

자율주행차와 윤리적 의사결정: 누가 사는 것이 더 합당한가?

Who Should Live? Autonomous Vehicles and Moral Decision-Making

신홍임†
Hong Im Shin†

Abstract

The reduction of traffic accidents is a primary potential benefit of autonomous vehicles (AVs). However, the prevalence of AVs also arouses a key question: to what extent should a human wrest control back from AVs? Specifically, in an unavoidable situation of emergency, should an AV be able to decide between the safety of its own passengers and endangered pedestrians? Should AV programming include well-accepted decision rules about actions to take in hypothetical situations? The current study (N = 103) examined individual/situational variables that could perform critical decision-making roles in AV related traffic accidents. The individual variable of attitudes toward AVs was assessed using the Self-driving Car Acceptance Scale. To investigate situational influences on decisional processes, the study's participants were assigned to one of two groups: the achievement value was activated in one group and the benevolence value was triggered in the other through the use of a sentence completion task. Thereafter, participants were required to indicate who should be protected from injury: the passengers of the concerned AV, or endangered pedestrians. Participants were also asked to record the extent to which they intended to buy an AV programmed to decide in favor of the greater good according to Utilitarian principles. The results suggested that participants in the "achievement value: driver perspective" group expressed the lowest willingness to sacrifice themselves to save several pedestrians in an unavoidable traffic accident. This group of participants was also the most reluctant to buy an AV programmed with utilitarian rules, even though there were significant positive relationships between members' acceptance of AVs and their expressed intention to purchase one. These findings highlight the role of the decisional processes involved in the "achievement value" pertaining to AVs. The paper finally records the limitations of the present study and suggests directions for future research.

Key words: Autonomous Vehicles, Utilitarianism, Decision Making, Value

요약

자율주행차는 운전자의 부주의로 인한 사고를 줄일 수 있는 반면, 기계에 어느 정도의 자율성을 허용해야 할 것인지의 문제를 제기한다. 특히 돌발상황의 발생시에 운전자와 보행자 중에서 누구를 선택해야 할 것인지에 대한 의사결정은 모두가 합당하다고 생각하는 도덕적 원칙을 기반으로 일관적으로 실시되어야 한다. 본 연구(N = 103)에서는 한국 사회에서 자율주행차의 윤리적 의사결정을 프로그래밍한다고 가정했을 때, 의사결정과정에 영향을 끼칠 수 있는 개인적/상황적 변인을 검증하였다. 이에 따라 개인적 변인으로서 참가자의 자율주행차의 자동화 기술에 대한 수용도를 측정 후, 상황적 변인으로 문장완성과제를 통해 참가자에게 성취 또는 배려의 가치를 활성화시켰다. 이후 자율주행차의 사고상황에서 공리주의 의사결정의 비율 및 자율주행차 구매의향이 어떻게 달라지는지를 설문지를 통

※ 이 논문은 2018년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2018S1A5A8026962).

†(교신저자) 신홍임: 영남대학교 교양학부 교육중점 전임교원 / E-mail : shin7038@naver.com / TEL : 053-810-7831

해 비교하였다. 그 결과 연구참가자에게 성취의 가치를 활성화시켰을 때, 자신이 차 안의 운전자로 가정된 상황에서 다수의 보행자를 배려하는 공리주의 의사결정의 비율이 가장 낮게 나타났다. 또한 연구참가자가 자율자동차 기술을 긍정적으로 생각할수록 공리주의 원칙으로 설계된 자율주행차의 구매의향이 전반적으로 높았지만, 성취-운전자 조건에서 참가자의 구매의향이 다른 조건보다 유의하게 낮게 나타났다. 이 결과는 자율주행차의 사고상황에서 의사결정의 과정이 개인적 변인뿐만 아니라 특정 상황에서 활성화된 가치 및 관점의 차이에 따라 영향을 받을 가능성을 시사한다. 논의에서는 연구결과의 제한점 및 후속 연구방향에 관해 논의하였다.

