시나리오를 제시하였고, 그 상황이 얼마나 부정적(-4점) 혹은 긍정적인지(+4점)를 시각적 아날로그 척도(VAS)로 평정하도록 하였다. 이들이 평정한 값의 평균이 절대값 0에 가까워서 가장 중립적이라고 평가된 17개의 사회적 상황 시나리오를 선정하였다. 이를 17개 사회적 상황 시나리오 정서가의 평균값은 0.29($SD=0.59$)였다.

얼굴표정 자극은 한국인 얼굴 데이터베이스인 YFace Database(Chong et al., 2019)에서 선택한 34명(여성 17명, 남성 17명)의 얼굴 사진을 사용하였다. 연구의 가설을 검증하기 위해서 행복(happy)과 중립(neutral) 정서를 표현하는 사진 총 68장(17명 모델 × 2개 정서 × 2개 성별)을 선택하였다. 자극을 제시하고 반응하는 과정에서 감정을 인식하는 사람(참여자)의 성별과 감정을 표현하는 사람(자극)의 성별, 그리고 이 둘의 상호작용이 정서 인식의 정확도나 반응 시간에 체계적으로

영향을 미치는 성별 효과가 발생한다(Kret & De Gelder, 2012). 이에 본 연구에서는 이를 통제하고 사회불안 변인의 효과를 더 명확히 확인하기 위해, 참여자의 생물학적 성별과 자극(얼굴 사진 및 음성)의 성별을 일치시키는 방법을 사용하였다. 즉, 여성 참가자에게는 여성의 얼굴과 음성을, 남성 참가자에게는 남성의 얼굴과 음성을 제시하였다. 실험에 사용된 얼굴표정 사진의 예는 초상권의 문제로 이와 유사하게 제작된 사진을 그림 1에 제시하였다.

음성 자극은 아마추어 배우와 아나운서가 녹음한 것으로, 총 4쌍(여성 2명, 남성 2명)으로 이루어져 있다. 음성은 사회적 상황에서 상대가 보이는 반응으로, 동일한 문구일지라도 톤을 달리 하여 행복(happy)과 중립(neutral) 정서를 띠도록 하였다. 얼굴표정 사진과 마찬가지로 성별 효과를 없애기 위해 참가자와 같은 성별의 음성을 들려주었다.

음성 자극 역시 사전에 16명의 심리학 전공 대학원생들에 의해 타당화되었다. 17개의 사회적 상황에서 녹음된 인물의 반응(예: “괜찮네”, “잘했어”) 음성 102개를 제시하고, 그 음성이 얼마나 부정적(-4점) 혹은 긍정적인지(+4점)를 VAS로 평정하였다. 긍정적 음성은 각 인물의 반응별로 평균값이 가장 큰 자극을 최종 자극으로 선택하였다. 그리고, 중립 정서를 표현하는 음성의 경우, 각 인물의 반응별로 평균이 0점(중립)과 가장 가까운 자극을 최종적으로 선택하였다. 긍정 음성의 평균값은 3.26($SD=0.42$)이었으며, 중립 음성의 평균값은 0.18($SD=0.58$)이었다.

실험 절차

실험을 시작할 때 모든 연구 참여자들은 본 연구의 목적과 절차에 대해 상세히 설명을 받고, 동의서를 작성했다. 이후 예비 음성을 듣고 40~45db 범위에서 소리 크기를 조절하였다. 모니터에서 약 50cm 정도의 거리를 두었고 헤드셋을 쓴 상태에서 컴퓨터 자판과 마우스를 이용해 실험 과제를 수행했다. 과제는 먼저 시나리오에 대한 평정을 하는 과제 1을 시행하였다. 이후 과제 2를 시행하기 전 연습 시행을 4회 실시하였고, 이후 과제 2에 대한 본 시행을 실시하였다.

첫 번째 과제에서는 참여자에게 사회적 상황 시나리오가 제시되었다. 문장으로 구성된 시나리오를 다 읽은 후 ‘space bar’를 누르면 시나리오에 대한 반응이 문자로 제시되었다. 참여자들은 제시된 반응이 얼마나 긍정적 혹은 부정적인지 마우스를 이용하여 점수를 평가하도록 요청받았다. 시

나리오의 순서는 무선적으로 제시되었으며, 평정은 -4점(부정적)부터 +4점(긍정적)까지의 VAS를 사용하였고, 0점을 기준으로 참가자가 선택한 지점까지의 거리를 기록했다. 이 과제는 총 17개 시행으로 구성되었다.

두 번째 과제에서는 문장으로 기술된 시나리오가 제시된 후 그 상황에서 상대의 반응이 문자로 제시된 첫 번째 과제와는 다르게 얼굴표정과 음성 자극이 동시에 제시되는 형태로 나타나도록 하였다. 즉, 시나리오를 다 읽은 후 참가자가 ‘space bar’를 누르면 이후 얼굴과 음성 자극쌍이 2000ms 동안 동시에 제시되었다. 자극이 사라진 이후 참여자들은 제시된 상대의 반응이 얼마나 긍정적 혹은 부정적인지 마우스를 이용해 점수를 평가하였다. 시나리오와 얼굴표정-음성 자극 세트는 무선적으로 제시되었고, 동일한 시나리오가 연속되어 제시되지 않도록 했다. 첫 번째 과제와 동일하게 평정은 -4점(부정적)부터 +4점(긍정적)까

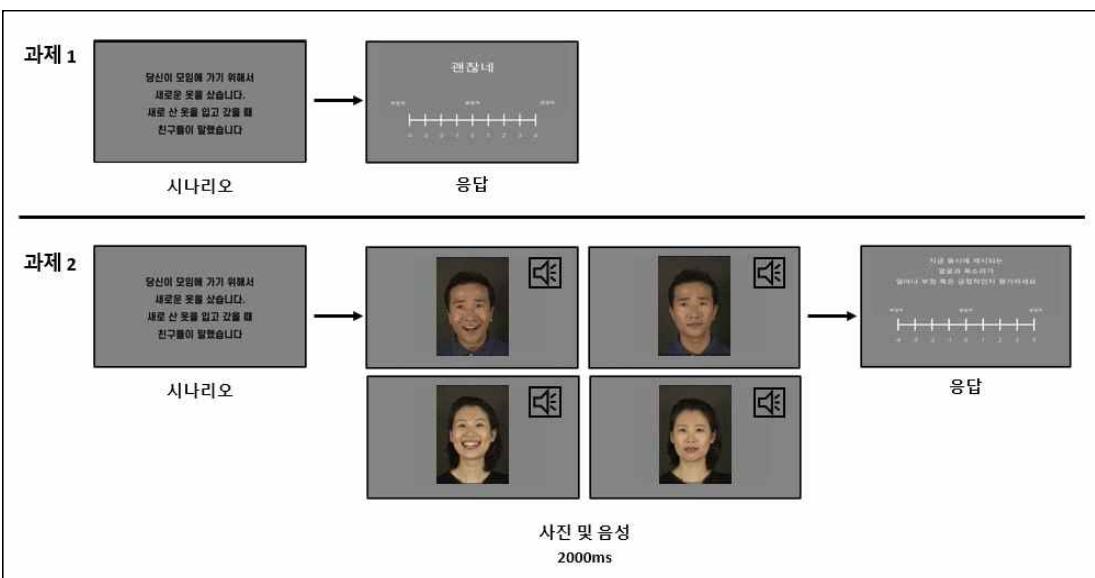


그림 2. 실험 시행의 예

지의 VAS를 사용하였다. 두 번째 과제는 총 68회의 시행(시나리오 17개 × 2개 정서 얼굴[행복, 중립] × 2개 정서 음성[행복, 중립])으로 이루어졌다.

