

## 로봇의 성별에 따른 행동표상, 마음지각과 실수에 대한 용서\*

신 홍 임†


국립순천대학교

선행연구에서는 여성로봇이 따뜻한 정서를 갖고 있는 대상으로 인식되어 남성로봇보다 더 선호됨을 보고한다. 본 연구에서는 여성로봇에 대한 선호가 주도적 행동차원, 마음지각 및 실수에 대한 용서에서도 나타나는지를 검증하려 한다. 연구 1(N = 62)에서는 로봇의 성별을 조작하기 위해 남성이름 또는 여성이름을 제시하는 조건으로 조작한 후, 상위 행동수준의 적합성, 인간과의 유사성 및 마음지각을 측정하였다. 연구 2(N = 60)에서는 네 가지 유형의 행동주체(남성인간 vs. 여성인간 vs. 남성로봇 vs. 여성로봇)에 대한 기술문을 제시한 후 행동주체의 성별에 따라 상위 행동수준의 적합성, 마음지각 및 실수를 용서하는 정도를 측정하였다. 그 결과 연구 1에서는 여성로봇의 행동이 남성로봇의 행동보다 행동주체의 의지와 목표가 부각되는 상위수준에서 더 낮게 평정되었다. 연구 2에서는 여성로봇의 실수를 받아들이는 정도가 남성로봇의 실수를 받아들이는 정도보다 더 낮게 나타났다. 본 연구는 과제수행의 차원에서 여성에 대한 고정관념이 여성로봇에게도 전이될 가능성을 보여준다.

주제어 : 여성로봇, 행동표상, 마음지각, 실수에 대한 용서

\* 본 연구는 2023년 한국여성심리학회 학술연구공모전의 최우수상에 선정되어 지원되었음.  
본 논문에 자세한 도움말을 주신 심사위원님께 감사드립니다.

† 교신저자: 신홍임, 국립순천대학교 자유전공학부, (57922) 전라남도 순천시 중앙로 255, 순천대학교 기초  
교육관 317호, E-mail: shin7038@naver.com

 Copyright © 2023, The Korean Society for Woman Psychology. This is an Open Access article This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial Licenses (<https://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.

최근 개봉된 “She, Chef(2022)”의 다큐멘터리에는 전통적으로 여성의 영역으로 구분되는 요리에서도 여성이 전문세프로서 인정받는 것이 거의 불가능함을 보여주고 있다. 여성 요리사와 남성 요리사에 대한 처우가 다를 뿐만 아니라 세계적으로 유명한 식당에서 메뉴의 개발, 계획 및 질관리를 총괄하는 최고 주방장은 거의 모두 남성이다. 이러한 현상은 남성과 여성의 과제수행능력이 성별에 따라 다르게 평가되고 있을 가능성을 보여준다. 선행연구(Heilman, 2001; Heilman, Block, & Martell, 1995; Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012)에 의하면 남성은 일반적으로 여성보다 더 주도적이고, 더 능력이 있는 것으로 평가된다.

지금까지 사회과학의 연구(Devine, 1989; Heilman, 2001)에서는 우리가 타인을 만났을 때, 남성과 여성에 대한 고정관념이 자동적으로 활성화됨을 보여주었다. 예를 들어, 남성에게는 주도적이고, 독립적이며, 강한 추진력이 연상되는 데 비해, 여성에게는 감정적이며, 따뜻한 돌봄과 공감의 속성이 연결된다. 여성은 소통의 영역에서는 남성보다 더 호감을 받지만, 과제수행에서는 남성보다 주도성에서 대체로 더 낮게 평가된다(Borau, Otterbring, & Laporte, 2021; Eyssel & Hegel, 2012). 그렇다면 남성과 여성에 대한 고정관념은 성별이 부여된 로봇에게도 적용될까? 본 연구에서는 전통적 고정관념이 로봇에게 적용된다면, 로봇의 성별에 따라 동일한 행동이 다르게 지각될 것인지를 문제를 검증하려 한다. 또한 과제수행상황에서 로봇이 실수를 했을 때, 로봇의 성별에 따라 실수가 받아들여지는 정도가 달라지는지를 분석하려 한다.

### 인간-로봇의 사회적 상호작용

인간-로봇의 상호작용에 대한 연구는 다양한 학문분야에서 점점 더 관심을 받고 있다(Huang & Rust, 2018). 선행연구(Esteva, Kuprel, & Novoa, 2017; Grove, Zald, Lebow, Snitz, & Nelson, 2012)에서는 로봇이 경제, 산업 또는 의학분야에서 증거기반 알고리즘(algorithm)을 토대로 인간보다 더 정확한 의사결정을 수행함을 보여주고 있다. Grove 등(2000)은 의료분야 및 인간행동분야의 136개 선행연구를 메타분석한 결과 로봇의 알고리즘기반 의사결정이 인간의 직관기반의 의사결정보다 10% 정도 더 정확한 수행도를 나타냄을 보고하였다.

그러나 인간-로봇의 상호작용에서 로봇에 대한 태도는 아직까지 부정적이다. 예를 들어, 로봇의 알고리즘기반 의사결정은 인간의 직관기반 의사결정보다 오류발생이후에 신뢰감을 더 잃는 경향이 있다(Dietvorst, Simmons, & Massey, 2015). 또한 선행연구에서는 로봇의 의사결정이 인간의 의사결정보다 덜 선호되고(Diab, Pui, Yankelevich, & Highhouse, 2011), 중요성의 정도에서 더 낮게 인식됨을 보여주고 있다(Önkal, Goodwin, Thomson, Gönül, & Pollock, 2014). 이에 따라 인간-로봇의 상호작용을 개선시키기 위해 로봇에게 인간의 속성을 부여하는 의인화 연구가 실시되고 있다. 예를 들어, Yam 등(2020)은 로봇이 감정을 느끼는 것처럼 지시문에 기술하고, Gray와 Wegner(2012)는 로봇의 외관을 인간과 유사하게 제시하면, 연구참가자들이 로봇을 인간의 마음과 유사하게 지각하며, 로봇에 대한 호감이 증가함을 보여주었다. 또한 선행연구(Laban, 2021; Lee, Lee, & Sah, 2020)에서는 대화형 인공지능(chat bot)과의 상호작용을 개선시킬 수 있는 변인들에 대해 보고하고 있다.

Lavan(2021)의 연구에서는 대화형 인공지능에 게 사회적 단서(예: 이름, 나이, 제스처 등)를 추가하여 의인화하였을 때, 인공지능에 대한 태도와 인공지능을 사용하려는 의향이 긍정적으로 변화함을 보여주었다. 이러한 결과는 다양한 사회적 단서가 인간-인공지능과의 상호작용에 영향을 끼칠 가능성을 보여준다. 본 연구에서는 사회적 단서 중의 하나인 성별이 로봇에게 부여되었을 때, 로봇에 대한 태도가 달라지는지를 검증하려 한다.