주제어: 자율주행차, 공리주의, 의사결정, 가치

1. 서론

자율주행차는 더 이상 상상속의 차가 아니다. 구글사의 자율주행차는 벌써 실제 도로에서 수천마일을 주행하였다(Waldrop, 2015). 자율주행차는 운전자 부주의로 인한 교통사고를 90%까지 줄일 것으로 예측되고 있으며, 환경오염의 감소 및 도로교통의 효율성에 획기적으로 기여할 것으로 기대되고 있다(Bonnefon, Shariff, & Rahwan, 2016).

우리나라에서 2016년 개정된 「자동차관리법」은 자율주행차를 “운전자 또는 승객의 조작 없이 자동차 스스로 운행이 가능한 자동차”로 정의하고 있다. 자율주행차가 3단계에서 4단계로 발전될 경우 운전대와 브레이크가 사라지는 등의 자동차 내부설계가 바뀌면서 자율주행차에 전면적인 자율성이 주어진다(Lee, 2016). 기존에 운전자의 주의 의무는 자율주행차의 자율성으로 대체되기 때문에 자율주행차에게 긴급한 사고상황에서 어떤 도덕적 원칙을 프로그래밍해야 할 것인지는 중요한 사안이다. 또한 한국은 자율주행차와 관련된 윤리적 입장이 거의 준비되어 있지 않기 때문에 대책마련이 시급한 것으로 보인다(Byun, 2017).

자율주행차가 대중화되기에 앞서 시급한 사안은 자율주행차의 도덕적 원칙뿐만 아니라 사회구성원 모두가 이 원칙을 받아들일 수 있어야 하는 것이다(Karnouskos, 2018). Fig. 1과 같이 돌발 사고상황에서 가능한 많은 사람을 구하는 것이 더 합당하다는 공리주의 원칙(Greene, 2016)을 따른다면, 다수의 보행자를 구하기 위해 길 위에 서있는 한 명의 행인에게 피해를 입히는 선택을 하게 될 것이다. 또한 다수의 보행자를 구하기 위해 다른 대안이 없다면, 운전자 스스

로 피해를 입는 선택이 윤리적으로 더 적절한 것으로 보여진다. 그러나 선행연구(예: Bonnefon et al., 2016; Sütfeld et al., 2017)에 의하면, 다수를 위해 피해를 입는 한 사람을 자신 또는 가족이라고 가정하면, 연구참가자가 공리주의 원칙을 따르려 하는 경향이 급격하게 떨어졌다. 이러한 차이는 공리주의 원칙으로 자율주행차의 의사결정 프로그램을 설계한다면, 일반대중이 합당하다고 생각하는 윤리적 결정과 운전자 개인이 내리는 선택이 충돌하게 될 것을 보여준다. 다양한 입장 간의 갈등은 향후 자율주행차의 전면적 도입에 심리적 장애물로서 작용하여, 자율주행차의 대중화를 방해할 가능성이 있다. 본 연구에서는 연구참가자가 내린 공리주의 결정의 비율이 자율주행차의 다양한 사고상황에 따라 어떻게 달라지는지를 상황적 변인 및 개인적 변인에 따라 분석하고자 한다. 우선은 지금까지 실시된 선행연구를 개관하고, 본 연구와 기존연구와의 차별성을 기술하려 한다.