두 과제의 시행은 합쳐서 총 85시행이었으며, 실험이 끝난 후에 참여자들은 사회불안의 경험적 측면과 인지적 측면을 측정하는 자기보고식 설문지인 SIAS와 BFNE를 실시하였다. 모든 절차가 끝난 뒤, 연구 참여자에게 사후 설명과 실험 참여에 대한 보상을 제공하였다. 실험은 총 25분 정도가 소요되었다. 실험 시행의 예를 절차의 도식은 그림 2에 제시하였다.

자료 분석

모든 통계 분석은 R 4.4.3(R Core Team, 2025)을 사용하여 수행하였다. 주요 패키지로는 선형 혼합 효과 모형(linear mixed-effects model [LMM]) 분석을 위해 lme4와 lmerTest를, 사후 분석 및 시각화를 위해 emmeans와 ggplot2를 사용하였다. 참가자의 평정값은 -4점(부정적)에서 +4점(긍정적) 범위였다. 이때 해석편향은 참가자가 평정한 VAS 점수의 방향성으로 간주하였다. 즉, 평정 점수가 0점(중립)에서 벗어나 -4점(부정적) 혹은 +4점(긍정적) 방향으로 치우치는 정도를 의미하며, 0점에 가까울수록 중립적인 해석을, -4점에 가까울수록 부정적 해석편향을, +4점에 가까울수록 긍정적 해석 경향을 나타내는 것으로 보았다. 또한 사회불안 척도인 SIAS와 BFNE의 점수는 평균 중심화 후 표준화(z 점수 변환)하였다.

과제 1에서는 사회불안 수준이 문자로 제시된 사회적 상황 해석에 미치는 영향을 확인하기 위

해 LMM을 적용하였다. 시나리오 반응(문자)에 대한 평정값을 종속변인으로 하고, SIAS와 BFNE를 각각 예측변인으로 투입하였다. 참가자와 항목(시나리오 × 참가자 성별)은 무선 절편(random intercept)으로 포함하였다.

과제 2에서는 시각(얼굴표정)과 청각(음성) 단서가 동시에 제시되는 사회적 상황에서 사회불안이 정서 인식에 미치는 영향을 확인하기 위해 공분산분석(analysis of covariance [ANCOVA]) 방식의 LMM을 적용하였다. 종속변인은 시나리오 반응(얼굴표정 및 음성)에 대한 평정값이며, 과제 1의 개인별 평균 평정값을 공변량으로 포함하였다. 이를 통해 문자 정보 해석의 개인차를 통제한 상태에서 시청각 정보의 순수한 효과를 알아보고자 하였다. 주요 예측변인으로는 얼굴표정(행복, 중립), 음성(행복, 중립), 그리고 사회불안(SIAS 또는 BFNE)을 포함하였으며, 이들 간의 모든 상호작용 효과를 확인하기 위해 삼원 상호작용 항(얼굴표정 × 음성 × 사회불안)을 모형에 포함하였다. 또한, 과제 1과 동일하게 참가자와 항목을 무선 절편으로 설정하였다.

LMM 분석에 따라 유의하거나 경계선 수준의 상호작용이 나타난 경우, emmeans 패키지를 사용하여 조건별 추정평균(estimated marginal means [EMMs])을 산출하고 Tukey 방법으로 다중비교를 보정하였다. 또한, 사회불안 수준에 따른 조건별 단순기울기(simple slope)를 산출하고 기울기 차이를 검증하였다.

사회불안 수준과의 상호작용 효과는 ggplot2를 이용해 시각화하였다. 그래프는 독자의 직관적인 이해를 돋기 위해 x축을 표준화 점수(z 점수)가 아닌 척도의 원점수로 역변환하여 제시하였으

며, 예측된 추세선과 95% 신뢰구간(confidence intervals [CI])을 함께 표시하였다.

결과

연구 참가자 특성

참가자 89명의 연령 평균은 22.04($SD=3.29$)세였다.

으며, 남성은 20명(22.5%), 여성 69명(77.5%)이었다. 사회적 상호작용에 대한 불안이라는 사회불안의 경험적 측면을 측정하는 SIAS의 평균은 34.52($SD=14.04$)였으며, 부정적 평가에 대한 두려움이라는 사회불안의 인지적 측면을 측정하는 BFNE의 평균은 37.44($SD=10.67$)였다.

표 1. 시청각 통합 과제에서 사회적 상호작용 불안에 따른 선형 혼합 효과 모형

예측변인	<i>b</i>	SE	<i>df</i>	<i>t</i>	<i>p</i>	95% 신뢰구간
(절편)	2.15	0.14	121.97	15.61	<.001	[1.88, 2.42]
공변량(과제 1 평정값)	0.29	0.08	85.98	3.84	<.001	[0.14, 0.44]
얼굴(행복/중립)	-2.46	0.04	5924.53	-55.40	<.001	[-2.55, -2.38]
음성(행복/중립)	-1.02	0.04	5924.53	-22.89	<.001	[-1.11, -0.93]
SIAS	-0.11	0.06	145.29	-2.01	0.047	[-0.22, -0.00]
얼굴 × 음성	0.12	0.06	5924.53	1.88	0.060	[-0.01, 0.24]
얼굴 × SIAS	0.04	0.04	5924.53	0.81	0.416	[-0.05, 0.12]
음성 × SIAS	-0.01	0.04	5924.53	-0.22	0.830	[-0.09, 0.07]
얼굴 × 음성 × SIAS	0.02	0.06	5924.53	0.27	0.789	[-0.11, 0.14]

주. 공변량으로 과제 1(문자 정보 조건)의 개인별 평균 평정값을 포함하였다. 얼굴 및 음성 변인은 행복/중립으로 코딩하였으며, 중립 조건이 기준(reference)으로 설정되었다. SIAS=사회적 상호작용 불안 척도(Social Interaction Anxiety Scale)