### 로봇의 성별, 사회적 범주화, 마음지각과 행동식별

사회적 범주화에 관한 선행연구(Devine, 1989)에서는 한 개인의 성별이 인상형성에서 중요한 단서로 작용함을 보여주었다. 로봇에게 시각적/청각적 단서(예: 이름, 얼굴, 신체 특징, 목소리 등)로 성별을 부여한다면 로봇의 성별에 따라 사회적 범주화가 활성화되어 인간-로봇의 상호작용이 달라질 수 있다. 예를 들어, Eyssel과 Hegel(2012)의 연구에서는 머리 길이가 긴 로봇은 참가자에게 감정이나 소통과 연관된 여성성을 더 많이 활성화시키고, 머리 길이가 짧은 로봇은 주도성과 연관된 남성성을 더 많이 떠오르게 함을 보여주었다. Weißel 등(2022)의 연구에서는 요양병원의 서비스로봇에게 여성의 외관적 특징을 부여하면, 환자들이 로봇에 대해 더 편안한 느낌을 갖고, 로봇에 대한 호감이 증가함을 보고하였다<sup>1)</sup>. Borau 등(2021)에서도 로봇에게 여성/남성의 이

름이나 얼굴사진을 제시하여 성별을 부여하면, 여성로봇에게 남성로봇보다 따뜻함이나 다정함과 같은 정서가 더 많이 연상되는 것으로 나타났다. 이 결과는 원래 성별이 없는 로봇에게 성별이 부여되면 성 고정관념이 활성화되어 인간-로봇의 상호작용이 달라지고 있음을 보여준다.

그렇다면 성 고정관념은 성별이 부여된 로봇의 행동을 인식하는 데도 영향을 끼칠까? 행동식별이론(action identification theory; Vallcher & Wegner, 1989)에 의하면, 인간은 상대방의 행동을 관찰하고, 작은 단서를 기반으로 행동을 판단한다. 이 과정은 상대방의 동작이 단순한 움직임이 아니라 전제에서 시작된다. 객체는 스스로 행동할 수 없기 때문에 상대방의 행동을 관찰한다는 것은 관찰자가 상대방의 행동의 주체로서 지각함을 의미한다. 예를 들어, 한 개인이 지우개를 물건으로 인식하며, 지우개의 행동을 관찰하고 이해하려 한다는 것은 말이 되지 않지만, 똑똑한 기계를 행동의 주체로 전제한다면, 기계의 행동에서 목표나 행동에 대한 책임을 식별하려 시도할 수 있을 것이다(Kozak, Wegner, & Marsh, 2006).

행동식별이론에 의하면 행동은 높은 행동수준 또는 낮은 행동수준에서 표상된다. 주체의 목표가 지각되면, 높은 행동수준에서 표상되고, 행동의 구체적 실행이 부각되면, 행동은 낮은 행동수준에서 표상된다. 예를 들어, 청소행동의 목표가 부각되면서, 청결한 공간을 유지하기 위해 집을 청소한다고 생각한다면 높은 행동수준인 반면, 청소기로 거실바닥의 먼지를 빨아들이는 특정 세부행동이 강조되면, 낮은 행동수준에서 표상될 것이다.

또한 동일한 행동도 행동주체가 누구인지에 따라 높은/낮은 표상수준의 적절성에 대한

1) Weißel 등(2021)은 로봇과의 상호작용을 촉진시키기 위해 로봇에게 성별을 부여하여 의인화하는 것은 로봇의 디자인을 통해 전통적 성 고정관념을 강화시키는 윤리적 문제를 발생시킬 가능성을 논의하였다.

평정이 달라질 수 있다. 예를 들어, 선행연구(Kozak, Marsh, & Wegner, 2006)에서는 관찰자가 다양한 행동주체(예: 좋아하는 vs. 싫어하는 사람, 경제적 곤궁에 처한 실업자, 자기 자신 등)의 마음을 어떻게 지각하는지에 따라 높은 행동수준의 적합성에 대한 평정이 달라짐을 보고하였다. 관찰자와의 관계가 긍정적일수록 행동주체가 목표와 의지를 갖고, 외부환경에 영향력을 주는 높은 수준에서 행동이 표상되는 경향이 증가하였다. 반면 관찰자가 행동주체에 대해 부정적으로 생각한다면, 동일한 행동도 주체의 자율적 의지 없이 단순히 반복적으로 수행하는 것으로만 인식되었다.

일반적으로 로봇은 자신 고유의 목표나 의지를 갖고 있지 않은 것으로 인식되고 있기 때문에(Ward, Olsen, & Wegner, 2013), 로봇의 행동은 미래의 계획이나 행동의 목표가 표상되는 높은 수준에서 인간의 행동보다 적절성에서 낮게 평정될 것이다. 그러나 로봇의 행동에 대한 표상수준은 관찰자가 로봇을 어떻게 생각하는지에 따라 달라질 수 있다. 로봇에 대한 긍정적 태도는 로봇이 버튼을 누르는 행동을 유익한 자료검색이라는 높은 표상수준에서 보게 할 수 있다. 반면 관찰자가 로봇을 단순한 플라스틱이나 고철몽치로 부정적으로 본다면, 버튼을 누르는 동작을 어떤 목표없이 단순한 프로그램이 실행되는 것으로 지각할 것이다. 지금까지 선행연구(Kim & Duhachek, 2020; Kozak et al., 2006)에서는 인간-인간의 상호작용에서 행동의 표상수준을 분석하였기 때문에, 본 연구에서는 인간-로봇의 상호작용으로 선행연구를 확장하여 로봇의 행동도 관찰자와 행동주체의 관계에 따라 상위수준 행동 표상의 식별이 달라지는지를 분석하려 한다.

### 로봇의 성별, 마음지각과 실수에 대한 용서

마음지각이론(Gray, Gray & Wegner, 2007)에서는 관찰자와 행동주체의 관계에 따라 주체의 마음을 지각하는 정도가 달라짐을 설명한다. 마음지각에는 경험성(experience)과 주도성(agency)의 두 가지 하위차원이 있다. 경험성은 감정을 느끼는 능력인데 비해 주도성은 주체가 자신의 의지대로 자율적으로 행동하는 능력이다. 선행연구(Gray et al., 2007)에서는 인간이 다른 비인간(예: 로봇, 동물)보다 마음지각에서 더 높게 평정됨을 보여주었다. 또한 인간의 범주에서도 성인 남성과 관찰자 자신은 성인 여성, 소년과 소녀에 비해 주도성에서 더 높게 평정되었다. Schein과 Gray(2012)의 연구에서는 얼굴모양에 따라서도 마음지각이 달라짐을 보였주었다. 눈과 코 중에서 눈이 없는 얼굴의 경우에 마음지각이 더 낮게 나타났다.

마음지각의 문제가 중요한 이유는 상대방의 마음을 어떻게 보는지는 상대방에 대한 도덕적 처우와 연결되기 때문이다(Gray, Young, & Waytz, 2012). 마음지각이론(Gray 등, 2007)에 의하면 주도성이 높을수록 상대방을 행위주체로 인식하고, 경험성이 높을수록 상대방을 피행위자로 지각한다. 관찰자가 상대방의 주도성을 높게 인식할수록 상대방을 자신의 행동에 대한 책임을 갖고 있는 존재로서 생각하며, 존중할 것이다. 또한 관찰자가 고통과 즐거움을 느끼는 상대방의 경험성을 높게 인식할수록 상대방에 대한 가해행동을 꺼리게 될 것을 예측할 수 있다. Gray 등(2007)의 연구에서는 로봇이 다른 대상보다 주도성과 경험성에서 가장 낮았다. 이 결과는 로봇은 행위주체나 피행위자라도 거의 인식되지 않기 때문에 지금까지 도덕적 고려의 대상이 아니었음을 보

여준다.

