

누가 마음을 갖고 있는가? 로봇에 대한 마음지각과 도덕 판단*

신 흥 입†

순천대학교

본 연구에서는 행위자와 피행위자에 대한 마음지각에 따라 도덕적 지위(moral standing)가 부여된다고 설명하는 양자간 도덕이론(dyadic moral theory)을 중심으로 다양한 대상에 대한 마음지각과 가해 행동의 비도덕성 판단의 관계를 분석하였다. 연구 1에서는 대학생 130명을 대상으로 다양한 대상에 대한 마음지각을 자기보고식 질문지를 통해 분석하였다. 그 결과 다양한 대상의 마음지각에서 성인 남자에 대한 마음지각이 가장 높게 나타났으며, 식물상태환자와 로봇에 대한 마음지각이 상대적으로 낮게 나타났다. 또한 대상에 대한 마음지각이 높을수록 이 대상에 피해를 주어서는 안 된다는 의사결정이 높게 나타났다. 연구 2에서는 대학생 114명을 대상으로 로봇의 기능적 가치의 제시방식(예: 경제적 로봇 vs. 사회적 로봇)에 따라 마음지각과 도덕 판단이 달라지는지를 분석하였다. 그 결과, 사회적 로봇에 대한 마음지각은 경제적 로봇에 대한 마음지각보다 더 높게 나타났다. 또한 사회적 로봇에 대한 도덕적 지위는 경제적 로봇에 대한 도덕적 지위보다 더 높게 나타났다. 이 결과는 마음지각이 인간뿐만 아니라 로봇에 대한 사회적 관계에 따라서도 달라질 수 있으며, 도덕성 판단에도 차별화된 영향을 끼칠 수 있음을 보여준다. 논의에서는 4차 산업혁명 이후 디지털화된 사회에서 마음지각이 인간-로봇의 상호작용에 끼치는 영향에 대한 시사점을 토론하고, 연구의 한계 및 후속연구의 방향을 다루었다.

주요어: 마음지각, 로봇, 양자간 도덕이론, 도덕적 지위

* 이 연구는 2019년 대한민국 교육부와 한국연구재단의 인문사회분야 신진연구지원사업의 지원을 받아 수행된 연구임(NRF-2019S1A5A8032211). 본 논문의 완성도를 높일 수 있도록 세심한 심사를 해준 익명의 심사위원님들께 감사드립니다.

† 단독저자 : 신흥입, 순천대학교 자유전공학부 부교수, 전라남도 순천시 중앙로 255 순천대학교 기초교육관317호, E-mail : shin7038@naver.com

■ 최초투고일 : 2020년 11월 29일 ■ 심사마감일 : 2021년 1월 19일 ■ 게재확정일 : 2021년 1월 22일

1. 서론

과학기술의 급속한 발달과 함께 로봇의 의미는 이제 우리에게 생산의 효율성 측면에서 경제적 이득만을 창출하는 기계에 국한되지 않는다. 경제적 로봇은 서서히 다양한 사회적 맥락(예: 교육, 의료, 복지서비스)에서 인간에게 연결의 욕구를 충족시켜주는 사회적 로봇으로 변화하고 있다(Wang & Krumhuber, 2018). 로봇과의 관계 변화는 로봇에 대한 마음지각과 도덕 판단에도 영향을 끼칠 수 있다. 지금까지 마음지각의 선행 연구(Gray, Gray, & Wegner, 2007)에서는 사회적 관계를 주로 인간-인간의 상호작용으로 전제함에 비해, 최근 로봇의 사회적 기능에 대한 관심은 사회적 관계를 인간-로봇으로 확장시켜 로봇에 대한 마음지각을 탐색해볼 필요성을 제시한다(Dahl, 2013). 이에 따라 본 연구에서는 도덕판단에서 행위자(agent)와 피행위자(patient)의 마음지각에 초점을 두는 양자간 도덕이론(dyadic moral theory)을 토대로 다양한 대상(예: 성인남자, 식물상태환자, 로봇)에 대한 마음지각을 탐색해보고자 하였다. 또한 로봇의 기능적 가치의 제시방식(예: 경제적 로봇 vs. 사회적 로봇)이 로봇에 대한 마음지각과 도덕적 지위)의 판단에 끼치는 영향을 분석하였다. 이를 통해 4차 산업혁명 이후 디지털화된 사회에서 인간-로봇의 상호작용에 대한 시사점을 도출하고자 하였다.

2. 이론적 배경

1) 양자간 도덕 이론(dyadic moral theory)과 마음지각

양자간 도덕이론(Gray & Wegner, 2012)에 의하면 도덕성은 행위자(agent)와 피행위자(patient)로 구성된다. 행위자는 특정한 행동을 하는 반면, 피행위자는 행위자의 행동에 의한 직접적 영향을 받게 된다. 예를 들어, 지갑을 훔치는 사람은 행위자이고, 이로 인해 돈을 분실하는 사람은 피행위자가 된다. 여기에서 옳고 그름에 대한 판단은 행위자와 피행위자의 마음에 대한 관찰자의 지각에 따라 달라지게 된다. 만일 내가 선하다고 생각했던 사람이 의도적으로 악한 행동을 하게 되면, 더 비도덕적으로 평가하고, 책임을 묻게 된다. 예를 들어, 지갑을 훔친 행위자가 테레사수녀였다면, 우리는 다른 일반인보다 더 비도덕적인 행위자로 판단을 내리게 될 것이다. 이것은 내가 상대방(테레사수녀)의 원래 마음상태(선한 행위자)를 알고 있다고 생각하는 것을 근거로 내리게 되는 판단이다. 한편 돈을 분실한 피행위자가 나의 친구라면, 내가 모르는 다른 사람에 비해 피행위자의 고통을 더 이해하게 될 것이다. 따라서 우리는 내가 안다고 생각하는, 상대방의 이전 마음상태 및 상대방과의 관계를 근거로, 상대방의 현재 마음상태에 대한 추론을 하게 된다(Ward et al., 2013). 우리는 이러한 추론을 근거로 행위자와 피행위자에 대한 도덕판단을 하고, 이들을 어느 정도로 도덕적으로 대할 것인지에 대해서도 결정을 내리게 된다(Loughnan et al., 2010).

1) 본 연구에서는 도덕적 지위(moral status)의 개념을 Miller(1994)에 따라 피행위자에게 도덕적으로 존엄하게 대하는 정도로 조작적으로 정의하였다. Miller는 피행위자가 감각능력을 갖고 있다고 지각된다면, 도덕적 지위의 부여는 결코 '영(0)'에 도달할 수 없다고 설명하였다.

Gray, Gray와 Wegner(2007)에 의하면 마음(mind)은 경험성(experience)과 주도성(agency)을 갖고 있는 능력으로 정의된다. Gray 등(2007)의 연구에서는 참가자들에게 다양한 대상(예: 6개월 된 아기, 성인 남자 어른, 식물상태 환자)을 제시하고, 이들 대상이 어느 정도 여러 가지 능력의 측면에 부합하는지를 평정하게 했다. 그 결과, 일반 성인의 마음은 다른 대상들과 비교할 때, 경험을 할 수 있는 능력(예: 감정을 느낌)과 주도적으로 행동할 수 있는 능력(예: 계획하기)에서 가장 높게 평정되는 것으로 나타났다. 이 연구에서 흥미로운 것은 특정 대상의 마음을 추상적으로 접근한 것이 아니라, 경험성과 주도성의 두 가지 독립된 차원²⁾에서 구체적으로 측정했다는 것이다. 경험성과 주도성의 측면은 인간성(humanness)의 두 가지 차원과도 연결된다. Fiske 등(2007)은 대인지각에서 따뜻함과 능력의 차원을 구분하여, 따뜻함과 능력이 높을수록 동일한 인간범주 안에서 가장 인간적인 처우를 받는다고 설명하였다. 또한 Haslam 등(2008)은 인간성을 인간 본성(human nature)과 인간 고유성(human uniqueness)으로 규정했다. 그에 의하면 인간의 정서는 기계와 구분되는 인간 본성(예: 고통을 느낌, 따뜻함)이며, 문화와 도덕성은 동물과 구분되는 인간 고유성(예: 합리적 이성, 깊이가 있음)이다. 따라서 마음지각에서 경험성의 차원은 ‘따뜻한 정서’ 및 ‘인간 본성’과 연결되며, 주도성의 차원은 ‘지적 능력’ 및 ‘인간 고유성’과 연관된다고 볼 수 있다(Gray & Wegner, 2012).