표 2. 시청각 통합 과제에서 부정적 평가에 대한 두려움에 따른 선형 혼합 효과 모형

예측변인	<i>b</i>	SE	<i>df</i>	<i>t</i>	<i>p</i>	95% 신뢰구간
(절편)	2.13	0.14	121.3	15.11	<.001	[1.85, 2.41]
공변량(과제 1 평정값)	0.31	0.08	86.13	3.94	<.001	[0.15, 0.47]
얼굴(행복/중립)	-2.46	0.04	5924.53	-55.41	<.001	[-2.55, -2.38]
음성(행복/중립)	-1.02	0.04	5924.53	-22.90	<.001	[-1.11, -0.93]
BFNE	-0.01	0.06	141.55	-0.08	0.937	[-0.12, 0.11]
얼굴 × 음성	0.12	0.06	5924.53	1.88	0.060	[-0.01, 0.24]
얼굴 × BFNE	0.01	0.04	5924.53	0.22	0.824	[-0.08, 0.10]
음성 × BFNE	-0.09	0.04	5924.53	-1.91	0.056	[-0.17, 0.00]
얼굴 × 음성 × BFNE	0.04	0.06	5924.53	0.68	0.494	[-0.08, 0.17]

주. 공변량으로 과제 1(문자 정보 조건)의 개인별 평균 평정값을 포함하였다. 얼굴 및 음성 변인은 행복/중립으로 코딩하였으며, 중립 조건이 기준(reference)으로 설정되었다. BFNE=단축형 부정적 평가에 대한 두려움 척도(Brief Fear of Negative Evaluation)

과제 1(문자 정보 조건)

과제 1에서 시나리오에 대한 반응이 문자로만 제시되었을 때, SIAS와 BFNE 모두 더 부정적인 정서 인식과 관련되었다. LMM 분석 결과, BFNE 점수가 높을수록 시나리오에 대한 반응을 유의미하게 더 부정적으로 인식하였다($b=-0.17$, $t(86.89)=-2.56$, $p=.012$, 95% CI [-0.31, -0.04]). SIAS의 효과 역시 통계적으로 유의한 수준에 가깝게 부정적인 경향성을 보였다($b=-0.14$, $t(86.64)=-1.97$, $p=.052$, 95% CI [-0.27, 0.00]).

과제 2(시청각 정보 조건)

얼굴 및 음성 단서의 주효과 및 상호작용

과제 2에서는 시나리오에 대한 반응으로 시청각 정보가 제시되었으며, SIAS와 BFNE를 각각 예측 변인으로 하는 LMM 분석 결과를 표 1과 2에 제시하였다. SIAS와 BFNE 모형 모두에서 얼굴 및 음성 단서의 주효과가 유의하였으며 ($p<.001$), 상호작용은 유의한 경향성을 보였다 (SIAS 모형: $p=.061$, BFNE 모형: $p=.060$). 사후 분석 결과, 얼굴-음성의 네 가지 조건(행복 얼굴-행복 음성, 행복 얼굴-중립 음성, 중립 얼굴-행복 음성, 중립 얼굴-중립 음성)의 예측 평균 평정값은 모두 서로 유의미한 차이를 보였다(표 3). 구

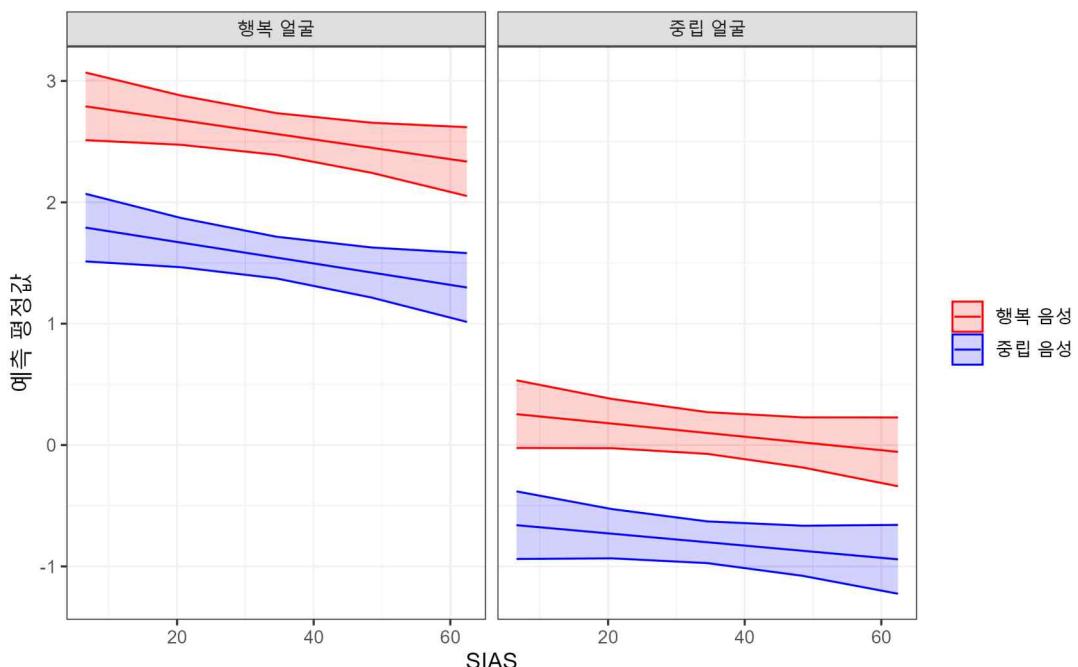


그림 3. 사회적 상호작용 불안 수준에 따른 얼굴-음성 단서 조건별 예측 평정값

주. Y축은 과제 2에서의 예측 평정값으로, 점수가 낮을수록 부정적인 평가를 나타낸다. X축은 SIAS의 원점수이며, 그림에는 95% 신뢰구간(음영 범위)이 함께 표시되었다. SIAS=사회적 상호작용 불안 척도(Social Interaction Anxiety Scale)

체적으로, 얼굴과 음성 모두 행복인 조건이 가장 긍정적으로 인식되었고, 다음으로 행복 얼굴-중립 음성, 중립 얼굴-행복 음성, 중립 얼굴-중립 음성 조건 순으로 낮아졌다.

사회적 상호작용 불안(SIAS)의 주효과

표 1(SIAS를 예측 변인으로 하는 LMM)에 제시된 바와 같이, SIAS는 유의미한 주효과를 보였다($b=-0.11$, $t(145.29)=-2.01$, $p=.047$, 95% CI [-0.22, -0.00]). 이는 전반적으로 SIAS 수준이 높을수록 사회적 상황에서 시청각 정보가 함께 제시될 때 자극에 대한 정서를 더 부정적으로 인식함을 의미한다. 이러한 주효과의 구체적인 양상을

파악하기 위한 사후 분석 결과(그림 3), SIAS의 부정적 효과는 행복한 얼굴이 제시된 두 조건, 즉 행복 얼굴-행복 음성($b=-0.11$, $p=.045$, 95% CI [-0.23, -0.00])과 행복 얼굴-중립 음성($b=-0.12$, $p=.030$, 95% CI [-0.23, -0.01]) 조건에서만 통계적으로 유의하게 나타났다. 중립 얼굴이 제시된 조건에서는 SIAS의 효과가 유의하지 않았다(중립 얼굴-행복 음성($b=-0.08$, $p=.171$, 95% CI [-0.19, 0.03])과 중립 얼굴-중립 음성($b=-0.07$, $p=.215$, 95% CI [-0.18, 0.04])). 한편, SIAS와 관련된 상호작용 효과는 유의하지 않았다(표 1 참고).