그러나 마음지각이 상대방에 대한 태도에 따라 변화하는 것이라면, 로봇을 마치 인간처럼 제시하는 방식이 로봇에 대한 마음지각에 영향을 끼칠 수 있을까? 선행연구(예: Gray & Wegner, 2012; Wang & Krumhuber, 2018; Yam et al., 2020)에서는 의인화된 로봇(예: 인간과 유사한 외관, 이름, 정서의 소통)이 의인화되지 않은 기계 로봇보다 마음지각에서 더 높게 나타남을 보여주었다. 또한 마음지각에서 고통과 즐거움을 느끼는 경험성이 자율적으로 계획하고 행동하는 주도성보다 로봇을 인간과 유사하게 지각하는 데 더 중요한 영향을 끼쳤다. 그렇다면 성별이 부여된 로봇에 대한 마음지각은 어떻게 나타날 것인가? Borau 등(2020)의 연구에서는 여성로봇이 남성로봇보다 인간과 더 유사하게 지각됨을 보고하였다. 이것은 전통적인 성 고정관념이 활성화되어 여성로봇이 따뜻함이나 돌봄의 정서를 연상시키는 결과로 해석할 수 있다. 또한 여성로봇을 남성로봇보다 인간과 더 유사하게 지각하는 데는 마음지각의 경험성이 주도성보다 더 중요한 역할을 하고 있음을 보여준다. 그러나 로봇을 인간과 유사하게 지각하는 것은 로봇의 행동을 자율적인 것으로 상위수준에서 표상하고, 로봇의 실수를 더 받아들여주는, 로봇에 대한 긍정적 태도와 반드시 연관될 것인가? 선행연구(Borau 등, 2020)에서는 여성로봇에 대한 호감을 여성로봇이 남성로봇보다 더 인간과 유사하게 인식되는 것으로 설명하였다. 본 연구에서는 선행연구를 확장하기 위해 여성로봇을 인간과 유사하게 지각하는 것과 연관된 행동표상의 수준, 마음지각 및 실수에 대한 용서를 분석하려 한다. 로봇을 인간과 유사하게 지각하는 것은 복잡한 과정이어서

긍정적 태도뿐만 아니라 부정적 태도가 연관될 수 있다. 예를 들어, Gray와 Wegner(2012)에 의하면, 로봇이 인간이 아니면서 인간처럼 보이는 것은 인간에게 기이한 느낌을 갖게 할 수 있다. 또한 Borau 등(2020)에서는 여성로봇에 대한 호감은 여성로봇에게 마치 인간과 같은 따뜻함의 감정이 연상되어 여성로봇이 인간의 욕구를 채워줄 것이라는 기대를 갖게 하기 때문이다. 그러나 여성로봇이 실수를 하는 상황에서는 능력의 차원에서 여성에 대한 부정적 고정관념(예: 소극적, 비효율적 등이 여성로봇에게 적용되어 남성로봇보다 실수를 받아들여주는 정도를 더 낮출 수 있다.

요약하면 본 연구에서는 로봇의 성별, 행동식별 수준 및 마음지각의 관계에 대해 검증하려 한다. 연구가설은 다음과 같다.

첫째, 남성로봇과 여성로봇은 행동주체의 목표가 부각되는 상위수준 행동표상과 마음지각의 경험성/주도성에서 차이가 나타나는가(연구 1)? 선행연구(예: Eyssel과 Hegel, 2012)에서는 로봇에게 성별을 부여하였을 때, 전통적인 성 고정관념이 활성화되어 남성로봇은 여성로봇보다 더 주도적으로 인식되고, 여성로봇은 남성로봇보다 더 감정적으로 잘 소통하는 것으로 평정됨을 보고하였다. 본 연구에서는 남성의 주도성/여성의 정서와 같은 속성이 로봇에게도 적용된다면, 행동주체의 자율적 의지가 부각되는 상위수준 행동식별에서 남성로봇이 여성로봇보다 더 높게 평정될 것을 예측한다.

둘째, 행동주체가 실수를 하는 상황에서 여성로봇은 남성로봇과 다르게 평정되는가(연구 2)? 선행연구에서는 여성로봇에 대한 선호가 남성로봇에 대한 선호보다 높은 것은 여성로봇에게 따뜻한 감정이 있는 것으로 인식되기

때문으로 해석하였다. 또한 Yam 등의 연구(2021)에서는 (로봇의 주도성이 아니라) 로봇에게 감정이 있다고 설명하는 의인화가 로봇의 실수를 받아들이는 정도를 증가시킴을 보고하였다. 본 연구에서는 선행연구와 차별화하기 위해 과제수행상황에서 로봇의 동일한 실수가 로봇의 성별에 따라 어떻게 받아들여지는지를 분석하려 한다. 선행연구와 같이 로봇의 감정을 강조하는 의인화가 실수의 용서에 영향을 끼친다면 여성로봇의 실수가 더 많이 받아들여질 것이다. 반면 본 연구에서는 과제수행의 상황이 능력의 차원을 부각시키기 때문에 과제수행의 상황에서 여성의 능력에 대한 부정적 고정관념이 활성화된다면, 여성로봇의 실수를 받아들이는 정도가 남성로봇보다 더 낮아질 것을 예측한다.

## 연구 1

연구 1의 목적은 전통적인 성고정관념이 의인화된 로봇에게 적용된다면 남성로봇과 여성로봇이 상위수준 행동표상의 적절성 및 마음지각에서 차이가 나타나는지를 분석하는 데 있다.

## 방 법

### 참가자

4년제 대학교의 대학생 62명(여 = 28명, 평균연령 만 21.67세, 표준편차 = 2.42)이 대학교 홈페이지 게시판 공고를 보고 자발적 신청을 통해 참가하였다.

### 연구도구 및 절차

참가자는 남성로봇조건 또는 여성로봇조건에 무선적으로 할당되었다. 본 연구에서는 로봇의 성별을 이름으로 조작하는 선행연구(예: Nomura, 2017)를 참조하여 남성로봇에게는 ‘마이클’의 이름을, 여성로봇에게는 ‘헤나’의 이름을 제시하였다. 참가자는 우선 각 조건에서 로봇에 대한 기술문을 읽었다. 이 글에서는 로봇에 대해 우리의 시간을 효율적으로 사용하도록 도와주는 기능적 로봇으로 소개하였다(예: 마트에서 우리가 원하는 제품을 찾아주며, 제품에 대한 질문사항이 있을 때, 신속하게 응답해줌). 참가자는 로봇에 대한 글을 읽은 후, 이 로봇이 인간과 유사한 정도를 7점 척도(1: 전혀 비슷하지 않음, 7: 아주 비슷함)에 따라 응답하였다. 또한 여성로봇과 남성로봇의 성별 조작이 잘 되었는지를 확인하기 위해 Eyssel과 Hegel(2012)의 연구를 참고하여 위의 글에 제시된 로봇이 남성과 여성 중에서는 어느 쪽에 더 가까운지를 7점 척도에서 작성하도록 하였다(1: 남성에 아주 가까움, 4: 남성도 여성도 아님, 7: 여성에 가까움). 이후 참가자는 행동식별 질문지(부록 1)에 응답하였다. 이 질문지에서 참가자는 열일곱 개의 항목을 읽고, 각 항목에 기술된 내용이 행동 주체가 스스로 하기에 적절하다고 생각하는 정도를 7점 척도에서 응답하였다(1: 전혀 적절하지 않음, 7: 매우 적절함).