마음지각에서 경험성의 차원과 주도성의 차원은 도덕성에 중요한 시사점을 제시한다. Gray 등

(2007)에서는 특정 대상이 경험성의 차원에서 점수가 높을수록 피행위자로서 인식되는 경향이 컸던데 비해, 주도성의 차원에서 점수가 높을수록 행위주체로서 지각될 가능성이 큰 것으로 나타났다. 또한 이 연구에서는 다양한 대상을 2개씩 짝지어 제시하면서, 이 중에서 도덕적 책임을 수행해야 할 것으로 보이는 대상과 가능하면 연구참가자가 해를 끼치고 싶지 않은 대상을 선택해보도록 했다. 그 결과 경험성의 차원에서 점수가 높은 대상일수록 피해를 입히고 싶지 않고, 보호받아야 할 가치가 있는 대상으로 선택되었으며, 주도성의 차원에서 점수가 높은 대상일수록 도덕적 책임을 져야 할 대상으로 선택되었다. 이 결과는 우리가 왜 특정 대상에게 피해를 입히는 것이 더 어렵고, 왜 특정 대상에게 도덕적 책임을 더 묻게 되는지를 설명해준다. 예를 들어, 성인 남자와 3살 영아가 모두 강아지를 해쳤을 때, 왜 성인남자에게 더 책임을 전가하게 되는지를 설명해준다. 또한 경험성의 차원에서 점수가 낮은 로봇에게 해를 가하는 것은, 경험성의 차원에서 점수가 높은 5살 소녀를 해치는 것보다 더 수월하다. 왜냐하면 마음지각의 경험성의 차원에서 점수가 낮은 대상일수록, 피해를 받을 때 고통을 적게 느끼는 것으로 인식되기 때문이다. 고통을 느끼는 능력이 적으면, 보호받아야 할 가치가 있는 대상으로 인식되지 않아서, 해치는 행동이 나쁜 행동으로 인식되지 않는다. 주도성과 경험성의 차원에서 모두 점수가 낮은 대상은 도덕성을 고려해야 하는 범위 밖으로 밀려나게 되며, 이 대상은 행위자 또는 피행위자로서의 어떤 도덕적 지위도 갖지 못하게 된다.

2) Gray 등(2007)은 요인분석을 수행하여, 마음과 연관된 열 여덟 개의 속성중에서 직각회전법(varimax rotation)에 의해 두 가지 독립된 요인인 주도성과 경험성을 추출하였다.

3) 인간본성은 주로 인간과 기계와의 구분에 사용되었다. 기계의 특성으로 가장 흔히 인식되는 정서결여, 경직성, 수동성 및 자동성은 비인간화되는 인간에게 자주 사용되었다(Montague & Matson, 1987).

2) 로봇에 대한 마음지각과 도덕적 지위의 부여

지금까지 로봇은 인간의 욕구를 충족시켜주는 수단으로만 인식되고, 인간과 동일한 위치에서 수평적으로 존엄하게 다루어야 할 대상으로 지각되지 않았다(Friedman et al., 2003). 로봇에 대한 마음지각에서 로봇이 인간과 특히 다른 것은 마음지각에서 경험성의 측면이다. 로봇이 문제를 해결하는 지적 능력은 인간보다 더 우월할 때도 있지만, 고통을 느끼거나 정서를 경험하는 능력은 부족하게 여겨지기 때문이다. 또한 Gray 등(2007)에서 보고한 마음지각과 도덕성 지위의 관계를 토대로 생각해볼 때, 로봇은 경험성의 측면에서 열등하다고 지각되기 때문에 도덕적으로 처우할 필요가 없다고 인식될 것을 예상할 수 있다.

그러나 최근 15년간 로봇에 대한 태도는 인간의 사회적 욕구에 대한 정서적 지지 기능이 강화되면서 변화되고 있다(Wang & Krumhuber, 2018). 지금까지 로봇의 기능이 경제적 측면에서 생산과정의 효율성을 극대화하는데 중점을 두었다면, 사회적 로봇이 점점 보편화될수록 로봇의 정서적 지원기능이 더 잘 인식될 것을 예상할 수 있다. 사회적 로봇은 의료서비스나 사회적 돌봄(Graf et al., 2009; Pollack et al., 2002) 또는 동반자의 기능을 수행하도록 설계되어 있다. 사회적 로봇의 정서적 지원 기능은 인간과 친밀한 관계를 맺도록 하기 때문에, 로봇의 사회적 가치가 증가함과 동시에 로봇에 대한 마음지각이 경험성의 측면에서 상승할 가능성이 있다.

우리가 상대방의 사회적 기능을 관찰하면서 그 사람의 특성을 추론한다는 사회적 역할이론(Hoffman & Hurst, 1990; Wood & Eagly, 2012)에서는 우리가 로봇에 대해서도 로봇이 어떤 사회적 기능을 수행하는지에 따라 로봇의 특성을 추론하게

될 가능성을 제시한다. Wood와 Eagly(2012)는 성 고정관념의 형성에 대한 설명에서 수렵채집사회의 전통적 역할분담에서 여성이 돌봄의 역할을 수행하였기 때문에 ‘따뜻함’의 용어가 남성보다 더 잘 연결되며, 남성은 사냥과 같은 역할을 더 많이 수행하였기 때문에 ‘주도적’의 특성이 더 많이 연상된다고 설명하였다. 또한 직업에 대한 고정관념의 한 예시로서 사업가의 범주에서는 예술가보다 ‘이성적’의 특성이 더 잘 연상되며, 예술가의 범주에서는 ‘감정적’의 특성이 사업가의 범주에서보다 더 잘 연상됨을 보고하였다. 따라서 특정 범주가 주로 어떤 기능을 수행하는 것으로 사전에 정의되었는지에 따라 이러한 유형의 범주와 연상되는 속성이 달라질 수 있다. 로봇은 전통적으로 생산과정을 효율화하고, 경제적 이득을 극대화시키는데 활용되었기 때문에 마음지각의 주도성과 연관된 지적 측면(예: 기억능력)에서는 대체로 높은 점수를 받을 것을 예측할 수 있다. 반면 로봇의 사회적 지지 기능은 최근에 와서야 추가된 새로운 기능적 가치로서, 지금까지 거의 개발되지 않았기 때문에 마음지각의 경험성 측면(예: 감정을 느끼는 능력)에서 점수가 낮을 것을 예측할 수 있다. 그러나 최근 들어 로봇의 사회적 기능이 부각되면서, 로봇에 대한 마음지각이 경험성의 측면에서 상승하고, 이에 따라 더 높은 도덕적 지위가 부여될 것을 예측할 수 있다. 따라서 로봇의 어떤 기능적 가치를 활성화시키는지에 따라 마음지각과 로봇에 대한 가해행동의 비도덕성 판단이 달라질 가능성을 생각해볼 수 있다.