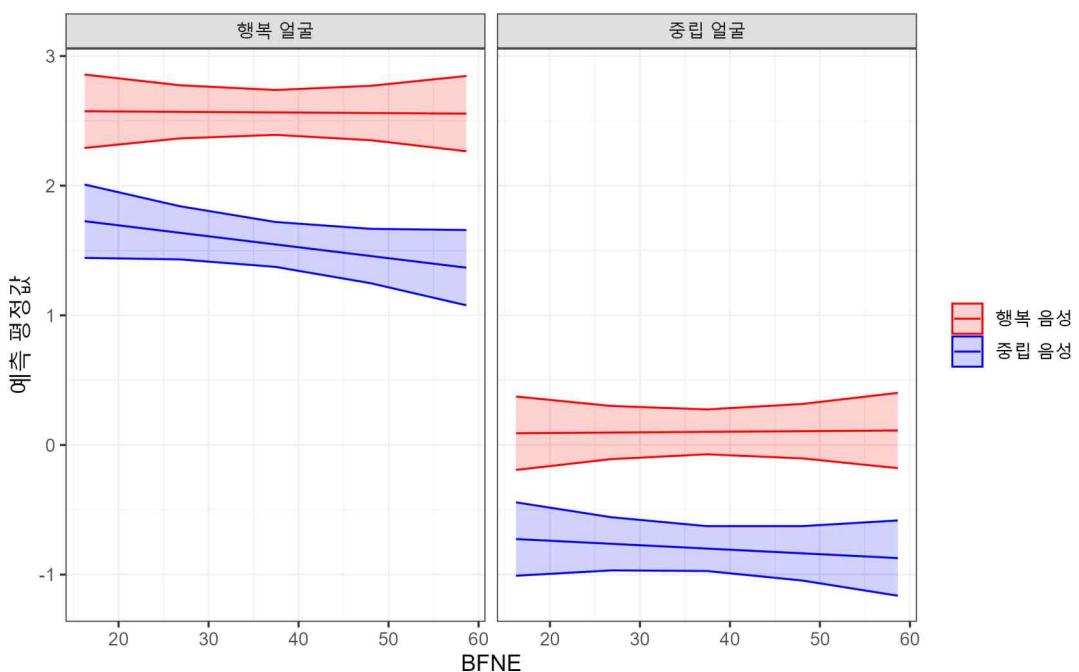


그림 4. 부정적 평가에 대한 두려움 수준에 따른 얼굴-음성 단서 조건별 예측 평점값

주. Y축은 과제 2에서의 예측 평점값으로, 점수가 낮을수록 부정적인 평가를 나타낸다. X축은 BFNE의 원점수이며, 그림에는 95% 신뢰구간(영어 범위)이 함께 표시되었다. BFNE=단축형 부정적 평가에 대한 두려움 척도(Brief Fear of Negative Evaluation)

부정적 평가에 대한 두려움(BFNE)과 음성 단서의 상호작용

표 2(BFNE를 예측 변인으로 하는 LMM)에 제시된 바와 같이 BFNE의 주효과는 유의하지 않았으나($p=.937$), 음성 단서와의 상호작용 효과는 통계적으로 유의한 수준에 가까웠다($b=-0.09$, $t(5924.5)=-1.91$, $p=.056$, 95% CI [-0.17, 0.00]). 이에 BFNE 수준에 따른 각 음성 조건에서의 단순 기울기를 산출하였으며(행복 음성: $b=0.00$, $p=.995$, 중립 음성: $b=-0.06$, $p=.242$), 두 기울기 간의 차이가 통계적으로 유의하였다($b=0.06$, $z=2.02$, $p=.043$, 95% CI [0.00, 0.12]). 즉, BFNE가 높을수록 행복 음성 조건에 비해 중립 음성 조건에서 정서 인식이 더 부정적이었다(그림 4).

논의

사회불안 증상이 사회적 상황을 부정적으로 해석하는 편향과 관련된다는 대한 연구 결과는 지속적으로 확인된 바 있다. 그러나 선행 연구들은 언어화된 시나리오나 얼굴표정과 같은 단일한 감각 자극만을 활용하였으며, 주로 중립적이거나 부정적 상황의 해석편향에 초점을 맞췄다는 한계를 보였다. 또한 사회불안 증상이 가지는 인지적, 경험적 측면의 차이에 대한 고려도 거의 이루어지지 않았다. 따라서 이번 연구에서는 시나리오를 이용하여 사회적 상황을 제시하고 그 상황에서 등장하는 인물의 얼굴표정이라는 시각 자극과 음성 자극을 동시 제시하여 사회불안이 시청각 자극의 정서 인식에 어떤 영향을 미치는지를 확인하고자 하였다. 이 과정에서 사회불안의 인지적 측면과 경험적 측면을 구분하여, 사회불안의 각

측면이 각기 다른 유형의 사회적 정보 처리에 차별적인 해석편향과 관련이 있는지 살펴보았다.

본 연구의 결과를 요약하면 다음과 같다. 첫째, 사회불안의 여부를 떠나 사람들은 동일하게 상황이 묘사되었더라도 비언어적 단서, 즉 얼굴표정과 음성에 따라 그 상황의 정서가를 인식하는 것이 달랐다. 구체적으로, 행복 얼굴-행복 음성, 행복 얼굴-중립 음성, 중립 얼굴-행복 음성, 중립 얼굴-중립 음성 쌍 순서대로 그 상황을 더 부정적으로 인식하였다. 얼굴표정과 음성 정보가 행복으로 일치할 때 가장 강한 긍정 정서로 인식하였고, 반면에 두 자극 모두 중립일 때 가장 부정적으로 인식한 것은 당연한 결과이다. 선행 연구에서도 시각과 청각 정보가 일치할 때 얼굴 정서 인식이 더욱 촉진된다고 보고된 바 있다(de Gelder & Bertelson, 2003). 흥미로운 결과는 행복 얼굴-중립 음성 쌍과 중립 얼굴-행복 음성의 쌍이 제시되었을 때 전자를 후자보다 더 긍정적으로 평가하였다는 것이다. 이는 각기 다른 정서의 자극이 제시된 경우에 얼굴표정이라는 시각 정보를 음성이라는 청각 정보보다 더 우선시하여 처리하였다는 것을 시사하며, 이는 선행 연구를 재확인하는 결과이다(Collignon et al., 2008; Ernst & Banks, 2002). 이에 더해 이번 연구에서는 행복과 중립 자극을 활용함으로써 시각 우세 효과가 긍정 정서에도 유사하게 나타나는 것을 보여주었다는 점에서 의미를 갖는다.