이 문항은 선행연구(Kozak et al., 2006)를 토대로 연구참가자가 로봇이 자유의지를 갖고 있다고 상상하면서, 로봇이 행동주체가 되었을 때, 부자연스럽지 않은 항목을 선정하여 사용하였다. 본 연구에서 사용한 열일곱 개 문항의 내적 합치도 알파계수는 .78로 나타났

다. 행동식별과제가 완료되면, 참가자는 마음지각척도(Gray et al., 2007)에 응답하였다. 이 질문지는 특정 대상에 대한 마음지각을 경험성(예: 분노를 느끼는 능력)과 주도성(예: 자신을 통제할 능력)의 하위영역에서 구분하여 측정한다(부록 1 참조). 참가자는 마음지각 질문지의 총 열 개 문항을 읽고, 참가자가 생각할 때, 위의 기술에서 제시된 로봇의 느낌과 생각이 이 문항에 제시된 능력과 연관된다고 생각되는 정도를 7점 척도(1: 전혀 연관되지 않음, 7: 매우 연관됨)에서 응답하였다. 마음지각 척도에서는 경험성과 주도성의 각 하위영역별로 문항의 총합을 구하여 문항 개수로 나눈 평균을 산출한다. 연구 1에서 마음지각 질문지의 내적 합치도는 경험성의 측면에서 .74, 주도성의 측면에서 .77로 나타났다. 마지막으로 참가자는 본 연구의 취지와 가설을 파악했는지에 대한 질문에 응답하였다<sup>2)</sup>. 이후 참가자에게 감사의 표시로 커피쿠폰을 제공하고, 본 연구를 종료하였다.

## 결과 및 논의

### 조작점검

연구 1에서는 이름을 통해 로봇의 성별 조작이 잘 되었는지를 확인하였다. 여성의 이름이 부여된 로봇( $M = 5.23, SD = 1.12$ )은 남성 이름이 제시된 로봇( $M = 2.56, SD = 1.32$ )보다 여성에 가까운 정도가 유의하게 높았다,  $t(60) = 4.23, p = .002$ .

2) 연구 1과 연구 2에서는 본 연구의 목적과 가설을 파악한 참가자가 없었기 때문에, 연구결과분석에서 모든 참가자의 자료가 사용되었다.

### 로봇의 성별, 상위수준 행동 적합성, 마음지각 및 인간과 유사한 정도

연구 1에서는 로봇의 성별에 따라 상위수준의 행동식별, 마음지각 및 인간과 유사하다고 인식되는 정도가 달라지는지를 분석하였다. 그 결과 남성로봇조건( $M = 4.42$ )에서 여성로봇조건( $M = 3.78$ )보다 상위수준의 행동식별에서 적합성점수가 유의하게 높게 나타났다,  $t(60) = 2.61, p = .012$ . 이것은 연구참가자가 여성로봇을 남성로봇보다 자율적 의지를 갖고 있는 행동주체로서의 능력을 상대적으로 부족하다고 평정한 결과로 해석할 수 있다. 또한 마음지각척도에서 남성로봇의 주도성( $M = 4.43$ )이 여성로봇의 주도성( $M = 3.79$ )보다 더 높게 평정되었다,  $t(60) = 2.31, p = .034$ . 마음지각의 경험성 차원에서는 여성로봇의 점수( $M = 2.79$ )가 남성로봇의 점수( $M = 2.02$ )보다 더 높게 나타났다,  $t(60) = 2.39, p = .020$ .

인간과 유사한 정도에서는 여성로봇( $M = 4.38$ )이 남성로봇( $M = 3.68$ )보다 더 높은 경향이 나타났다,  $t(60) = 1.838, p = .074$ . 인간과의 유사성에서 여성로봇이 남성로봇보다 더 높은 점수를 얻은 것은 여성로봇이 마음지각의 주도성보다는 경험성에서 더 높은 것과 연관되었을 가능성이 있다. 이것은 선행연구(Borau et al., 2021)의 결과와도 일치한다. 이 결과는 기계를 인간과 유사하게 지각하는 데 주도성보다 경험성이 더 중요한 역할을 하는 것으로 해석할 수 있다.

표 1과 같이 변인들간의 상관분석결과를보면, 상위수준의 행동식별과 마음지각의 주도성간에 유의한 상관관계가 나타났다,  $r = .385, p = .002$ . 이 결과는 관찰자가 보았을 때, 상위수준 행동표상의 적절성은 행동주체

표 1. 주요 변인의 평균과 상관분석결과(연구 1)

변인	M	SD	1	2	3	4
1. 상위수준 행동식별	4.07	.94		.175	.385**	-.007
2. 경험성(마음지각)	2.39	1.31			-.074	.247*
3. 주도성(마음지각)	4.12	1.39				.023
4. 인간과 유사한 정도	3.98	1.44				

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

의 마음에 주도적으로 생각하고 계획하는 능력이 있다고 믿는 정도와 관계가 있음을 보여준다. 또한 마음지각에서 감정을 느끼는 능력인 경험성의 차원이 인간과의 유사성과 유의한 관계가 나타났다,  $r = .247, p = .032$ .

연구 1의 결과는 여성에 대한 일반적인 고정관념(예: 감정적, 소극적 등)이 여성로봇에게도 적용되고 있을 가능성을 보여준다. 선행연구(Eysssel & Hegel, 2012)에서 제시한 여성로봇에 대한 호감이 여성에 대한 긍정적 고정관념(예: 공감, 따뜻함)이 적용된 것이라면, 연구 1의 여성로봇조건에서 상위수준의 행동식별과 주도성의 낮은 점수는 여성에 대한 부정적 고정관념(예: 소극적, 비효율적)이 여성로봇에게 적용되었을 가능성이 있다. 이에 따라 연구 2에서는 로봇이 실수를 하는 상황에서 여성로봇/남성로봇에 대한 태도가 어떻게 달라지는지를 비교하려 한다.

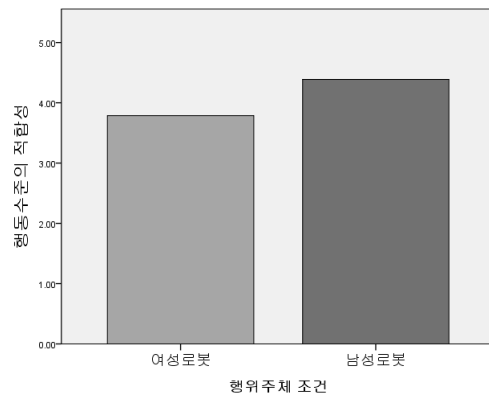


그림 1. 로봇의 성별에 따른 행동수준의 적합성 평정결과(연구 1)

용된다면, 여성로봇의 실수를 받아들이는 정도가 남성로봇에 비해 더 낮을 것을 예측할 수 있다.