요약하면, 본 연구에서는 대상에 대한 마음지각과 가해 행동의 비도덕성 판단의 관계에 대해 다음과 같이 두 개의 연구를 통해 검증하였다. 첫째, 연구 1에서는 다양한 대상에 대한 마음지각에서 로봇이 다른 대상과 비교할 때 어떻게 차이가 나타나는지를 검증하였다. 선행연구(예: Gray 등,

2007)에서는 마음지각이 다양한 대상에 따라 달라짐을 보고하였다. Gray 등의 연구에서는 마음지각의 차이에만 초점을 맞추었기 때문에, 본 연구에서는 마음지각의 두 차원을 구분하여 로봇과 다양한 대상에 대한 마음지각이 피행위자에 대한 도덕적 지위의 부여와 유의한 관계에 있는지를 검증하였다. 본 연구에서는 대상에 대한 마음지각의 경험성이 높을수록 해를 끼치는 행동에 대한 심리적 불편감을 더 많이 갖게 되고, 옳지 않은 행동으로 판단하는 경향이 더 높을 것을 예측하였다. 또한 로봇에 대한 마음지각이 다른 대상에 비해 상대적으로 낮고, 가해행동에 대한 심리적 불편감과 비도덕성 평정도 더 낮을 것을 예측하였다. 둘째, 연구 2에서는 로봇의 기능적 가치를 제시하는 방식에 따라 로봇에 대한 마음지각이 달라지는지를 분석하였다. 지금까지 선행연구(Gray & Wegner, 2012)에서는 인간에 대한 마음지각을 중심으로 마음지각연구가 주로 수행되었고, 로봇에 대한 마음지각이 거의 연구되지 않았다. 또한 로봇에 관한 선행연구(Belpaeme et al., 2018)에서는 로봇의 외적인 측면(예: 외형, 성별, 목소리)을 중심으로 주로 연구되었고, 로봇의 내적 측면에 대한 경험적 연구는 거의 수행되지 않았다. 경험적 자료를 보고하지는 않았지만, 로봇의 의인화에 관한 선행연구(Damiano & Dumouchi, 2018; Duffy, 2003)에서는 사회적 로봇이 경제적 로봇보다 내적 속성이 인간과 더 유사하게 지각되어, 로봇에게 인간의 속성(예: 자율적인, 공감하는)이 부여되는 의인화가 나타날 가능성을 제시하였다. 로봇의 내적 측면은 향후 인간-로봇의 상호작용에 끼칠 영향이 크기 때문에 본 연구에서는 로봇의 기능을 제시하는 방식에 따라 로봇의 내적 상태를 추론하는 마음지각과 로봇을 해치는 행위의 비도덕성 판단이 달라지는지를 검증하였다. 본 연구에서는 로봇의 기능적 가치를 사회적 로봇으

로 제시하는 조건에서 경제적 로봇으로 제시하는 조건보다 로봇에 대한 마음지각이 더 높고, 로봇에게 더 높은 도덕적 지위가 부여될 것을 예측하였다.

3. 연구 1

1) 방법

(1) 연구설계

연구 1에서는 열 세 개의 다양한 대상에 대한 마음지각 및 도덕성 판단을 참가자내 설계로 측정하였다.

(2) 참가자

4년제 대학교의 대학생 130명(평균연령 만 23.16세, 표준편차=4.02세; 남=67명)이 학교 홈페이지 게시판을 통해 참가하였다. 연구 참가에는 약 10분의 시간이 소요되었으며, 참가자들은 오천원 상당의 커피쿠폰 보상을 받았다.

(3) 도구 및 절차

본 연구는 연구참가자에게 사전동의를 얻은 후, 온라인 설문지로 진행되었다. 참가자에게는 총 열 세 개의 다양한 대상에 대한 마음지각 질문지([부록 1] 참조)가 제시되었다. 참가자에게는 개, 소, 돼지, 성인 남자, 성인 여자, 84세 노인, 5살 아동, 장애인, 노숙자, 식물상태 환자, 4개월 된 태아, 신, 로봇에 대한 마음지각을 평정하도록 지시하였다. 참가자는 각 대상을 마음지각 질문지(Gray 등, 2007)를 기반으로 7점 척도에 따라 평정하였

다. 마음지각 질문지는 2개의 하위영역으로 구성되어 있다. 첫 번째 하위영역은 정서와 연관된 경험성(experience)(예: 감정, 즐거움)인데 비해, 두 번째 하위영역은 지적 능력과 연관된 주도성(agency)(예: 계획, 자기조절)이다. 즉 특정 대상이 마음을 갖고 있는지를 평가하는 것은 그 대상이 정서적 경험을 할 능력이 있는지와 어느 정도 지적으로 주도적으로 활동을 할 수 있는 지로 측정된다. 이 영문판 질문지(Gray 등, 2007⁴⁾)는 본 연구의 연구자에 의해 한국어로 번역된 후, 전문 번역인이 다시 교차번역하는 문항검수절차를 거쳤다. 연구 1에서 마음지각 질문지의 내적 합치도(Cronbach's Alpha)는 경험성의 측면에서 .82, 주도성의 측면에서 .91로 나타났다. 이후 참가자들은 열 세 개의 다양한 대상에 대한 도덕성 과제를 수행하였다. 이 대상은 마음지각과제에 사용된 열 세 개의 대상과 동일하였다. 도덕성과제에는 Bastian 등(2012)을 토대로 2개의 문항을 사용하였다. 이 중 첫 번째 문항에서는 이 대상을 해친다면 마음이 불편한 정도를 7점 척도에 따라 평정하고(1: 전혀 일치하지 않음, 7: 매우 일치함), 두 번째 문항에서는 이 대상을 해치는 행동이 도덕적으로 옳지 못하다고 생각하는 정도를 참가자들에게 7점 척도에 따라 응답하도록 했다(1: 전혀 나쁘지 않다, 7: 매우 나쁘다). 도덕성 과제에서 문항 간의 내적 합치도(Cronbach's Alpha)는 .68로 나타났다. 도덕성과제가 완료되면, 참가자들에게 연구목적과 취지를 파악했는지에 대한 질문을 한 후, 참가에 대한 보상으로 오천 원 상당의 커피쿠폰을 온라인으로 지급하였다.

(4) 분석방법

연구 1에서는 다양한 대상에 따라 마음지각 및 도덕성 판단이 달라지는지를 검증하고자 하였다. 이를 위해 SPSS 18.0 통계분석프로그램을 사용하여 기술통계, 상관분석 및 반복측정 변량분석을 수행하였다.

2) 연구 1의 결과 및 논의

(1) 다양한 대상에 대한 마음지각

<표 1>과 같이 다양한 대상에 대한 경험성과 주도성의 마음지각, 해치는 행동에 대한 불편한 마음 및 해치는 행동에 대한 도덕성 평정의 기술통계를 수행하였다. 선행연구(Gray 등, 2007; Bastian et al., 2012)를 참조하여 각 대상에 따라 마음지각을 분석한 결과, 다양한 대상에 대한 마음지각의 경험성과 주도성에서 유의한 차이가 나타났다. 열 세 개의 대상에 대한 마음지각을 반복측정 변량분석으로 분석한 결과 대상의 주효과가 유의하였다, $F(12, 1546) = 11.23, p < .001, \eta^2 = .41$. 마음지각은 성인남자에서 가장 높았고, 로봇에서 가장 낮게 나타났다. 또한 동일한 범주안에서도 마음지각의 차이가 나타났다. 인간의 범주안에서 성인남자에 대한 마음지각이 가장 높았고, 식물상 태환자에 대한 마음지각이 가장 낮게 나타났다. 동물의 범주에서는 개에 대한 마음지각이 가장 높았고, 돼지에 대한 마음지각이 가장 낮게 나타났다.

4) Gray 등(2007)의 연구에서는 경험성과 연관된 문항 11개, 주도성과 연관된 문항 7개를 사용하였다. 본 연구에서는 경험성과 연관된 문항중에서 Gray 등(2007)의 요인분석결과를 토대로 요인부하량이 높은 일곱 개 문항을 선택하여 사용하였다.