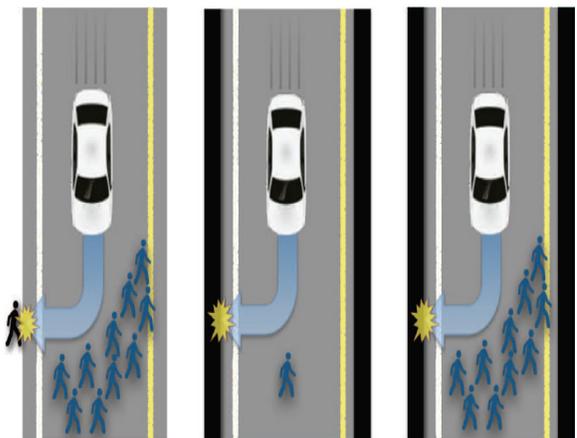


Fig. 1. Driverless dilemmas of autonomous vehicles in emergency situation

2. 이론적 배경

2.1. 공리주의적 입장

윤리적 의사결정에 관한 선행연구(Conway & Gawronski, 2013; Greene, 2009; Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Haidt, 2001)에 의하면, 공리주의 결정은 가능한 많은 사람들의 이득을 도출하는 행동을 도덕적으로 더 적절한 행동으로 판단하며(예: 다섯 명을 구하는 것이 한 사람을 살리는 것보다 더 옳다), 결과를 과정보다 더 중요하게 부각시킨다¹⁾. 자율주행차의 주제를 직접 다룬 연구는 아니지만, Greene 등(2001)의 연구에서는 스위치를 눌러 타인에게 해를 가하는 비개인적 상황에서는 뇌에서 사고를 담당하는 영역이 기저선보다 더 많이 활성화되고, 공리주의 결정이 증가하였다. 반면, 개인이 직접적인 신체적 접촉을 통해 타인에게 상해를 입히는 개인적 갈등상황(예: 열 명의 환자를 살리기 위해 한 명의 건강한 방문객을 직접 해치는 병원 딜레마)에서는 공리주의 결정이 감소하였다. 이 연구는 한 개인이 동일한 상황에 대해서도 관점의 차이에 의해 상황과의 연관성을 다르게 느끼기 때문에 윤리적 의사결정이 변화할 수 있음을 보여준다(Joo & Lee, 2013). 자율주행차의 사고상황에서도 관점의 차이(예: 운전자 vs. 관찰자)는 행동의 도덕적 적절성을 판단하는데 영향을 끼칠 가능성이 있다. 자신이 운전자라면, 상황과의 직접적 연관성을 경험하며, 자신의 자율주행차가 길 위의 다수 보행자를 위해 자신을 희생시키

는 것에 직관적으로 동의하지 않을 것이다. 반면 관찰자 입장에서는 상황으로부터 직접적 연관성을 덜 느끼며, 자율주행차는 운전자 한 명 보다는 다수의 보행자 중심으로 운행해야 한다고 생각할 것이다. 따라서 자율주행차의 돌발사고상황에서 두 가지 서로 다른 입장에서의 판단을 비교해보는 것은 모두가 보편적으로 받아들이는 윤리적 원칙을 수립하는데 있어서 필요할 것으로 보인다.

또한 윤리적 의사결정은 자율주행차의 다양한 사고 상황에 따라 달라질 것이다. 예를 들어 Bonnefon 등(2016)에서는 참가자들에게 다양한 사고상황을 제시하고, 각 상황에서 운전자가 자신보다는 다수의 보행자를 보호하기 위해 스스로에게 피해를 가하는 선택을 할 수 있는지에 대해 온라인 설문조사를 실시했다. 그 결과 참가자의 76% 이상이 다수의 보행자를 구하기 위해 차 안에 있는 한 명의 승객이 피해를 입는 것이 도덕적으로 더 적절하다고 응답했다. 반면 자율주행차의 공리주의 원칙에 대한 호감도(평균 85점, 범위: 0-100점)는 높았지만, 공리주의 원칙을 기반으로 자율주행차가 실제 설계되는 것에 대한 동의비율은 33%였다. 따라서 공리주의 원칙에 대한 호감도와 실제 적용에 대한 동의정도는 서로 연관되지 않았다. 또한 연구참가자 자신이 사고상황에서 운전자로 가정될 경우, 보행자를 구하기 위해 운전자 자신을 희생하는 결정을 하는 경향은 거의 나타나지 않았다. 이와 더불어 정부의 주도하에 공리주의 원칙으로 설계된 자율주행차의 구매의향이 정부가 프로그램 설계에 개입하지 않은 자율주행차보다 더 낮게 나타났다. 연구자들은 이 결과를 공리주의 결정은 일반적으로 합리적으로 인식되고, 이 원칙을 기반으로 다른 사람들이 결정을 내려야 한다고 생각되지만, 자율주행차에 앓은 승객이 바로 자신일 경우, 상황에 대해 직접적 연관성을 느끼기 때문에, 한 개인이 공리주의 결정을 선택하는 것은 거의 불가능한 것으로 설명했다. 따라서 단순히 공리주의 원칙만을 기반으로 자율주행차 의사결정 프로그램을 설계할 때, 실제 사고상황에서 다양한 입장 간의 사회적 갈등(예: 운전자, 일반 대중, 정부, 소비자)이 예측된다고 볼 수 있다.