## 방 법

### 연구 2

연구 2의 목적은 과제수행상황에서 여성로봇/남성로봇의 실수를 받아들이는 정도를 비교하는 데 있다. 능력의 차원에서 여성에게 적용되는 부정적 고정관념이 여성로봇에게 적

### 참가자

4년제 대학교의 대학생 60명<sup>3)</sup>(여 = 27명,

3) 각 조건별 참가자수는 다음과 같았다: 남성인간(n = 13, 여 = 6), 여성인간(n = 16, 여 = 6), 남성로봇(n = 15, 여 = 8), 여성로봇(n = 16,



평균연령 만 22.72세, 표준편차 = 3.14)이 대학교 홈페이지 공고를 통해 자발적으로 참가하였다. 연구참가에는 약 5분의 시간이 소요되었다.

#### 연구도구 및 절차

연구 2에서 참가자는 우선 네 가지 행동주체 조건 중의 한 조건에 무선적으로 배정되어 짧은 글을 읽었다. 글은 행동주체의 사진과 함께 제시되었다(부록 2 참조). 네 가지 조건의 글은 성별과 이름만 다르고, 글에 기술된 상황은 모두 동일하였다. 글에는 행동주체가 호텔 프론트에서 일하고 있으며, 객실 체크인, 짐 보관 및 근처 여행지에 대한 정보 안내 등의 업무를 수행하고 있음을 기술하였다. 또한 참가자가 행동주체의 실수를 받아들이는 정도를 측정하기 위해 행동주체가 객실 체크인 과정에서 실수로 객실번호를 혼동하여 다른 객실로 고객을 안내했다는 상황을 추가적으로 기술하였다. 이 글의 상황은 Yam 등(2020)의 연구를 참고하여 제시하였다. 각 조건의 연구 참가자에게는 이 글을 읽고 난 후, 만약 자신이 이 호텔을 방문했다면, 이 호텔직원의 실수를 용서할 의향을 7점 척도(1: 전혀 그렇지 않다, 7: 매우 그렇다)에서 기입하도록 지시하였다. 또한 여성로봇과 남성로봇의 조건에서 성별조작이 잘 되었는지를 점검하기 위해 연구 1과 동일한 문항에 응답하도록 하였다. 이후 참가자는 연구 1과 같은 상위수준 행동표상 질문지와 마음지각 질문지에 응답하였다. 마지막으로 참가자에게 연구의 목적과 가설을 파악했는지에 대해 작성하도록 한 후, 감사의 표시로 커피쿠폰을 제공하고, 연구를

여 = 7)

종료하였다.

## 결과 및 논의

#### 조작점검

연구 2에서 글과 사진을 통해 인간/로봇의 조건에 따라 성별 조작이 잘 되었는지를 확인하였다. 여성( $M = 6.13, SD = 0.92$ )은 남성( $M = 2.37, SD = 1.02$ )보다 여성에 가까운 정도가 유의하게 높았다,  $t(27) = 5.27, p < .001$ . 여성 이름의 로봇( $M = 4.93, SD = 0.97$ )은 남성 이름의 로봇( $M = 2.76, SD = 1.21$ )보다 여성에 가까운 정도가 유의하게 높았다,  $t(29) = 3.87, p = .014$ . 이에 따라 각 조건에서 이름과 사진을 통해 성별이 적절하게 조작된 것으로 보인다.

#### 행동주체, 참가자 성별<sup>4)</sup> 및 실수에 대한

4) 연구 2에서는 참가자 성별의 효과가 유의하게 나타나서, 이 변인을 포함시켜 이원변량분석을 수행하였다. 선행연구에서는 연구참가자 성별에 따라 성 고정관념 활성화 및 실수에 대한 용서의 정도가 다르게 나타날 수 있음을 보고하였다. 예를 들어, Gupta 등(2008)의 연구에서는 남성참가자에게 성공적인 남성 기업가에 대한 글을 제시하여, 암묵적으로 남성에 대한 고정관념(공격적, 주도적)을 활성화시키면, 남성참가자들이 향후 기업가가 되고자 하는 의향이 유의하게 증가하였다. (반면 여성참가자들에게는 직접적으로 여성기업가의 긍정적 특성(돌봄, 따뜻함)을 강조하면, 기업가가 되려는 의향이 증가하였다). 또한 Kalera와 Mróz (2022)의 연구에서는 여성참가자가 남성참가자보다 실수에 대한 용서에서 더 관대하게 나타났다. 연구자들은 이 결과에 대해 남성은 자신을 주도적이고, 타인으로부터 독립적으로 표상하는 반면,

용서

연구 2에서 행동주체와 참가자 성별에 따라 실수에 대한 용서의 정도가 달라지는지를 확인하기 위해 이원변량분석을 수행하였다. 그 결과 행동주체의 주효과가 유의하였다,  $F(3, 52) = 7.429, p < .001, \eta^2 = .300$ . 표 3과 같이 Scheffe 사후검정에 의하면 여성로봇의 실수를 받아들이는 정도가 가장 낮게 나타났다. 참가자 성별의 주효과도 유의하였다,  $F(1, 52) = 14.574, p < .001, \eta^2 = .219$ . 여성 참가자 ( $M = 5.62$ )는 남성 참가자( $M = 4.32$ )보다 행위주체의 실수를 받아들이는 정도가 유의하게 더 높았다. 또한 그림 2와 같이 행동주체조건(인간남성 vs. 인간여성 vs. 남성로봇 vs. 여성로봇)과 참가자성별의 상호작용이 유의하였다,  $F(3, 52) = 15.674, p < .001, \eta^2 = .475$ . 남성 참가자는 남성로봇의 실수 ( $M = 5.97$ )를 여성로봇의 실수( $M = 1.03$ )보다 더 받아들이는 경향이 있었다,  $t(14) = 6.94, p < .001$ . 이에 비

해 여성 참가자는 여성로봇의 실수( $M = 6.11$ )를 남성로봇의 실수( $M = 4.75$ )보다 더 용서하는 경향이 나타났다,  $t(13) = 2.09, p = .045$ .

연구 2의 결과는 참가자수가 총 60명으로 네 가지 조건에서 참가자 성별로 각각 분석할 때, 참가자수가 불충분하다는 한계점이 있다. 표본크기가 충분히 크지 않아서 정규성 검증을 실시하여 Shapiro-Wilk의 유의확률을 각 조건별로 검토하였다. 그 결과 참가자의 성별에 따라 남성/여성으로 구분한 조건에서는 유의확률이 .05보다 높아서 분포의 정규성이 확인되었다. 이에 비해 행위주체조건의 정규성 검증에서는 인간남성조건과 남성로봇 조건에서 Shapiro-Wilk의 유의확률이 .05보다 낮아서 정규분포가 아닌 것으로 나타났다. 이에 따라 실수에 대한 용서를 행동주체와 참가자 성별로 구분하여 분석한 결과에는 타당도의 한계가 있다.