<표 1> 다양한 대상에 대한 마음지각과 해치는 행동에 대한 비도덕성 판단(연구 1)

대상	마음지각의 주도성	마음지각의 경험성	마음지각 평균	심리적 불편함	해치는 행동의 비도덕성
성인남자	6.24 (1.09)	5.98 (1.01)	6.14 (1.21)	6.81 (1.30)	6.89 (1.10)
성인여자	5.89 (1.12)	6.07 (1.29)	6.03 (1.27)	6.63 (1.54)	6.72 (1.36)
5살 아동	5.56 (1.32)	6.09 (1.23)	5.89 (1.42)	6.52 (1.39)	6.65 (1.18)
신	5.89 (1.32)	5.16 (1.08)	5.48 (1.24)	5.26 (1.83)	5.50 (1.64)
장애인	5.47 (1.12)	5.98 (1.32)	5.71 (1.27)	6.24 (2.21)	6.12 (1.80)
84세 노인	4.95 (1.01)	4.72 (1.13)	4.83 (1.07)	6.54 (1.23)	6.37 (1.58)
개	4.47 (1.23)	4.82 (1.07)	4.68 (1.17)	6.33 (2.18)	5.83 (1.07)
노숙자	4.43 (1.24)	4.70 (1.54)	4.58 (1.42)	5.92 (1.72)	6.13 (1.22)
소	4.16 (1.02)	4.32 (1.23)	4.24 (1.15)	5.18 (1.44)	4.83 (2.16)
돼지	3.89 (1.12)	4.01 (1.20)	3.95 (1.17)	4.91 (1.41)	4.33 (1.75)
4개월 태아	3.24 (2.12)	3.78 (1.94)	3.51 (2.01)	5.28 (1.45)	5.10 (1.89)
식물상태환자	3.23 (1.01)	3.48 (1.39)	3.36 (1.20)	5.89 (2.28)	5.72 (1.77)
로봇	2.51 (1.34)	2.02 (1.21)	2.28 (1.29)	3.83 (1.15)	3.86 (1.66)

(2) 마음지각과 해치는 행동의 비도덕성 평정

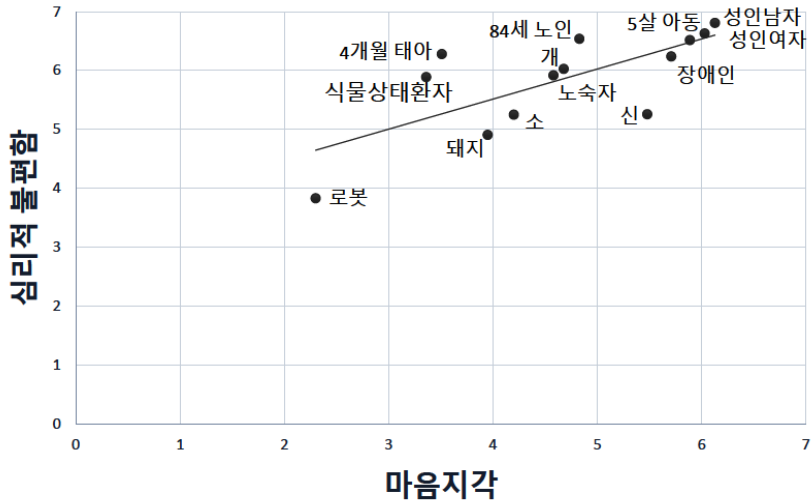
다양한 대상에 대한 마음지각과 해치는 행동에 대한 도덕성 평정의 관계는 전반적으로 유의하게 나타났다([부록 2] 참조). 성인남자에 대한 마음지각이 가장 높았고, 해치는 행동에 대한 심리적 불편감이 가장 높았다. 이에 비해 로봇에 대한 마음지각이 가장 낮았고, 해치는 행동에 대한 심리적 불편감과 비도덕성에 대한 평정도 가장 낮았다.

또한 <그림 1>과 같이 다양한 대상에 대한 마음지각의 평균 점수와 이 대상을 해치는 행동에 대한 불편한 마음의 점수 간에 평균 $r = .38$ 의 정적 상관이 나타났다. 예를 들어, 식물상태 환자에 대한 마음지각과 식물상태 환자를 해친다면 마음이 불편한 것이라고 평정한 점수 간에는 유의한 정적 상관관계가 나타났다, $r = .41, p = .002$. 또한 다양한 대상에 대한 마음지각의 평균 점수와 이 대상을 해치는 행동이 도덕적으로 옳지 못하다고 평정한 점수 간에도 평균 $r = .28$ 의 정적 상관이 나타났다. 예를 들어, 5세 아동에 대한 마음

지각이 높을수록 이 아동을 해치는 행동이 도덕적으로 옳지 못하다고 평정한 점수가 더 커지는 경향이 나타났다, $r = .37, p = .002$. 또한 로봇에 대한 마음지각이 낮을수록 로봇을 해치는 행동이 옳지 못하다고 평정하는 경향이 낮게 나타났다, $r = .31, p = .004$.

(3) 논의

연구 1에서는 예상했던 대로 로봇이 다른 대상과 비교할 때, 마음지각과 가해 행동에 대한 비도덕성 판단에서 차이가 나타났다. 성인남자에 대한 마음지각이 가장 높았고, 로봇에 대한 마음지각이 가장 낮게 나타났다. 또한 마음지각과 가해행동의 비도덕성 판단 간에 정적 상관관계가 나타나서, 마음지각이 가장 높게 나타났던 성인남자로부터 각 대상에 대한 마음지각이 어느 정도 떨어져 있는지에 따라 피행위자에 대한 도덕적 처우가 달라짐을 보여주었다. 로봇에 대한 마음지각은 성인남자에 대한 마음지각과 가장 멀리 떨어져 있었



<그림 1> 연구 1의 마음지각과 해치는 행동의 심리적 불편함의 관계

으며, 로봇을 해치는 행동에 대해 심리적 불편감이 가장 적었고, 해치는 행동을 옳지 못하다고 평정하는 경향이 가장 낮게 나타났다. 이 결과는 연구 참가자가 열 세 개의 대상 중에서 로봇과의 관계에서 마음의 경험성과 주도성에 대한 유사성을 가장 적게 경험하였고, 도덕적 처우의 필요성을 거의 느끼지 않았음을 보여준다. 연구 2에서는 연구 1의 마음지각에서 점수가 가장 낮았던 로봇의 기능적 가치를 어떻게 제시하는지에 따라 로봇에 대한 마음지각과 로봇을 해치는 행동에 대한 도덕성 평정이 달라질 수 있는지를 검증하였다.

4. 연구 2

1) 방법

(1) 연구설계

연구 2에서는 로봇의 제시방식에 따라 마음지각이 달라지는지를 검증하였다. 이에 따라 연구 2는

2(제시방식: 경제적 로봇 vs. 사회적 로봇)*2(마음지각: 경험성 vs. 주도성)의 혼합설계로 구성되었다. 참가자간 변인은 로봇의 기능적 가치의 제시방식이었고, 참가자내 변인은 마음지각이었다. 또한 로봇의 제시방식(경제적 로봇 vs. 사회적 로봇)에 따라 로봇에 대한 가해행동의 도덕성 판단이 달라지는지를 검증하였다. 참가자간 변인은 로봇을 제시방식이었고, 측정변인은 도덕성 판단이었다.

(2) 참가자

4년제 대학교의 대학생 114명(평균연령 만 23.14세, 표준편차 = 2.26세; 남 = 54명)이 교양수업을 수강하면서 수업계시판의 모집공고를 통해 자발적으로 참가하였다. 연구참가에는 약 5분의 시간이 소요되었으며, 참가자들은 연구참가에 대한 보상으로 오천원 상당의 기프티콘을 받았다.