둘째, 본 연구에서 주되게 관심을 가진 것은 사회불안을 구성하는 타인의 부정적인 평가에 대한 두려움이라는 인지적 측면과 사회적 상호작용에 대한 불안이라는 경험적 측면 사이의 차별적인 영향이 나타난다는 점이다. 우선 언어적으로 제시

된 모호한 사회적 상황에 대한 평정을 해야 하는 경우(과제 1)에서 사회불안의 인지적 측면, 즉, BFNE의 척도 점수가 높을수록 시나리오에 대한 정서가를 부정적으로 평가하는 부정적 해석편향을 보였다. 반면, 사회적 상황 및 시청각 자극으로 제시된 과제 2에서는 사회적 상호작용에 대한 불안 즉, SIAS 점수가 높을수록 부정적으로 평가하는 경향성을 보였다. 이러한 차이는 두 척도가 측정하는 사회불안의 특징과 각 실험 과제가 요구하는 심리적 과정이 서로 다른 것에서 비롯된 것으로 해석할 수 있다. 본 연구의 과제 1은 참가자가 언어적 자극으로 제시되는 사회적 상황과 응답을 읽고, 이에 대한 의미를 추론하는 인지적 처리를 요구하는 과제이다. 따라서 부정적 평가에 대한 두려움을 인지적 수준에서 측정하는 BFNE 점수는 상기한 과제에서 관련성을 크게 보였을 가능성이 있다. 반면, 본 연구의 과제 2는 얼굴 표정과 음성이라는 사회적인 감각 단서를 실시간으로 마주하는 실제 상호작용과 유사한 경험을 제공한다. 이에 사회적 상호작용에서 나타나는 불안과 긴장이 반영되었을 가능성이 크고, 이러한 측면이 사회불안의 경험적 측면을 측정하는 SIAS와 연관될 수 있다. 결론적으로, 본 연구의 결과는 해석편향이 단일한 과정이 아님을 시사한다.

셋째, 상기한 바에 이어서 사회불안의 인지적 측면과 경험적 측면이 어떠한 해석편향과 연관되어 있는지에 대해서도 보다 구체적으로 살펴보았다. 우선 사회불안을 구성하고 있는 부정적 평가에 대한 두려움이라는 인지적 측면을 살펴보면, 중립 정서의 음성이 제시되는 조건에서 부정적으로 해석하는 경향성이 유의하다. 음성은 다차원적 단서로 구성되어 있어 본질적으로 모호성이 큰데,

특히 중립적인 정서가를 띤 음성 자극은 더더욱 모호성이 크다(Belin et al., 2004; Juslin & Laukka, 2003). 모호성을 처리하는 과정에서 사회불안의 인지적 측면이 두드러지는 집단은 이러한 불확실성에 대해 자신의 부정적인 스키마를 기반으로 하여 해석할 가능성이 크다. 즉, 불확실성에 대한 의미를 만들어내는 과정에서 모호성을 위협적인 단서로 왜곡하게 되며, 이로부터 해석편향이 비롯될 수 있음을 시사한다(Clark & Wells, 1995). 한편, 사회불안의 경험적 측면인 사회적 상호작용에 대한 불안이 높은 사람들은 행복한 얼굴 표정이 나타나는 조건에서 덜 긍정적으로 평가하는 경향을 보였다. 상호작용 상황에서 경험하는 불안과 불편감이 외부에서 주어지는 긍정적 단서의 힘을 약화시키거나 불신하게 만드는 요인으로 작용하는 것으로 보인다. 이에 사회적 상호작용에 대한 불안이 높은 사람들이 보이는 해석편향은 긍정성을 수용하지 못하는 문제 즉, 긍정적 해석편향의 부족과 연관된 것으로 보인다.

본 연구의 결과는 사회불안의 해석편향이 단일한 과정이 아님을 보여준다. 선행 연구들은 종종 사회불안을 단일 개념으로 측정하거나 부정적 해석편향에만 초점을 맞추었으나(Gutiérrez-García & Calvo, 2017; Maoz et al., 2016), 본 연구는 사회불안의 인지적 측면(BFNE)과 경험적 측면(SIAS)이 각기 다른 유형의 편향을 유발함을 시사하였다. 평가에 대한 두려움이 높은 개인들은 해석이 필요한 언어적 시나리오와 모호성이 높은 중립적 음성에 대해 선택적인 부정적 편향을 보였다. 이는 이들의 문제가 의미를 생성하는 과정, 즉 불확실한 정보를 내재된 부정적 스키마로 채워 넣는 인지적 오류에서 비롯됨을 시사한다. 반

면, 상호작용에 대한 불안이 높은 개인들은 모호한 자극이 아닌, 명백히 긍정적인 행복한 얼굴에 대한 평가를 절하하는 경향을 보였다. 이는 이들의 문제가 정서를 수용하는 과정, 즉 실제 상호작용 상황에서 경험하는 압도적인 불안감이 상대의 긍정적 신호를 인식하지 못하게 만드는 테에 있음을 보여준다. 이처럼 사회불안의 두 하위 요인이 부정적인 해석편향과 긍정 편향의 부족으로 분리된다는 발견은 사회불안의 인지 모델을 더욱 정교하게 확장한다는 이론적 중요성을 가질 수 있다.

더불어 본 연구는 기존의 언어 및 단일 자극 중심 연구에서 나아가, 얼굴과 음성이 결합된 다각 자극을 사용하여 생태학적 타당도를 높였다. 이번 연구를 통해 시각 우세 효과를 재확인하는 동시에, 청각 단서의 모호성이 평가에 대한 두려움을 가진 개인들의 부정적 해석편향을 유발하는 주요 요인이 될 수 있음을 밝혔다. 이들은 실제 사회적 상호작용에서 상대방의 말투나 억양을 부정적으로 인식하여 불안이 증가하는 악순환을 보일 가능성이 크다. 즉, 이번 결과는 사회불안이 높은 사람들이 의사소통 과정에서 단순히 내용에 대한 왜곡 뿐만 아니라 음성 단서 처리 과정에서 발생하는 체계적 왜곡에서 편향이 비롯될 수 있음을 시사한다.

이는 치료적 개입의 구체적인 목표를 설정하는데 중요한 단서를 제공한다. 기존의 인지 편향 수정(cognitive bias modification) 훈련이 주로 시각 자극(얼굴표정, 문장)에 초점을 맞춰왔다면, 본 연구 결과는 음성 자극에 초점을 맞춘 인지 편향 수정 훈련의 개발 필요성을 시사한다(Hertel & Mathews, 2011). 최근 한 연구에서는 사회불안

집단에게 얼굴표정을 제시하고 객관적인 정서 정보를 피드백으로써 단회기만으로도 정서 인식의 정확도를 높일 수 있음을 보여주었다(Choi et al., 2025). 이와 유사하게, 모호한 톤의 음성을 들려주고 그 의도를 판단하게 한 뒤 객관적인 피드백을 제공하거나, 대안적인 비위협적 해석을 생성하도록 유도하는 훈련을 할 수 있다. 이러한 개입은 일상적인 대화에서 상대방의 말투에 대해 자동적인 부정적 해석편향을 수정하고, 사회불안 증상을 완화하는 데 실질적인 도움을 줄 수 있을 것이다.