위에 서술한 다양한 입장 간의 갈등 및 의사결정의 비밀관성은 한 개인이 의사결정의 상황에서 선택을

1) 선행연구(Conway & Gawronski, 2013; Greene, 2009; Greene et al., 2001; Haidt, 2001)에서는 두 가지 유형의 윤리적 의사결정을 구분한다. 공리주의 결정은 가능한 많은 사람들의 이득을 도출하는 결과를 과정보다 더 중요하게 부각시킨다. 이에 비해 의무론에서는 의도적으로 타인을 위해하지 않는다는 도덕적 원칙을 따르는 것이 인간의 의무라고 주장한다(Byun, 2017). 따라서 최소한의 피해일지라도 인간을 수단으로 활용하려는 의도가 행동에 담겨있다면, 옳지 않은 행동으로 보기 때문에(예: 한 명의 보행자라도 무고한 인간의 생명을 의도적으로 빼앗을 수 없다), 도덕판단에서 결과보다는 과정을 더 중요시하는 입장이다. 본 연구에서는 이 두 가지 입장 중에서 공리주의 결정에 제한하여 연구참가자의 윤리적 의사결정을 살펴보고자 한다.

할 때, 그 상황에 어느 정도 직접적으로 연관된 것으로 인식하느냐와 관계가 있다. 자신이 차안의 승객인 경우와 아닌 경우 또는 공리주의원칙의 프로그램이 자신의 차에 설계되는 경우와 아닌 경우에 따라 의사결정의 차이가 나타났다. 본 연구에서는 모두가 합당하다고 인정할 수 있는 자율주행차의 의사결정 프로그램을 설계하기 위해 의사결정의 차이에 영향을 끼치는 상황적 변인과 개인적 변인을 탐색하고자 한다.

2.2. 윤리적 의사결정과 가치

Schwartz(1992)의 사회적 가치의 원형모형(circular model of social values)에 의하면, 가치(value)는 한 개인의 신념이며, 행동을 동기화한다(Schwartz, 1992; 1996). Schwartz(1992)에 의하면, 성취의 가치는 타인보다는 자기(self)를 더 중요시하는 자기향상(self-enhancement)의 영역에 속하며, 능력, 성공 및 영향력으로 정의된다. 반면, 배려의 가치는 자기초월(self-transcendence)의 영역에 속하며, 타인의 행복에 대한 자발적 관심, 도움과 책임으로 정의된다. 예를 들어, 다수의 보행자와 자신을 선택해야 하는 갈등상황에서 성취의 가치가 활성화되면, 타인보다는 자기를 더 중요시하기 때문에 다수의 보행자보다 자신을 선택하는 행동을 하게 될 가능성이 더 높다. 배려의 가치는 자기초월 및 타인에 대한 관심과 연관되기 때문에, 한 명의 보행자와 자신을 선택해야 하는 상황에서 보행자를 선택하는 개인의 결정비율이 성취의 가치가 활성화되는 조건보다 더 커질 가능성이 높다. 따라서 한 개인에게 특정 상황에서 활성화되는 가치를 통해 윤리적 의사결정의 과정을 더 잘 이해할 수 있을 것이다.

또한, 가치는 한 개인의 고정불변의 성향이 아니라 상황에 의해 영향을 받아 무의식적으로 활성화될 수 있다(Maio, Pakizeh, Cheung, & Rees, 2009). Maio 등(2009)에서는 성취 또는 배려와 연관된 단어를 포함시켜 참가자들에게 단어분류과제를 수행하도록 했을 때, 성취조건에서는 배려조건보다 과제수행도가 증가했다. 반면 배려 조건에서는 통제조건보다 도움행동이 증가하였다. 또한 Hart와 Albarracin(2009)에서는 성취와 연관된 단어를 나열한 후, 참가자들에게 문장

완성과제를 지시하였을 때, 참가자들의 성취동기가 통제조건에 비해 더 높아졌다. 따라서 한 개인이 일상적으로 추구하는 가치 이외에도 외부 상황에서 특정한 가치가 무의식적으로 활성화된다면, 의사결정과정에서 영향을 받을 가능성을 추정해볼 수 있다(Chae, 2017). 이에 따라 본 연구에서는 성취/배려의 가치를 실험상황에서 활성화시킨 후, 자율주행차의 사고상황에서 연구참가자의 선택이 영향을 받는지를 살펴보려 한다. 이 결과를 통해 향후 발생할 수 있는 돌발사고 상황에서 상황적으로 활성화된 가치가 어떤 영향을 끼치는지를 추정할 수 있을 것이다.