(3) 도구 및 절차

연구 2에서는 참가자 집단을 무선적으로 두 집

단으로 나누어, 로봇의 기능에 대한 각각 다른 글을 제시하였다. 경제적 로봇의 집단에서 참가자는 로봇의 실용적 측면을 기술하는 글을 읽었으며, 사회적 로봇의 집단에서 참가자는 로봇이 인간의 삶을 위해 정서적 지지를 수행하는 글을 읽었다 ([부록 2] 참조). 이후 참가자들은 연구 1과 동일하게 로봇에 대한 마음지각 질문지에 응답하였다. 연구 2에서 마음지각 질문지의 신뢰도(Cronbach's alpha)는 경험성의 측면에서 .78, 주도성의 측면에서는 .84로 나타났다. 마음지각 질문에 응답을 완료하면, 참가자는 연구 1과 동일한 도덕성 문제를 수행하였다. 참가자는 경제적 로봇 또는 사회적 로봇에게 해를 끼치게 될 경우 마음이 불편한 정도와 이러한 행동을 옳지 못하다고 생각하는 정도를 7점 척도에 따라 응답하였다. 연구 2에서 도덕성 문제의 신뢰도(Cronbach's alpha)는 .71로 나타났다. 마지막으로 참가자들은 연구의 목적과 취지를 파악했는지에 대한 질문에 응답한 후, 기프트콘을 온라인으로 제공받았다.

(4) 분석방법

연구 2에서는 경제적 로봇 또는 사회적 로봇에 대한 마음지각과 도덕성 평정에 관한 연구참가자의 자료를 자기보고식 질문지를 통해 수집하였다. 연구 2의 자료분석을 위해 SPSS 18.0 프로그램을 사용하여 기술통계, 상관분석, t-검정 및 혼합변량 분석을 수행하였다.

2) 결과

(1) 로봇의 기능적 가치제시방식과 마음지각

연구 2에서는 로봇의 기능적 가치 제시방식에 따른 마음지각과 로봇을 해치는 행동의 비도덕성

에 대한 평균분석을 수행하였다(<표 2> 참조). 마음지각에서는 경험성과 주도성에 따라 각 차원에 속하는 일곱 개 문항에 대한 평균값을 각각 산출하여 결과분석을 수행하였다. 로봇의 기능적 가치를 제시하는 방식이 마음지각에 영향을 끼쳤는지를 분석하기 위해 2(제시방식: 경제적 로봇 vs. 사회적 로봇) x 2(마음지각: 경험성 vs. 주도성)의 혼합변량분석을 수행하였다. 제시방식은 참가자간 변인이었고, 마음지각의 두 가지 차원은 참가자내 변인이었다.

그 결과, 제시방식의 주효과가 유의하였다, $F(1, 112) = 11.23, p = .001, \eta^2 = .091$. 독립표본 t-검정을 수행한 결과, 사회적 로봇에 대한 마음지각($M = 3.75, SD = 1.15$)은 경제적 로봇에 대한 마음지각($M = 3.08, SD = 0.94$)보다 전반적으로 더 높았다, $t(112) = 3.35, p = .001$. 또한 마음지각의 주효과가 유의하였다, $F(1, 112) = 387.42, p < .001, \eta^2 = 77.60$. 대응표본 t-검정을 수행한 결과, 로봇에 대한 마음지각은 주도성($M = 4.80, SD = 1.41$)에서 경험성의 차원($M = 2.02, SD = 1.25$)보다 전반적으로 더 높게 나타났다, $t(113) = 19.70, p < .001$. 이 결과는 로봇에 대한 마음지각이 지적능력에 더 집중되어 있음을 보여준다. 제시방식과 마음지각의 상호작용은 유의하지 않았다, $F < 1$.

주목할 것은 <표 2>와 같이 사회적 로봇이 경제적 로봇보다 경험성의 차원에서도 마음지각이 더 높았고, $t(112) = 3.57, p = .001$, 주도성의 차원에서도 더 높았던 것이다, $t(112) = 2.01, p = .045$. 이 결과는 사회적 로봇이 정서를 경험하는 능력에서뿐만 아니라 지적 능력에서도 더 뛰어난 로봇으로 인식되고 있음을 보여준다.

(2) 로봇에 대한 마음지각과 도덕성 평정

로봇의 기능적 가치를 제시하는 방식이 해치는

<표 2> 로봇에 대한 마음지각과 해치는 행동에 대한 비도덕성 평정의 관계(연구 2)

변인		평균	표준편차	1	2	3
경제적 로봇	1. 경험성	1.62	.91			
	2. 주도성	4.55	1.41	.29*		
	3. 해치는 행동의 심리적 불편감	3.91	2.05	.31*	.23	
	4. 해치는 행동의 비도덕성	3.17	1.88	.26*	.09	.743**
사회적 로봇	1. 경험성	2.42	1.41			
	2. 주도성	5.07	1.37	.36**		
	3. 해치는 행동의 심리적 불편감	5.01	1.82	.44**	.50**	
	4. 해치는 행동의 비도덕성	3.82	1.87	.54**	.42**	.68**

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

행동에 대한 비도덕성 평정에 영향을 끼쳤는지를 분석하기 위해 독립표본 t-검정을 수행하였다. 그 결과 <표 2>와 같이 사회적 로봇을 해치는 행동에 대한 심리적 불편감이 경제적 로봇을 해치는 행동에 대한 심리적 불편감보다 더 높았다, $t(112) = 3.35, p = .001$.

또한 로봇에 대한 마음지각과 해치는 행동에 대한 심리적 불편감 및 비도덕성 평정에 대한 상관 분석을 수행하였다. <표 2>와 같이 경제적 로봇에 대한 마음지각이 경험성에서 높을수록 해치는 행동에 대한 심리적 불편감이 더 높았다, $r = .31, p = .016$. 또한 경제적 로봇에 대한 마음지각이 경험성에서 높을수록 해치는 행동이 옳지 못하다고 평정하는 경향이 더 높았다, $r = .26, p = .044$. 반면, 경제적 로봇에 대한 마음지각이 지적 능력과 연관된 주도성에서 높은 것은 심리적 불편감이나 비도덕성 판단과 유의한 관계가 나타나지 않았다. 따라서 경제적 로봇과 도덕성 평정의 관계에서는 경제적 로봇에 대한 경험성 지각이 더 중요한 역할을 수행함을 추정해볼 수 있다.

이에 비해 사회적 로봇에서는 마음지각의 경험성과 주도성의 차원에서 모두 해치는 행동의 비도덕성 평정과 유의한 관계가 나타났다. 사회적

로봇에 대한 마음지각이 경험성에서 높을수록 가해행동에 대한 심리적 불편감이 더 높았다, $r = .44, p = .001$. 또한 사회적 로봇에 대한 마음지각이 경험성에서 높을수록 가해행동이 옳지 못하다고 평정하는 경향이 더 높았다, $r = .54, p = .001$. 또한 사회적 로봇은 지적 능력과 연관된 주도성에서도 심리적 불편감이나, $r = .50, p = .001$, 비도덕성 판단과 유의한 관계가 나타났다, $r = .42, p = .001$. 이 결과는 사회적 로봇의 주도성이 높게 지각될수록 로봇에 대한 가해행동의 비도덕성 평정도 커질 가능성을 보여준다.

마지막으로 로봇에 기능에 따른 마음지각이 로봇에 대한 가해행동의 비도덕성 평정에 조절효과를 갖는지를 분석하였다. <표 3>과 같이 로봇의 기능을 독립변인(연)으로, 마음지각을 조절변인으로, 비도덕성 평정을 종속변인으로 설정하였다. 위계적 회귀분석의 결과, 마음지각의 경험성 차원은 조절변인으로서 유의한 영향을 끼치지 않았다. 반면 로봇의 기능/마음지각 주도성의 상호작용항이 유의하여 주도성 차원은 조절변인으로 유의하였다. 따라서 로봇의 기능에 따른 주도성의 지각은 로봇에 대한 가해행동의 비도덕성의 평정에 차별화된 영향을 끼치는 것으로 보인다.