실수에 대한 용서, 상위수준 행동표상 및 마음지각의 관계

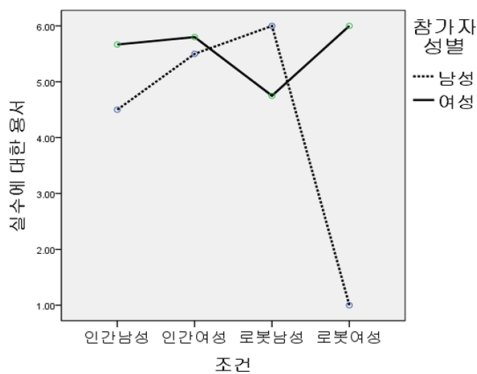


그림 2. 행동주체조건과 참가자 성별의 상호작용 (연구 2)

연구 2에서 실수에 대한 용서, 상위수준 행동표상 및 마음지각 간에 유의한 상관관계가 나타나는지를 분석하였다. 표 2와 같이 상관분석을 수행한 결과, 실수에 대한 용서와 상위수준 행동표상 간에 유의한 정적 상관관계가 나타났다,  $r = .42, p < .01$ . 또한 행동주체의 주도성을 지각하는 정도가 클수록 실수를 받아들이는 정도가 더 높게 나타났다,  $r = .43, p < .01$ . 또한 상위수준 행동표상은 마음지각의 주도성과 유의한 정적 상관관계에 있었다,  $r = .58, p < .001$ .

여성은 자신을 타인과 소통하는 관계성에 더 초점을 맞추어 표상하는 것으로 해석하였다.

표 2. 주요 변인들의 평균 및 상관분석 결과(연구 2)

변인	M	SD	2	3	4
1. 실수에 대한 용서	5.12	1.53	.42**	.17	.43**
2. 상위수준 행동표상	3.76	1.11		.13	.58***
3. 마음지각(경험성)	3.12	.89			.37*
4. 마음지각(주도성)	4.66	1.19			

\*  $p < .05$ , \*\*  $p < .01$

표 3. 행동주체 조건에 따른 실수에 대한 용서의 분석결과(연구 2)

(I)조건	(J)조건	평균차(I-J)	표준오차	유의확률
인간남성	인간여성	-.454	.492	.360
	남성로봇	-.118	.463	.800
	여성로봇	1.702*	.505	.001
인간여성	인간남성	.454	.492	.360
	남성로봇	.337	.457	.465
	여성로봇	2.156*	.499	< .001
남성로봇	인간남성	.118	.463	.800
	인간여성	-.337	.457	.465
	여성로봇	1.819*	.470	.000
여성로봇	인간남성	-1.702*	.505	.001
	인간여성	-2.156*	.499	.000
	남성로봇	-1.819*	.470	.000

\*  $p < .05$ , \*\*  $p < .01$

행동주체, 참가자 성별 및 상위수준 행동표상

연구 2에서 행동주체와 참가자 성별에 따라 상위수준 행동표상의 적절성이 달라지는지를 확인하기 위해 이원변량분석을 수행하였다. 그 결과 행동주체의 주효과가 유의하였다,  $F(3, 52) = 6.648, p = .001, \eta^2 = .277$ . 행동주체 조건 간의 차이를 분석하기 위해 독립표본

t-검정을 수행한 결과, 여성로봇의 상위수준 행동에 대한 적합성 평정( $M = 3.19$ )은 남성로봇(4.06)보다 유의하게 낮았다,  $t(29) = 2.13, p = .045$ . 이에 비해 남성로봇의 상위수준 행동에 대한 적합성평정은 남성인간( $M = 4.42$ )이나 여성인간( $M = 3.92$ )에 비해 유의한 차이가 나타나지 않았다. 또한 참가자 성별의 주효과도 유의하였다,  $F(3, 52) = 5.850, p = .019$ ,

$\eta^2 = .277$ . 남성 참가자( $M = 3.41$ )는 여성 참가자( $M = 4.02$ )보다 상위수준 행동표상의 적합성을 전반적으로 더 낮게 평정했다. 행동주체조건(인간남성 vs. 인간여성 vs. 남성 로봇 vs. 여성로봇)과 참가자성별의 상호작용도 유의하게 나타났다,  $F(3, 52) = 5.032, p < .001, \eta^2 = .225$ . 남성참가자( $M = 2.21$ )는 여성로봇의 행동을 여성참가자( $M = 4.15$ )보다 상위수준에 덜 적합한 것으로 평정하였다,  $t(14) = -3.13, p = .010$ . 이에 비해 남성로봇의 행동에 대한 평정에서는 남성참가자( $M = 4.50$ )와 여성참가자( $M = 4.38$ ) 간에 유의한 차이가 나타나지 않았다,  $t(13) = .224, p = .826$ .

연구 2에서는 네 가지 행동주체의 조건(남성인간, 여성인간, 남성로봇, 여성로봇) 중에서 여성로봇조건의 실수에 대한 용서가 가장 낮았다. 이 결과는 여성로봇의 실수가 남성로봇의 실수보다 더 엄격하게 처리될 가능성을 보여주었다. Yam 등(2020)의 연구에서는 로봇의 정서성을 강조하면, 실수에 대한 용서의 정도가 증가함을 보여준 데 비해, 본 연구에서는 여성로봇에게서 연상되는 따뜻함의 정서가 실수에 대한 용서를 증가시키는 데 영향을 끼치지 않은 것으로 보인다. 반면 행동주체의 마음지각에서 주도성은 실수에 대한 용서와 정적 상관관계를 갖고 있음이 나타났다. 이와 일치하게 상위수준 행동표상에서 더 높은 점수를 받은 남성로봇의 실수는 여성로봇의 실수보다 더 관대하게 받아들여지는 경향이 나타났다. 이 결과는 과제수행능력과 연관된 전통적 성고정관념이 여성로봇과 남성로봇에게 활성화되었을 가능성을 보여준다. 여성로봇은 업무를 수행할 때 소극적이고, 남성로봇은 주도적이라는 고정관념이 활성화되면, 과제수행 상황에서 오류가 발생했을 때, 여성로봇이 남

성로봇보다 더 불리한 위치를 차지할 가능성이 있다.

## 전체 논의

본 연구에서는 인간에게 적용되는 기존의 성 고정관념이 성별이 부여된 로봇에게도 활성화되는지를 분석하였다. 연구 1에서는 남성로봇이 여성로봇보다 행동주체의 자율적 의지가 부각되는 상위수준 행동표상에서 적합성이 더 높은 것으로 평정되었다. 또한 남성로봇은 자기주도적으로 계획하고 행동을 조절하는 마음지각의 주도성에서 더 높게 평정되었다. 이에 비해 여성로봇은 감정을 느끼는 능력인 경험성에서 남성로봇보다 더 높은 것으로 평정되었다. 이 결과는 성별이 부여된 로봇에게 나타나는 사회적 범주화를 보고한 선행연구(Eyssel & Hegel, 2012)의 결과와도 일치한다. 연구 2에서는 여성로봇의 실수를 받아들이는 정도가 다른 행동주체(남성인간, 여성인간, 남성로봇)보다 가장 낮게 평정되었다. 또한 참가자의 성별이 로봇의 성별과 일치하는 경우 로봇의 실수를 더 잘 용서하는 경향이 나타났다. 이 결과는 인간에게 적용되었던 전통적인 성 고정관념이 성별이 부여된 로봇에게도 적용되며, 자신이 속한다고 인식하는 집단에 대한 내집단편향(ingroup bias)이 인간-로봇의 상호작용에서도 나타날 가능성을 보여준다.