본 연구의 필요성 및 선행연구와의 차별성은 다음과 같다. 첫째, 자율주행차의 의사결정에 관한 연구(Bonnefon et al., 2016)는 지금까지 국외 및 국내에서 거의 실시되지 않았다. 자율주행차의 의사결정은 생명과 직결된 문제이기 때문에 자율주행차의 대중화에 앞서서 충분히 연구되고, 논의되어야 할 문제이다. 둘째, 본 연구에서는 의사결정과 연관된 상황적 변인으로서 한 개인에게 특정 상황에서 활성화되는 가치의 문제를 분석하려 한다. 이 결과를 통해 개인이 자율주행차의 윤리적 결정에 대한 판단을 내릴 때, 사회적으로 활성화된 가치가 어떤 영향을 끼치는지를 추정할 수 있을 것이다. 셋째, 지금까지 자율주행차의 윤리적 의사결정에 대한 연구에서는 의사결정의 상황에서 발생하게 될 다양한 입장 간의 갈등(예: 운전자 관점, 관찰자 관점)에 영향을 끼치는 개인적 변인에 대해 거의 다루지 않았다. 최근 Nees(2016)는 자율주행차에 대한 의사결정과정에 영향을 끼칠 수 있는 개인적 변인을 탐색하기 위해 자율주행차 수용도 검사(SCAS: self-driving car acceptance scale)를 개발했는데, 본 연구에서는 이 검사를 기반으로 자율주행차의 자동화 기술에 대한 개인의 태도가 사고상황에서 윤리적 의사결정과 구매의향에 어떤 영향을 끼치는지를 살펴보려 한다.

본 연구의 연구문제는 다음과 같다.

첫째, 자율주행차의 사고상황에서 성취의 가치 또는 배려의 가치가 활성화된다면, 운전자와 다수의 보행자를 선택하는 행동에서 차이가 나타나는가? 본 연구에서는 성취의 가치가 활성화된 상황에서 배려의 가치가 활성화된 조건보다 운전자와 다수의 보행

자 중에서 다수의 보행자를 선택하는 비율이 더 낮을 것을 예측한다. 둘째, 자율주행차의 사고상황에서 관점의 차이에 따라 공리주의적 결정비율이 달라지는가? 본 연구에서는 자율주행차와 더 직접적 연관성을 느끼게 되는 운전자 관점에서 관찰자 관점보다 다수의 보행자를 선택하는 행동의 비율이 더 낮아질 것을 예측한다. 셋째, 상황적으로 활성화된 가치와 사고상황에 대한 관점의 차이는 다수의 보행자를 우선시하는 공리주의 원칙으로 설계된 자율주행차의 구매의향에 어떤 영향을 끼칠 것인가? 본 연구에서는 성취의 가치가 활성화된 상황에서 운전자 관점으로 판단할 때, 공리주의 원칙의 자율자동차 구매의향이 가장 낮을 것을 예측한다. 넷째, 연구참가자의 자율주행차에 대한 사전 태도는 다수의 보행자를 우선시하는 공리주의 원칙으로 설계된 자율주행차의 구매의향에 어떤 영향을 끼칠 것인가? 본 연구에서는 자율주행차에 대한 긍정적 태도가 강할수록 공리주의 원칙으로 설계된 자동차를 구매하려는 의향이 더 높을 것을 예측한다.

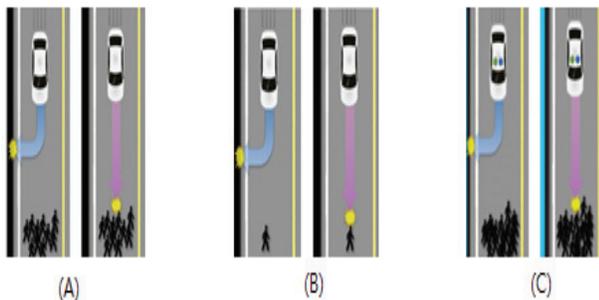


Fig. 2. Three traffic situations involving imminent unavoidable harm. The car must decide between (A) killing several pedestrians or its own passenger, (B) killing one pedestrian or its own passenger, and (C) killing several pedestrians or its own two passengers.