5) 로봇의 기능을 더미변수로 변환하여 분석하였음(경제적 로봇=0, 사회적 로봇=1).

<표 3> 로봇의 기능에 따른 마음지각의 주도성이 로봇을 해치는 행동에 대한 비도덕성 평정의 조절효과 분석(연구 2)

	독립변인	β	t	R^2	ΔR^2	F
1단계	(상수)		2.89	.16	.14	10.65***
	로봇의 기능	.18	2.64*			
	마음지각의 주도성	.32	3.67***			
2단계	(상수)		3.35	.18	.16	8.25***
	로봇의 기능	.34	1.10			
	마음지각의 주도성	.17	1.47			
	로봇의 기능 x 마음지각의 주도성	.65	2.24*			

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

(3) 논의

연구 2에서는 로봇의 기능에 따라 마음지각과 로봇을 해치는 행동에 대한 비도덕성 판단의 관계를 분석하였다. 그 결과 로봇에 대한 마음지각에서 경험성의 차원은 로봇의 기능과 관계없이 로봇을 해치는 행동에 전반적으로 유의한 영향을 끼치는 것으로 나타났다. 이것은 선행연구(Wang & Krumhuber, 2018)에서 보고하는 것과 같이 특정 대상이 정서를 느낄 수 있는 것으로서 인식될 때, 해치는 행동에 대한 심리적 불편감이 증가하는 것으로 해석할 수 있다. 이에 비해 마음지각의 주도성은 로봇의 기능과 상호작용하면서 비도덕성 평정에 조절변인으로서 영향을 끼치는 것으로 보인다. 이것은 본 연구에서 도출된 새로운 결과로 볼 수 있다. 지금까지 선행연구에서는 마음지각의 경험성과 피행위자에 대한 가해행동의 도덕성 평정의 관계가 주로 보고되는데 비해, 본 연구 결과는 마음지각의 주도성 역시 가해행동의 비도덕성 평정에 영향을 끼칠 가능성을 보여주었다.

로봇의 기능을 구분하여 제시했을 때, 사회적 로봇에 대한 마음지각은 경제적 로봇에 대한 마음지각보다 더 높게 나타났다. 이 결과는 경제적 로봇보다는 사회적 로봇이 사회적 상호작용에서

더 수월하기 때문에 인간의 마음과 더 비슷하게 지각되며, 더 긍정적 관계가 성립될 가능성을 시사한다. 또한 경제적 로봇에서는 마음지각의 경험성 차원과 해치는 행동의 관계만 유의했던데 비해, 사회적 로봇에서는 마음지각의 경험성 차원뿐만 아니라 주도성 차원에서도 해치는 행동의 비도덕성 판단과 유의한 관계가 나타났다. 마음지각의 주도성은 자기조절능력과 상황에 대한 영향력과 연관된다. 따라서 연구참가자들이 사회적 로봇을 행위를 하는 주체로서 인식할수록 로봇을 해치는 행동을 비윤리적인 것으로 판단한 경향이 커진 것으로 해석할 수 있다. 이 결과는 사회적 로봇은 경제적 로봇보다 피행위자로서 뿐만 아니라 주체적 행위자로서도 공감과 같은 도덕적 정서를 더 많이 유발할 가능성을 추정해볼 수 있다.

연구 2의 결과를 다음과 같이 두 가지로 요약해볼 수 있다. 첫째, 로봇에 대한 마음지각은 인간에 대한 마음지각과 유사하게 대상의 기능에 의해 영향을 받고 있음을 알 수 있다. 인간의 마음지각과 비슷하게 로봇에 대한 마음지각에서도 경험성과 주도성의 차원이 분리되어 작동하고, 이 두 가지 차원에서 사회적 로봇은 경제적 로봇보다 인간의 마음지각과 더 유사하며, 경제적 로봇보다 인간의 특성에 더 근접한 것으로서 인식되고 있

음을 추정해볼 수 있다. 둘째, 사회적 로봇을 해치는 행동이 경제적 로봇을 해치는 행동보다 도덕적으로 더 옳지 않다고 생각하는 것은 사회적 로봇이 경제적 로봇보다 인간과 더 비슷하게 인식되고 있음을 알 수 있다. 따라서 향후 인간-로봇의 상호작용에서 사회적 로봇이 경제적 로봇보다 더 높은 도덕적 지위를 갖게 되고, 로봇과의 연결 욕구에도 더 큰 영향력을 끼칠 가능성이 있다. 그러나 본 연구에서는 연구참가자의 다양한 개인차 변인(예: 개인의 관계욕구, 인공지능기술에 대한 사전 태도)을 수집하지 않았기 때문에 로봇-인간의 관계에 대한 문제에 명료하게 답하기에는 한계가 있다.

5. 종합논의

본 연구에서는 로봇과 다양한 대상에 대한 마음지각의 차이를 비교하고, 마음지각과 가해 행동에 대한 비도덕성 판단의 관계를 분석하였다. 그 결과 성인 남자에 대한 마음지각이 가장 높게 나타났고, 로봇에 대한 마음지각이 가장 낮게 나타났다. 또한 대상에 대한 마음지각이 높을수록 그 대상을 해치는 행동에 대한 심리적 불편감 및 그 행동을 옳지 못하다고 판단하는 정도가 높게 나타났다(연구 1). 연구 2에서는 연구 1에서 마음지각이 가장 낮게 나타난 로봇에 대해 경제적 이득을 지원하는 경제적 로봇 또는 정서적 기능을 지원하는 사회적 로봇으로 차별화하여 제시하였을 때, 로봇의 기능적 가치에 따라 마음지각이 달라지는지를 검증하였다. 그 결과 마음지각에서 상대적으로 가장 낮았던 로봇도 사회적 로봇으로 제시하게 될 경우에는 경험성과 주도성의 차원에서 마음지각의 점수가 모두 상승하고, 해치는 행동에 대한 비도덕성 평정도 증가하는 것이 관찰되었다.

본 연구는 다음의 세 가지 중요한 시사점을 제시한다. 첫째, 마음지각의 과정은 직관적이고 주관적인 과정임을 알 수 있다. 우리는 대상의 마음을 알지 못하기 때문에 대상과의 관계를 통해 대상의 마음을 추론하고, 이 과정에는 대상과의 관계에 영향을 끼치는 기능적 가치의 변인(예: 정서적 지원)이 중요한 역할을 할 가능성을 시사한다. 둘째, 로봇에 대한 마음지각과 도덕적 지위의 관계에서는 로봇의 기능을 어떻게 제시하는지에 따라 마음지각의 경험성과 주도성의 측면이 영향을 끼치는 정도가 달라질 가능성을 보여준다. 선행연구(Wang & Krumhuber, 2018)의 결과와 같이 경험성의 측면은 로봇을 해치는 행위의 비도덕성 판단에 중요한 영향을 끼쳤다. 이에 비해 사회적 로봇의 경우에는 로봇의 지적 능력과 연관된 주도성이 높게 지각될수록 로봇에 대한 가해행위의 비도덕성 판단이 증가하였다. 이것은 사회적 로봇이 행위를 하는 주체로서도 인간에게 도덕적 정서를 더 유발시킬 가능성을 보여준다. 셋째, 로봇의 기능적 가치를 다르게 제시함에 따라 마음지각이 변화된 것을 통해 인간에 대한 마음지각이나 로봇에 대한 마음지각이 유사한 추론과정을 기반으로 작동할 것을 추정해볼 수 있다. 따라서 특정 대상의 사회적 기능에 대해 사전에 정의내린 것을 기반으로 대상의 속성과 마음상태가 결정된다는 사회적 역할의 이론이 인간뿐만 아니라 로봇에게도 적용가능함을 알 수 있다.