특히 본 연구 결과는 사회불안에 대한 심리치료적 개입이 어떤 종류의 불안이 핵심인지에 따라 맞춤화되어야 함 역시 시사한다. 평가에 대한 두려움이 높은 개인에게는 모호한 음성에 대해 위협적이지 않은 대안 해석을 생성하고 겸증하도록 돋는 인지 편향 수정 훈련, 특히 청각 자극을 활용한 훈련이 효과적일 수 있다(Heeren et al., 2015). 반면, 상호작용에 대한 불안이 높은 개인의 경우, 단순히 부정적 해석편향을 줄이는 것을 넘어 실제 상호작용과 유사한 상황에서 상대방의 긍정적 신호에 주의를 기울여 의미를 수용하는 연습이 더 중요할 수 있다. 이러한 결과는 긍정 단서에 대한 주의 편향 수정이 보다 효과적일 수 있음을 시사한다(Hertel & Mathews, 2011).

결론적으로, 본 연구는 사회불안의 다차원적 특성을 다감각적 정보 처리 과정 속에서 살펴봄으로써 이들의 복잡한 내면세계를 더 깊이 이해하고 효과적인 치료 전략을 개발하는 데 중요한 이론적, 실증적 근거를 제공하였다라는 의의가 있다. 그러나 다음과 같은 한계점을 가지며 이를 보완하기 위한 후속 연구를 제안한다. 첫째, 연구 표본

이 대학생 집단에 국한되어 있어, 이 결과를 임상적으로 사회불안장애 진단을 받은 집단으로 일반화하기에는 한계가 있다. 따라서 후속 연구에서는 실제 환자 집단을 대상으로 반복 검증하여 결과의 임상적 타당도를 높일 필요가 있다. 또한, 다양한 연령대를 대상으로 연구를 확장하여 해석편향의 발달적 측면을 확인하는 것 역시 필요하다. 둘째, 참여자의 성비가 주로 여성에 편중되어 있다는 점이다. 사회불안장애가 남성보다 여성에게서 더 높은 유병률을 보이기 때문에(APA, 2013), 표집에서 여성이 더 많은 점이 실제 환경에서의 성별 비율을 반영하는 결과일 수 있다. 그러나 보다 일반화 가능성을 높이고 다감각에 대한 해석편향에 성차가 있는지를 확인하기 위해서는 추후 연구에서 보다 균형 잡힌 성비의 표본을 모집하는 것이 필요하다.셋째, 자극의 정서 범위가 행복과 중립에 한정되어 있다. 사회적 평가와 밀접하게 관계된 분노, 슬픔, 혐오, 경멸과 같은 부정 정서 자극을 포함한 후속 연구가 필요하다. 넷째, 사회적 맥락 시나리오가 주로 대학생의 일상 경험을 반영하고 있다. 직장 생활 등 다른 연령대의 사람들이 흔히 경험하는 다양한 사회적 맥락으로 시나리오를 확장하여 연구 결과의 생태학적 타당도를 높일 필요가 있다. 다섯째, 본 연구는 실험 과제를 모두 수행한 후에 사회불안 척도를 측정하였다. 실험 과제, 특히 사회적 평가를 연상시키는 자극에 노출되는 경험이 참가자에게 일시적인 불안, 긴장 상태를 유발하였을 수 있다. 이러한 점이 개인의 설문 응답에 영향을 미쳤을 가능성을 완전히 배제하기 어렵다. 따라서 후속 연구에서는 실험 과제로 인한 잠재적 영향을 통제하기 위해, 실험 과제 시작 전에 자기보고식 척도를 측정하

거나, 충분한 시간 간격을 두고 측정하는 방식을 사용할 필요가 있다.

이러한 한계에도 불구하고 본 연구는 사회불안을 단일 차원이 아닌 인지적 측면과 경험적 측면으로 구분하여, 각 하위 요인이 다감각적 정보 처리 과정에서 서로 다른 해석편향을 유발한다는 점을 실증적으로 확인하였다는 점에서 중요한 의의를 갖는다. 평가에 대한 두려움이 모호한 청각 자극을 부정적으로 왜곡하는 해석편향과 관련된다면, 상호작용에 대한 불안은 명백한 긍정적 시각 단서를 온전히 수용하지 못하는 긍정 편향의 부족과 관련됨을 밝혔다. 이러한 발견은 사회불안의 인지 모델을 다감각적 차원에서 확장하고, 개인의 특성에 맞춘 차별화된 임상적 개입의 필요성을 제기하였다는 의의를 갖는다.

참고문헌

- 김향숙 (2001). 사회공포증 하위 유형의 기억편향. 석사 학위논문. 서울대학교, 서울. <https://doi.org/10.23170/snu.000000064217.11032.0000820>
- 백용매, 이대현 (2013). 사회불안 집단의 얼굴표정 정서자극에 대한 해석편향. 한국심리학회지: 임상, 32(1), 243-257. <https://doi.org/10.15842/kjcp.2013.32.1.013>
- 이시형, 신영철, 오강섭 (1994). 사회공포증에 관한 10년간의 임상연구. 신경정신의학, 33(2), 305-312. <https://scholar.kyobobook.co.kr/article/detail/4050026314609>
- 이정윤, 최정훈 (1997). 한국판 사회공포증척도(K-SAD, K-FNE)의 신뢰도 및 타당도 연구. 한국심리학회지: 임상, 16(2), 251-264. <https://accesson.kr/kjcp/assets/pdf/26569/journal-16-2-251.pdf>
- 이하영, 남종호, 양재원 (2020). 사회불안 증상과 얼굴