본 연구의 시사점과 제한점은 다음과 같다. 첫째, 선행연구(Borau 등, 2021)에서는 여성로봇이 따뜻함과 공감의 정서를 활성화시켜 남성로봇보다 더 선호됨을 보여주었다. 이에 비해 본 연구에서는 남성로봇이 행동주체의 의지가 부각되는 상위수준의 행동표상에서 여성

로봇보다 더 적합하게 인식됨을 보여주었다. 또한 여성로봇의 실수가 가장 잘 받아들여지지 않고 있는 것은 여성로봇에 대한 부정적 태도를 보여주며, 선행연구결과를 확장하는 결과로 볼 수 있다. 이 결과는 기존의 인간-인간의 상호작용에서 나타나는 성 고정관념이 유사한 방식으로 성별이 부여된 로봇에게 적용될 가능성을 보여준다. 또한 본 연구는 로봇에게 적용되는 전통적인 성 고정관념이 로봇에 대한 태도에 어떤 영향을 끼칠 것인지를 보여준다. 여성로봇의 감정성과 남성로봇의 주도성에 대한 지각은 여성/남성의 단순한 구분을 뛰어넘어 여성로봇과 남성로봇에 대한 편견을 재생산할 수 있음을 보여준다.

둘째, 본 연구에서는 로봇에게 적용되는 성 고정관념의 활성화를 자기보고식 질문지를 통해 측정하였다. 질문지 기법은 참가자가 사회적 긍정적 인상추구로 인해 솔직하게 응답하지 않을 가능성이 있기 때문에 결과해석에 한계가 있다. 또한 명시적 편견을 거의 갖고 있지 않은 참가자도 암묵적 편견을 갖고 있을 가능성이 있다(Lemus et al., 2018). 후속연구에서는 성 고정관념과 편견을 정교하게 측정할 수 있는 암묵적 도구(예: Borau et al., 2021)을 활용하여 참가자의 무의식적, 자동적 반응을 수집하는 것이 필요할 것으로 보인다.

셋째, 연구 2에서는 참가자 성별에 따라 여성로봇과 남성로봇에 대한 내집단편향이 나타났다. 여성참가자는 여성로봇의 실수를 더 잘 받아들이는 경향이 나타났고, 남성참가자는 남성로봇의 실수를 더 잘 받아들이는 경향이 나타났다. 선행연구(Rudman & Goodwin, 2004)에서는 여성에게서 일반적으로 내집단편향이 더 많이 나타남을 보고하는 데 비해 본 연구에서는 여성과 남성 모두에게서 내집단편향이

나타났다. 후속연구에서는 성별에 따른 내집단편향이 인간-로봇의 상호작용에서 성공과 실패의 다양한 상황유형에 따라 나타나는 양상을 세분화하여 분석해볼 필요가 있다.

최근 인간-로봇의 상호작용을 개선하기 위한 한 가지 방안으로 로봇을 의인화하고, 로봇에게 여성적인 특징(예: 얼굴, 신체형태, 목소리 등)을 부여하는 시도가 증가하고 있다(예: 소피아, 시리와 알렉스 등). 이것은 여성로봇이 돌봄과 소통의 속성을 연상시켜 호감을 주기 때문에 설명할 수 있다. 이에 비해 본 연구에서는 여성로봇의 행동이 과제수행의 차원에서 부정적으로 인식될 가능성을 보여주었다. 로봇에게는 원래 성별이 없지만, 이름이나 머리모양과 같은 작은 단서를 통해 성별이 부여된 여성로봇은 남성로봇보다 주도성에서 낮게 인식되었다. 이것은 성별이 부여된 로봇의 대중적 보급이 기존의 성 고정관념을 더 강화시키고, 성 불평등 시스템을 가속화시킬 가능성을 시사한다. 성별이 없는 로봇에게 디자인을 통해 성별의 외관적 특징을 부여하여 사용자에게 성 고정관념을 활성화시키는 것은 윤리적인가? 성별이 부여된 로봇에 대해 실수를 용서하는 정도가 달랐던 것은 로봇의 디자인이 고정관념의 재생산에 중요한 사회적 영향을 끼칠 가능성을 보여준다. 향후 로봇 디자인은 기존의 성 고정관념과 반대되는 로봇(예: 주도적 여성로봇, 어르신을 돌보는 남성로봇 등)을 제작하는 데 기여해야 되지 않는가? 후속연구에서는 로봇기술과 성평등 시스템의 관계에 대해 정교한 실험연구를 수행할 필요가 있다.

### 참고문헌

- Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition, 181*, 21-34.
- Borau, S., Otterbring, T., Laporte, S., & Fosso Wamba, S. (2021). The most human bot: female gendering increases humanness perceptions of bots and acceptance of AI. *Psychology & Marketing, 38*(7), 1052-1068.
- Diab, D. L., Pui, S.-Y., Yankelevich, M., & Highhouse, S. (2011). Lay perceptions of selection decision aids in US and Non-US samples. *International Journal of Selection and Assessment, 19*(2), 209-216.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114-126.
- Esteve, A., Kuprel, B., Novoa, R. et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*(7660), 115-118.
- Eyssel, F., & Hegel, F. (2012). (S)he's got the look: gender stereotyping of robots. *Journal of Applied Social Psychology, 42*(9), 2213-2230.
- Eyssel, F., Kuchenbrandt, D., Hegel, F., & de Ruitter, L. (2012). Activating elicited agent knowledge: how robot and user features shape the perception of social robots. *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, Paris, France, pp. 851-857.
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science, 315*(5812), 619-619.
- Gray, K., & Wegner, D. M. (2009). Moral typecasting: divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology, 96*(3), 505-520.
- Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: mind perception and the uncanny valley. *Cognition, 125*(1), 125-130.
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry, 23*(2), 101-124.
- Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*, 19-30.
- Gupta, V. K., Turban, D. B., & Bhawe, N. M. (2008). The effect of gender stereotype activation on entrepreneurial intentions. *Journal of Applied Psychology, 93*(5), 1053-1061.
- Heilman, M. E. (2001). Description and prescription: how gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues, 57*(4), 657-674.
- Heilman, M. E., Block, C. J., & Martell, R. (1995). Sex stereotypes: do they influence perceptions of managers? *Journal of Social Behavior and Personality, 10*(6), 237-252.
- Huang, M.-H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research, 21*(1), 155-172.
- Kaletka, K., & Mróz, J. (2022). Gender differences in forgiveness and its affective correlates. *Journal of Religion and Health, 61*, 2819-2837.