본 연구의 제한점 및 후속연구에 대한 제안은 다음과 같다. 첫째, 연구 2에서는 사회적 로봇에 대한 마음지각이 경제적 로봇에 대한 마음지각보다 유의하게 높았지만, 연구 1의 인간이나 동물에 대한 마음지각보다 상대적으로 낮았다. 또한 연구 2에서 사회적 로봇의 세분화된 유형에 대해 마음지각의 차이를 비교하지 않은 것은 연구결과의 한계로서 남는다. 로봇에 대한 마음지각은 다양한

현장에서 사회적 로봇을 활용할 때, 인간과 로봇의 상호작용에 직접적 영향을 끼칠 수 있기 때문에, 후속연구에서는 사회적 로봇의 다양한 유형을 활용하여 마음지각의 차이를 비교하는 것이 필요할 것으로 보인다. 예를 들어, Belpaeme, Kennedy와 Ramachandran(2018)에서는 사회적 로봇을 교육현장에서 활용할 때, 인간과 정서적으로 상호작용하는 튜터 로봇으로서 학습자와 상호작용할 때 지식전달뿐만 아니라 정서경험도 가능하게 하기 때문에 교육효과가 더 증가함을 보고하였다. 따라서 로봇을 인간과 유사하며, 실제 중요한 역할을 수행하는 존재로서 가시화하는 구체적 방안에 대한 후속연구가 필요할 것으로 보인다.

둘째, 연구 1에서 다양한 대상에 대한 마음지각의 차이가 나타났지만, 이 차이는 자기보고식 질문지를 통해 측정하였기 때문에 마음지각의 추론과정을 객관적으로 정교하게 분석하지 못했다. 최근 연구에서는 마우스추적 컴퓨터 프로그램 기법을 활용하여 참가자가 스스로 의식하지 못하는 자동적 추론 과정을 측정하고 있는데, 후속연구에서는 이러한 기법을 토대로 대상에 대한 마음지각의 경험성과 주도성을 추론하는 과정을 분석해볼 필요가 있다(Scherbaum & Dshemuchadse, 2020).

셋째, 본 연구에서는 대학생들을 대상으로 온라인 질문지를 실시하여 자료를 수집하였다. 본 연

구에서는 연구참가자의 다양한 개인차 변인(예: 과학기술에 대한 태도, 사회적 바람직성)을 사전에 측정하여 통제하지 못하였고, 참가자에 대한 인구통계학적 정보를 충분히 수집하지 못하여 연구결과를 해석하는데 한계가 있다. 또한 온라인 설문기법은 연구참가자의 불성실응답으로 인해 연구결과의 질을 떨어뜨릴 위험이 있다. 후속 연구에서는 불성실응답의 사전 방지기법(예: 김우영, 이태현, 장재운, 2019)을 활용하여 온라인 설문기법의 문제점을 감소시킬 필요가 있다.

4차 산업혁명이후 디지털화된 사회에서 인간-로봇의 상호작용에 대한 논의가 활발히 진행되고 있다. 따뜻한 정서는 인간-인간의 관계에서 뿐만 아니라 인간-로봇의 상호작용에도 중요한 역할을 할 속성으로 보인다(Gray & Wegner, 2012). 본 연구에서는 정서적 기능을 지원하는 사회적 로봇에 대한 마음지각이 경제적 로봇에 대한 마음지각보다 경험성과 주도성에서 모두 높으며, 도덕성 지위에서도 우위를 차지할 가능성을 보여주었다. 이 결과는 사회적 로봇을 활용하여 인간-로봇의 상호작용을 통해 정서와 연관된 도덕성 교육이 가능할 수 있음을 시사한다. 후속연구에서는 다양한 사회적 로봇의 활용이 인간-로봇의 상호작용에 기반한 인성교육의 효과에 끼치는 영향을 세분화하여 탐색해볼 필요가 있다.

참 고 문 헌

- 김우영 · 이태현 · 장재운 (2019). 온라인 설문에서 거울 이미지를 활용한 불성실 응답 방지 방법에 대한 연구.<한국심리학회지: 일반>, 38(4), 669-698.
- Bastian, B., Costello, K., Loughnan, S., & Hodson, G. (2012). When closing the human-animal divide expands moral concern. *Social Psychological and Personality Science*, 36, 100-107.
- Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, 3, 1-9.

- Bera, A., Randhavane, T., Kubin, E., Wang, A., Gray, K., & Manocha, D. (2018). Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS), 4468-4475.
- Dahl, T. S., & Boulos, M. N. K. (2013). Robots in health and social care: a complementary technology to home care and telehealthcare? *Robotics*, 3, 1-21.
- Damiano, L., & Dumouchi, P. (2018). Anthropomorphism in Human-Robot Co-evolution. *Frontiers in Psychology*, 9, 468.
- Duffy, B. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, 42, 177-190.
- Eagly, A. H., & Wood, W. (2011). Feminism and the evolution of sex differences and similarities. *Sex Roles*, 64, 758-767.
- Eagly, A. H., & Wood, W. (2012). Social role theory. In P. van Lange, A. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories in social psychology* (pp. 458-476). Thousand Oaks, CA: Sage Publications.
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11, 77-83.
- Friedman, B., Kahn, P., & Hagman, J. (2003). "Hardware companions?: What online AIBO discussion forums reveal about the human-robotic relationship," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, (New York, NY: ACM), 273-280.
- Graf, B., Reiser, U., Hägele, M., Mauz, K., & Klein, P. (2009). "Robotic home assistant Care-O-bot R 3-product vision and innovation platform," in Proceedings of the 2009 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO), (Piscataway, NJ: IEEE), 139-144.
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315, 619.
- Gray, K., & Wegner, D. M. (2008). The sting of intentional pain. *Psychological Science*, 19, 1260-62.
- Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, 96, 505-520.
- Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, 125, 125-130.
- Gray, K., & Wegner, D. M. (2012). Morality takes two: Dyadic morality and mind perception. In M. Mikulincer & P. R. Shaver (Eds.), *Herzliya series on personality and social psychology*(pp. 109-127). Washington, DC, US: American Psychological Association.
- Haslam, N., Bain, P., Douge, L., Lee, M., & Bastian, B. (2005). More human than you: Attributing humanness to self and others. *Journal of Personality and Social Psychology*, 89, 973-950.
- Haslam, N., Loughnan, S., Kashima, Y., & Bain, P. (2008). Attributing and denying humanness to others. *European Review of Social Psychology*, 19, 55-65.
- Hoffman, C., & Hurst, N. (1990). Gender stereotypes: Perception or rationalization? *Journal of*

- Personality and Social Psychology*, 58, 197-208.
- Loughnan, S., Haslam, N., Murnane, T., Vaes, J., Reynolds, C., & Suitner, C. (2010). Objectification leads to depersonalization: The denial of mind and moral concern to objectified others. *European Journal of Social Psychology*, 40, 709-717.
- Miller, H. B. (1994). Science, ethics, and moral status. *Between the Species*, 10, 1-10.
- Montague, A., & Matson, F. (1983). *The dehumanization of man*. New York: McGraw-Hill.
- Pollack, M. E., Brown, L., Colbry, D., Orosz, C., Peintner, B., & Ramakrishnan, S. (2002). "Pearl: a mobile robotic assistant for the elderly," in Proceedings of the AAAI Workshop on Automation as Eldercare (Menlo Park, CA: AAAI), 85-91.
- Scherbaum, S., & Dshemuchadse, M. (2020). Psychometrics of the continuous mind: Measuring cognitive sub-processes via mouse tracking. *Memory and Cognition*, 48, 436-454.
- Stafford, R. Q., MacDonald, B. A., & Jayawardena, C. (2014). Does the robot have a mind? Mind perception and attitudes towards robots predict use of an eldercare robot. *International Journal of Social Robotics*, 6, 17-32.
- Wang, X., & Krumhuber, E. G. (2018). Mind perception of robots varies with their economic versus social function. *Frontiers in Psychology*, 9, 1-9.
- Ward, A. F., Olsen, A. S., & Wegner, D. M. (2013). The Harm-made mind: Observing victimization augments attribution of minds to vegetative patients, robots and, the dead. *Psychological Science*, 24, 1437-1445.
- Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. *Advances in Experimental Social Psychology*, 46, 55-123.