- 군집의 평균 정서 추론. *재활심리연구*, 27(4), 101–113. <https://doi.org/10.35734/karp.2020.27.4.006>
- 임효정, 양재원 (2021). 사회불안이 웃음, 비웃음 자각에 미치는 영향: 우울 증상의 통제 효과. *한국심리학회지: 건강*, 26(3), 491–507. <https://doi.org/10.17315/kjhp.2021.26.3.004>
- 전겸구, 최상진, 양병창 (2001). 통합적 한국판 CES-D 개발. *한국심리학회지: 건강*, 6(1), 59–76. <https://www.accesson.kr/healthpsy/v.6/1/59/27362>
- 최성영, 양재원 (2019). 맥락이 사회불안의 사회적 유능성 인식에 미치는 영향. *한국심리학회지: 임상*, 38(4), 369–381. <https://doi.org/10.15842/kjcp.2019.38.4.003>
- Abbott, M. J., & Rapee, R. M. (2004). Post-event rumination and negative self-appraisal in social phobia. *Behaviour Research and Therapy*, 42(8), 985–1002. <https://doi.org/10.1016/j.brat.2003.07.005>
- Alden, L. E., Taylor, C. T., Mellings, T. M., & Laposa, J. M. (2008). Social anxiety and the interpretation of positive social events. *Journal of Anxiety Disorders*, 22(4), 577–590. <https://doi.org/10.1016/j.janxdis.2007.05.007>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual for mental disorder* (5th ed.). American Psychiatric Publishing. <https://doi.org/10.1176/appi.books.9780890425596>
- Amin, N., Foa, E. B., & Coles, M. E. (1998). Negative interpretation bias in social phobia. *Behaviour Research and Therapy*, 36(10), 945–957. [https://doi.org/10.1016/S0005-7967\(98\)00060-6](https://doi.org/10.1016/S0005-7967(98)00060-6)
- Aviezer, H., Hassin, R. R., Bentin, S., & Trope, Y. (2008). Putting facial expressions back in context. In N. Ambady & J. J. Skowronsky (Eds.), *First Impressions* (pp. 255–286). Guilford Press. <https://psycnet.apa.org/record/2008-09248-011>
- Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*, 338(6111), 1225–1229. <https://doi.org/10.1126/science.1224313>
- Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, 10(1), 20–46. https://doi.org/10.1207/s15327957pspr1001_2
- Barrett, L. F., & Kensinger, E. A. (2010). Context is routinely encoded during emotion perception. *Psychological Science*, 21(4), 595–599. <https://doi.org/10.1177/0956797610363547>
- Bean, C. A., Everaert, J., & Ciesla, J. A. (2023). Positive interpretation bias predicts longitudinal decreases in social anxiety. *Behavior Therapy*, 54(2), 290–302. <https://doi.org/10.1016/j.beth.2022.09.003>
- Belin, P., Fecteau, S., & Bedard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Science*, 8(3), 129–135. <https://doi.org/10.1016/j.tics.2004.01.008>
- Benito-Ruiz, G., Luzón-Collado, C., Arrillaga-González, J., & Lahera, G. (2022). Development of an ecologically valid assessment for social cognition based on real interaction: Preliminary results. *Behavioral Sciences*, 12(2), Article 54. <https://doi.org/10.3390/bs12020054>
- Bronfman, Z. Z., Brezis, N., Lazarov, A., Usher, M., & Bar Haim, Y. (2018). Extraction of mean emotional tone from face arrays in social anxiety disorder. *Depression and Anxiety*, 35(3), 248–255. <https://doi.org/10.1002/da.22713>
- Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, 11(12), 535–543. <https://doi.org/10.1016/j.tics.2007.10.001>
- Chen, Y. P., Ehlers, A., Clark, D. M., & Mansell, W. (2002). Patients with generalized social phobia direct their attention away from faces. *Behaviour Research and Therapy*, 40(6), 677–687. [https://doi.org/10.1016/S0006-296X\(01\)00183-7](https://doi.org/10.1016/S0006-296X(01)00183-7)

- doi.org/10.1016/S0005-7967(01)00086-9
- Chen, J., Short, M., & Kemps, E. (2020). Interpretation bias in social anxiety: A systematic review and meta-analysis. *Journal of Affective Disorders*, 276, 1119–1130. <https://doi.org/10.1016/j.jad.2020.07.121>
- Choi, J., Kim, G., & Yang, J.-W. (2025). A single session feedback training modifies interpretation bias in individuals with high social anxiety: A randomized controlled trial. *British Journal of Clinical Psychology*, 64(2), 403–414. <https://doi.org/10.1111/bjcp.12512>
- Chung, K. M., Kim, S. J., Jung, W. H., & Kim, V. Y. (2019). Development and validation of the Yonsei Face Database (Yface DB). *Frontiers in Psychology*, 10, Article 2626. <https://doi.org/10.3389/fpsyg.2019.02626>
- Clark, D. M., & McManus, F. (2002). Information processing in social phobia. *Biological Psychiatry*, 51(1), 92–100. [https://doi.org/10.1016/S0006-3223\(01\)01296-3](https://doi.org/10.1016/S0006-3223(01)01296-3)
- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heinberg, M. Liebowitz, D. Hope, & F. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 69–93). New York: Guilford Press. <https://psycnet.apa.org/record/1995-98887-000>
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge. <https://doi.org/10.4324/9780203771587>
- Cole, J. (2015). Prosody in context: A review. *Language, Cognition and Neuroscience*, 30(1–2), 1–31. <https://doi.org/10.1080/23273798.2014.963130>
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research*, 1242, 126–135. <https://doi.org/10.1016/j.brainres.2008.04.023>
- Constans, J. I., Penn, D. L., Ihen, G. H., & Hope, D. A. (1999). Interpretive biases for ambiguous stimuli in social anxiety. *Behaviour Research and Therapy*, 37(7), 643–651. [https://doi.org/10.1016/S0005-7967\(98\)00180-6](https://doi.org/10.1016/S0005-7967(98)00180-6)
- Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, 16(2), 409–412. <https://doi.org/10.3758/BF03203962>
- De Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, 7(10), 460–467. <https://doi.org/10.1016/j.tics.2003.08.014>
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429–433. <https://doi.org/10.1038/415429a>
- Falagiarda, F., & Collignon, O. (2019). Time-resolved discrimination of audio-visual emotion expressions. *Cortex*, 119, 184–194. <https://doi.org/10.1016/j.cortex.2019.04.017>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. <https://doi.org/10.3758/BF03193146>
- Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, 11(5), 473–490. <https://doi.org/10.1162/08989299563544>
- Gomez, R., & Watson, S. D. (2017). Confirmatory factor analysis of the combined social phobia scale and social interaction anxiety scale: Support for a bifactor model. *Frontiers in Psychology*, 8, Article 70. <https://doi.org/10.3389/fpsyg.2017.00070>
- Gutiérrez-García, A., & Calvo, M. G. (2017). Social anxiety and threat-related interpretation of dynamic facial expressions: Sensitivity and