- Kim, T. W., & Duhachek, A. (2020). Artificial intelligence and persuasion: A construal-level account. *Psychological Science, 31*(4), 363-380.
- Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing?: action identification and mind attribution. *Journal of Personality and Social Psychology, 90*(4), 543-555.
- Laban, G. (2021) Perceptions of anthropomorphism in a chatbot dialogue: The role of animacy and intelligence. *International Conference on Human Agent Interaction, 21*, 305-310.
- Lee, S., Lee, N., & Sah, Y. (2020). Perceiving a mind in a chatbot: effect of mind perception and social cues on co-presence, closeness, and intention to use. *International Journal of Human - Computer Interaction, 36*(10), 930-940.
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences of the United States of America, 109*(41), 16474-16479.
- Nomura, T. (2017). Robots and gender. *Gender and the Genome, 1*(3), 18-25.
- Önköl, D., Goodwin, P., Thomson, M. E., Gönül, S., & Pollock, A. C. (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making, 22*(4), 390-409.
- Roesler, E., Heuring, M., & Onnasch, L. (2023). (Hu)man-like robots: the impact of anthropomorphism and language on perceived robot gender. *International Journal of Social Robotics, in press.*
- Rudman, L. A., & Goodwin, S. A. (2004). Gender differences in automatic in-group bias: why do women like women more than men like men? *Journal of Personality and Social Psychology, 87*(4), 494-509.
- Schein, C., & Gray, K. (2015). The eyes are the window to the uncanny valley: mind perception, autism and missing souls. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems, 16*(2), 173-179.
- Seyama, J., & Nagayama, R. S. (2007). The uncanny valley: effect of realism on the impression of artificial human faces. *Presence, 16*(4), 337-351.
- Vallacher, R., & Wegner, D. (1989). Levels of personal agency: individual variation in action identification. *Journal of Personality and Social Psychology, 57*(4), 660-671.
- Weßel, M., Ellerich-Groppe, N. & Schweda, M. (2021). Gender stereotyping of robotic systems in eldercare: an exploratory Analysis of ethical problems and possible solutions. *International Journal of Social Robotics, in press.*
- Weßel, M., Ellerich-Groppe, N., Koppelin, F., & Schweda, M. (2022). Gender and age stereotypes in robotics for elder care: ethical implications of stakeholder perspectives from technology development, industry, and nursing. *Science and Engineering ethics, 28*, 34.
- Yam, K. C., Bigman, Y. E., Tang, P. M., Ilies, R., De Cremer, D., Soh, H., & Gray, K. (2020). Robots at work: people prefer-and forgive-service robots with perceived feelings.

한국심리학회지 : 여성

*The Journal of Applied psychology*, 106(10),  
1557-1572.

1차원고접수 : 2023. 09. 04.

심사통과접수 : 2023. 09. 26.

최종원고접수 : 2023. 10. 15.



## Action identification of gendered robots, mind perception and forgiving mistakes

Hong Im Shin

Sunchon national university

Recent research reported that female robots are favored over male robots, mainly because female robots are judged as warmer and having emotions. This study examined whether positive attitudes toward female robots are changing depending a competence context such as identification of high level actions, agency of mind perception and forgiving mistakes. In Study 1, the participants (N=62) were randomly assigned to one of the two agent conditions (a male robot vs. a female robot) through gendering by the use of a male or a female name. The participants had to conduct an action identification task while rating how appropriate they found each action for the agent. In addition, they completed a mind perception scale. In Study 2, the participants (N=60) had to indicate the extent of how appropriate they rated each action in one of four agent conditions (male humans vs. female humans vs. male robots vs. female robots), mind perception and the extent of forgiving mistakes. According to the results, the extent of the appropriateness of high-level actions was rated much lower in the female robot condition than in the male robot condition. These findings implicate that negative attitudes toward women regarding competence can be transferred to female robots in the human-robot interaction.

*Key words* : female robots, action identification, mind perception, forgiving mistakes

**부록 1. 행동식별척도와 마음지각척도(연구 1, 연구 2)**

☞ 다음의 문항에 제시된 행동을 “로봇”이 스스로 하기에 적절한 정도를 7점 척도(1: 전혀 적절하지 않다, 4: 보통, 7: 매우 적절하다)에 따라 판단해주시시오.

1	로봇 마이클이 해야 할 일을 조직화하며 리스트를 작성한다	1-2-3-4-5-6-7
2	로봇 마이클이 지식을 습득하기 위해 읽는다.	1-2-3-4-5-6-7
3	로봇 마이클이 옷냄새를 제거하기 위해 세탁한다.	1-2-3-4-5-6-7
4	로봇 마이클이 땀감에 쓸 나무를 자르고 있다.	1-2-3-4-5-6-7
5	로봇 마이클이 카펫을 깔기 위해 방의 크기를 재고 있다.	1-2-3-4-5-6-7
6	로봇 마이클이 집을 깨끗하게 보여주기 위해 청소하고 있다.	1-2-3-4-5-6-7
7	로봇 마이클이 방을 새집처럼 보이도록 칠을 새로 하고 있다.	1-2-3-4-5-6-7
8	로봇 마이클이 거주하는 곳에서 계속 살기위해 방세를 낸다.	1-2-3-4-5-6-7
9	로봇 마이클이 집이 좋아보이도록 화분의 식물을 돌본다.	1-2-3-4-5-6-7
10	로봇 마이클이 집의 안전을 위해 문을 잠근다.	1-2-3-4-5-6-7
11	로봇 마이클이 좋은 경관을 보려고 나무에 오른다.	1-2-3-4-5-6-7
12	로봇 마이클이 자신의 지식을 보여주려고 시험을 본다.	1-2-3-4-5-6-7
13	로봇 마이클이 다정하게 보이려고 인사를 건넨다.	1-2-3-4-5-6-7
14	로봇 마이클이 신선한 야채를 얻으려고 정원을 가꾼다.	1-2-3-4-5-6-7
15	로봇 마이클이 시골을 보려고 차로 여행을 한다.	1-2-3-4-5-6-7
16	로봇 마이클이 아이를 가르치려고 말을 건다.	1-2-3-4-5-6-7
17	로봇 마이클이 집에 누가 있는지를 보려고 현관벨을 누른다.	1-2-3-4-5-6-7

☞ 아래의 문항을 읽고, 각 문항이 로봇의 느낌이나 생각과 연관된다고 생각하는 정도에 대해 자신의 의견을 표시해주시시오(1: 전혀 연관되지 않음, 4: 보통, 7: 매우 연관됨).

1	자신의 욕구를 가질 능력	1-2-3-4-5-6-7
2	즐거움을 느낄 능력	1-2-3-4-5-6-7
3	베고픔을 느낄 능력	1-2-3-4-5-6-7
4	고통을 느낄 능력	1-2-3-4-5-6-7
5	분노를 느낄 능력	1-2-3-4-5-6-7
6	자신을 통제할 능력	1-2-3-4-5-6-7
7	기억을 할 능력	1-2-3-4-5-6-7
8	자신의 느낌을 인식하는 능력	1-2-3-4-5-6-7
9	자신의 계획을 세울 능력	1-2-3-4-5-6-7
10	도덕적으로 행동할 능력	1-2-3-4-5-6-7

부록 2. 자극재료 예시(여성로봇조건(왼쪽)과 남성로봇조건(오른쪽))(연구 2)



로봇 헤나(여성, 31세)는 호텔 프론트에서 일을 합니다. 투숙객의 예약을 확인하고, 객실의 열쇠를 건네주며, 호텔의 시설과 근처의 좋은 여행지를 안내합니다. 그런데 오늘은 로봇 헤나가 실수로 투숙객의 이름을 혼동하여 다른 객실을 안내했습니다. 이 객실에는 이미 다른 투숙객이 있어서 늦은 저녁시간에 혼란이 발생했습니다.