[부록 1] 마음지각 질문지

	하위영역	능력 특성
1	경험성	위의 대상이 어느 정도 자신의 욕구를 가질 능력이 있다고 생각하십니까?
		위의 대상이 어느 정도 자신의 감정을 가질 능력이 있다고 생각하십니까?
		위의 대상이 어느 정도 자신의 정서를 가질 능력이 있다고 생각하십니까?
		위의 대상이 어느 정도 자신의 즐거움을 가질 능력이 있다고 생각하십니까?
		위의 대상이 어느 정도 자신의 두려움을 가질 능력이 있다고 생각하십니까?
		위의 대상이 어느 정도 자신의 배고픔을 느낄 능력이 있다고 생각하십니까?
		위의 대상이 어느 정도 자신의 성격을 가질 능력이 있다고 생각하십니까?
2	주도성	위의 대상이 어느 정도 스스로 계획을 할 능력이 갖고 있다고 생각하십니까?
		위의 대상이 어느 정도 자기조절능력을 갖고 있다고 생각하십니까?
		위의 대상이 어느 정도 기억능력을 갖고 있다고 생각하십니까?
		위의 대상이 어느 정도 정서인식능력을 갖고 있다고 생각하십니까?
		위의 대상이 어느 정도 도덕적 이해능력을 갖고 있다고 생각하십니까?
		위의 대상이 어느 정도 의사소통능력을 갖고 있다고 생각하십니까?
		위의 대상이 어느 정도 상황에 대한 영향력을 갖고 있다고 생각하십니까?

[부록 2] 다양한 대상에 대한 마음지각과 해치는 행동에 대한 심리적 불편감 및 비도덕성 평정의 관계(연구 1)

대상	변인	평균	표준편차	1	2	3
성인남자	1. 주도성	6.24	1.09			
	2. 경험성	5.98	1.01	.56**		
	3. 심리적 불편감	6.81	1.30	.35*	.39**	
	4. 비도덕성	6.89	1.10	.21	.21	.45*
성인여자	1. 주도성	5.89	1.12			
	2. 경험성	6.07	1.29	.49**		
	3. 심리적 불편감	6.63	1.54	.37**	.47**	
	4. 비도덕성	6.72	1.36	.19	.39**	.55**
5살 아동	1. 주도성	5.56	1.32			
	2. 경험성	6.09	1.23	.63****		
	3. 심리적 불편감	5.89	1.42	.18	.41**	
	4. 비도덕성	6.52	1.39	.13	.28*	.56**
신	1. 주도성	5.89	1.32			
	2. 경험성	5.16	1.08	.32**		
	3. 심리적 불편감	5.48	1.24	.18	.41**	
	4. 비도덕성	5.26	1.83	.13	.28*	.46**

대상	변인	평균	표준편차	1	2	3
장애인	1. 주도성	5.47	1.12			
	2. 경험성	5.98	1.32	.54**		
	3. 심리적 불편감	5.71	1.27	.20	.57**	
	4. 비도덕성	6.24	2.21	.12	.37**	.51**
84세 노인	1. 주도성	4.95	1.01			
	2. 경험성	4.72	1.13	.57**		
	3. 심리적 불편감	6.54	1.23	.27*	.50**	
	4. 비도덕성	6.37	1.58	.20	.19	.38**
개	1. 주도성	4.47	1.23			
	2. 경험성	4.82	1.07	.50**		
	3. 심리적 불편감	6.33	2.18	.17	.44**	
	4. 비도덕성	5.83	1.07	.70***	.67***	.35**
노숙자	1. 주도성	4.43	1.24			
	2. 경험성	4.70	1.54	.50**		
	3. 심리적 불편감	4.58	1.42	.08	.38**	
	4. 비도덕성	5.92	1.72	.03	.15	.37**
소	1. 주도성	4.16	1.02			
	2. 경험성	4.3172	1.23	.48**		
	3. 심리적 불편감	4.24	1.15	.35**	.41**	
	4. 비도덕성	5.18	1.44	.05	.17	.38**
돼지	1. 주도성	3.89	1.12			
	2. 경험성	4.01	1.20	.58**		
	3. 심리적 불편감	4.91	1.41	.27*	.34**	
	4. 비도덕성	4.33	1.75	.05	.19	.37**
4개월 태아	1. 주도성	3.24	1.12			
	2. 경험성	3.78	1.94	.34**		
	3. 심리적 불편감	5.28	1.45	.37**	.42**	
	4. 비도덕성	5.10	1.89	.15	.16	.40**
식물상태환자	1. 주도성	3.23	1.01			
	2. 경험성	3.48	1.39	.43**		
	3. 심리적 불편감	5.89	2.28	.38**	.41**	
	4. 비도덕성	5.72	1.77	.18	.20	.45**
로봇	1. 주도성	2.51	1.34			
	2. 경험성	2.02	1.21	.37**		
	3. 심리적 불편감	3.83	1.15	.21	.46**	
	4. 비도덕성	3.86	1.66	.19	.31**	.56**

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

[부록 3] 경제적 로봇과 사회적 로봇(연구 2)

경제적 로봇: 우리의 생활을 편리하게 해 주는 로봇

이 로봇은 우리가 시간을 더 효율적으로 사용할 수 있도록 도와줍니다.

예를 들어, 우리가 슈퍼마켓에 가서 원하는 제품을 찾지 못할 때 편리하게 도와줍니다. 또한 제품에 대한 질문사항이 있을 때, 언제나 다양한 정보를 신속하게 제공해줍니다.

사회적 로봇: 우리와 친구처럼 지내는 로봇

이 로봇은 우리가 혼자 있을 때, 우리와 같이 시간을 보내줍니다.

예를 들어, 우리가 슈퍼마켓에 갈 때, 같이 동행합니다. 사회적으로 고립되어 외로운 사람들과 함께 이야기를 나누며, 일상의 중요한 일정을 상기시켜 주기도 합니다.

Who Has a Mind?: Mind Perception and Moral Decision Toward Robots

Hong-Im Shin

Sunchon National University

According to the dyadic moral theory, ascribing a moral status to the specific entity depends on perceptions of a mind toward the entity. Based on this assumption, this research explored relationships between mind perception and moral decisions to provide theoretical and practical implications for future education. With the use of mind-perception scales, Study 1 (N=130) analyzed perceptions of a mind of various entities. According to the results, an adult man's mind is rated as the most highly valued, whereas a relatively lower mind is ascribed to vegetative patients and robots. In addition, the higher scores of mind perception correlated to the increasing tendency of perceived immorality in harming acts. Study 2 (N=114) analyzed differences between economic robots and social robots in mind perception. The participants evaluated that a social robot deserved a higher moral status than an economic robot. Additionally, harming acts were rated as more immorally in the condition of a social robot than those in the condition of an economic robot. These findings implicated that mind perception varies with a function of an entity and is linked to the essence of morality. Future research is needed to explore the applicability of social values of robots for morality education.

Keywords: Mind Perception, Robots, Dyadic Moral Theory, Moral Status