- response bias. *Personality and Individual Differences*, 107, 10–16. <https://doi.org/10.1016/j.paid.2016.11.025>
- Heeren, A., Mogoase, C., Philippot, P., & McNally, R. J. (2015). Attention bias modification for social anxiety: A systematic review and meta-analysis. *Clinical Psychology Review*, 40, 76–90. <https://doi.org/10.1016/j.cpr.2015.06.001>
- Heimberg, R. G., Brozovich, F. A., & Rapee, R. M. (2014). A cognitive-behavioral model of social anxiety disorder. In S. G. Hofmann & P. M. DiBartolo (Eds.), *Social anxiety: Clinical, developmental, and social perspectives* (pp. 705–728). Academic Press. <https://doi.org/10.1016/B978-0-12-394427-6.00024-8>
- Heinrichs, N., & Hofmann, S. G. (2001). Information processing in social phobia: A critical review. *Clinical Psychology Review*, 21(5), 751–770. [https://doi.org/10.1016/S0272-7358\(00\)00067-2](https://doi.org/10.1016/S0272-7358(00)00067-2)
- Hertel, P. T., & Mathews, A. (2011). Cognitive bias modification: Past perspectives, current findings, and future applications. *Perspectives on Psychological Science*, 6(6), 521–536. <https://doi.org/10.1177/1745691611421205>
- Huppert, J. D., Pasupuleti, R. V., Foa, E. B., & Mathews, A. (2007). Interpretation biases in social anxiety: Response generation, response selection, and self-appraisals. *Behaviour Research and Therapy*, 45(7), 1505–1515. <https://doi.org/10.1016/j.brat.2007.01.006>
- Jefferies, P., & Ungar, M. (2020). Social anxiety in young people: A prevalence study in seven countries. *PLoS One*, 15(9), e0239133. <https://doi.org/10.1371/journal.pone.0239133>
- Juslin, P. N., & Laukka, P. (2003). Communication of emotion in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129(5), 770–814. <https://doi.org/10.1037/0033-295X.129.5.770>
- Kret, M. E., & De Gelder, B. (2012). A review on sex differences in processing emotional signals. *Neuropsychologia*, 50(7), 1211–1221. <https://doi.org/10.1016/j.neuropsychologia.2011.12.022>
- Leary, M. R. (1983). A brief version of the fear of negative evaluation scale. *Personality and Social Psychology Bulletin*, 9, 371–375. <https://doi.org/10.1177/0146167283093007>
- Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, 36(4), 455–470. [https://doi.org/10.1016/S0005-7967\(97\)10031-6](https://doi.org/10.1016/S0005-7967(97)10031-6)
- Maoz, K., Eldar, S., Stoddard, J., Pine, D. S., Leibenluft, E., & Bar-Haim, Y. (2016). Angry-happy interpretations of ambiguous faces in social anxiety disorder. *Psychiatry Research*, 241, 122–127. <https://doi.org/10.1016/j.psychres.2016.04.100>
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Peschard, V., Maurage, P., & Philippot, P. (2014). Towards a cross-modal perspective of emotional perception in social anxiety: review and future directions. *Frontiers in Human Neuroscience*, 8, Article 322. <https://doi.org/10.3389/fnhum.2014.00322>
- Peschard, V., & Philippot, P. (2017). Overestimation of threat from neutral faces and voices in social anxiety. *Journal of Behavioral Therapy and Experimental Psychiatry*, 57, 206–211. <https://doi.org/10.1016/j.jbtep.2017.06.003>
- Pittam, J., & Scherer, K. R. (1993). Vocal expression and communication of emotion. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp.

- 185 - 197). The Guilford Press. <https://psycnet.apa.org/record/1993-98937-007>
- Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy*, 35(8), 741-756. [https://doi.org/10.1016/S0005-7967\(97\)00022-3](https://doi.org/10.1016/S0005-7967(97)00022-3)
- Romano, M., Moscovitch, D. A., Saini, P., & Huppert, J. D. (2020). The effects of positive interpretation bias on cognitive reappraisal and social performance: Implications for social anxiety disorder. *Behaviour Research and Therapy*, 131, Article 103651. <https://doi.org/10.1016/j.brat.2020.103651>
- Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, 9(4), 255 - 266. <https://doi.org/10.1038/nrn2331>
- Stopa, L., & Clark, D. M. (2000). Social phobia and interpretation of social events. *Behaviour Research and Therapy*, 38(3), 273-283. [https://doi.org/10.1016/s0005-7967\(99\)00043-1](https://doi.org/10.1016/s0005-7967(99)00043-1)
- Taylor, C. T., Bomyea, J., & Amir, N. (2010). Attentional bias away from positive social information mediates the link between social anxiety and anxiety vulnerability to a social stressor. *Journal of Anxiety Disorders*, 24(4), 403 - 408. <https://doi.org/10.1016/j.janxdis.2010.02.004>
- Taylor, C. T., Bomyea, J., & Amir, N. (2011). Malleability of attentional bias for positive emotional information and anxiety vulnerability. *Emotion*, 11(1), 127 - 138. <https://doi.org/10.1037/a0021301>
- Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, 33(4), 448-457. <https://doi.org/10.1037/h0027806>
- Weeks, J. W., Jakatdar, T. A., & Heimberg, R. G. (2010). Comparing and contrasting fears of positive and negative evaluation as facets of social anxiety. *Journal of Social and Clinical Psychology*, 29(1), 68 - 94. <https://doi.org/10.1521/jscp.2010.29.1.68>
- Wieser, M. J., & Moscovitch, D. A. (2015). The effect of affective context on visuocortical processing of neutral faces in social anxiety. *Frontiers in Psychology*, 6, Article 1824. <https://doi.org/10.3389/fpsyg.2015.01824>
- Yang, J.-W., Yoon, K. L., Chong, S. C., & Oh, K. J. (2013). Accurate but pathological: Social anxiety and ensemble coding of emotion. *Cognitive Therapy and Research*, 37(3), 572-578. <https://doi.org/10.1007/s10608-012-9500-5>
- Yoon, K. L., Yang, J.-W., Chong, S. C., & Oh, K. J. (2014). Perceptual sensitivity and response bias in social anxiety: An application of signal detection theory. *Cognitive Therapy and Research*, 38(5), 551-558. <https://doi.org/10.1007/s10608-014-9619-7>
- Yoon, K. L., & Zinbarg, R. E. (2007). Threat is in the eye of the beholder: Social anxiety and the interpretation of ambiguous facial expressions. *Behaviour Research and Therapy*, 45(4), 839-847. <https://doi.org/10.1016/j.brat.2006.05.004>
- Yoon, K. L., & Zinbarg, R. E. (2008). Interpreting neutral faces as threatening is a default mode for socially anxious individuals. *Journal of Abnormal Psychology*, 117(3), 680-685. <https://doi.org/10.1037/0021-843X.117.3.680>

원고접수일: 2025년 9월 11일

논문심사일: 2025년 10월 2일

게재결정일: 2025년 11월 18일

한국심리학회지: 건강
The Korean Journal of Health Psychology
2026. Vol. 31, No. 1, 1 - 21

Interpretation Bias Toward Social Situations in Social Anxiety: Using Combined Facial and Vocal Cues

Se-eun You

Department of Psychiatry

Dongguk University Ilsan Hospital

Jisu Choi

Department of Psychology

The Catholic University of Korea

Previous studies on interpretation bias in individuals with social anxiety have mainly used tasks that ask participants to interpret written scenarios or assess the emotional valence of facial stimuli. This study aimed to investigate the unique patterns of interpretation bias linked to different sub-factors of social anxiety by utilizing multisensory stimuli, specifically combining facial and vocal cues in social scenarios. Eighty-nine undergraduate and graduate students participated, completing measures assessing the cognitive (fear of negative evaluation) and experiential (social interaction anxiety) dimensions of social anxiety. In the experiment, participants first viewed an ambiguous social scenario, followed by simultaneous presentation of a facial photograph and a voice recording. The facial and vocal stimuli expressed either happiness or neutrality, resulting in four distinct conditions. Participants rated the perceived positivity or negativity of the combined audiovisual stimuli within the social context using a visual analogue scale (VAS). The results indicated a significant association between the cognitive aspect of social anxiety and a negative interpretation bias, particularly when neutral vocal cues were used. Conversely, the experiential aspect was linked with a tendency to perceive happy facial expressions as less positive. These findings imply that interpretation bias in social anxiety is not a singular process. Implications of these findings for theory and clinical practice are discussed.

Keywords: social anxiety, ambiguous situation, audiovisual stimuli, multisensory processing, interpretation bias