3. 연구방법 및 절차

3.1. 참가자

4년제 대학교의 대학생 103명(남 = 49, 평균연령 만 24.56세, 표준편차 = 2.51)이 대학교 홈페이지 자유게시판 공고를 통해 자발적으로 참가하였다. 연구참가

자들은 총 4개의 집단에 무선적으로 배정되었다. 각 집단의 연령 및 성별분포는 다음과 같다: 성취-운전자 집단(총 26명, 남 = 12, 평균 연령 만 23.92세), 성취-관찰자 집단(총 27명, 남 = 13, 평균 연령 만 25.26세), 배려-운전자 집단(총 26명, 남 = 12, 평균 연령 만 24.77세), 배려-관찰자 집단(총 24명, 남 = 11, 평균 연령 만 24.25세). 연구 참가에는 약 10분의 시간이 소요되었으며, 참가자들은 오천원 상당의 커피쿠폰의 보상을 받았다.

3.2. 연구도구 및 절차

본 연구는 온라인 설문지를 기반으로 실시되었다. 참가자는 우선 Nees(2016)가 개발한 자율주행차 수용검사(Self-driving Car Acceptance Scale)에 응답하였다(부록 1 참조). 이 질문지는 총 24개 문항으로 구성되어 있으며, 본 연구에서는 영어 문항을 한국어 문항으로 번역하여 사용하였다. 자율주행차 수용검사는 총 여덟 개의 하위영역으로 다음과 같이 구성되어 있다: (1) 자동화에 대한 신뢰, (2) 비용, (3) 자동화의 적절성 여부, (4) 자동화된 기계를 좋아하는 정도, (5) 자동화의 유용성을 인식하는 정도, (6) 자동화의 편리함을 인식하는 정도, (7) 자동화에 대한 경험 및 (8) 자동화를 적용하려는 의지. 참가자는 각 문항에 대해 자신이 동의하는 정도를 7점 척도(1: 전혀 동의하지 않음, 7: 아주 동의함)에 따라 응답하였다. 전체 검사문항 24개 중에서 13개의 문항은 역채점(reverse-scored)문항이었다(문항 3번, 6번, 7번, 8, 9, 10, 11, 12, 17, 18, 20, 23, 24번). 문항의 점수가 높을수록 자동화 기술에 대한 긍정적 태도로 해석할 수 있다. Nees(2016)의 연구에서 검사문항의 내적합치도(Cronbach's α)는 .901이었으며, 본 연구에서는 .754로 나타났다.

자율주행차에 대한 인식 검사 이후 실험상황에서 특정한 가치의 활성화가 참가자의 의사결정과정에서 영향을 끼치는지를 검증하기 위해 성취 또는 배려의 가치가 무의식적으로 연상되도록 하였다. 이를 위해 참가자는 성취집단과 배려집단에 무선적으로 배정되었다. 문장완성과제에서는 선행연구(Fitzsimons & Shah, 2008)를 참조하여, 참가자에게 목표단어(예: 성취)를 무의식적으로 연상시키는 단어 및 여러 다른 단어를

제시하고, 문장을 완성하도록 지시하였다. 이 과제에서는 참가자가 제시된 단어(예: 경쟁, 능력)를 활용하여 문장을 만들어내는 과정에서 연관된 목표단어(예: 성취)를 무의식적으로 연상하게 되는 원리를 활용하였다.

본 연구의 문장완성과제에서 참가자들은 총 열 개의 문장을 완성하도록 지시받았다. 성취집단과 배려집단에서 참가자들은 한 문장 당 다섯 개의 단어(예: 영어, 외국, 하늘, 성공, 전화하다)를 제시받았으며, 이 단어를 활용하여 문장을 작성해야 했다. 성취집단에서는 성취의 가치와 연관된 열 개의 단어(예: 성공, 능력)가 제시되었고, 배려집단에서는 배려의 가치와 연관된 열 개의 단어(예: 도움, 용서)가 사용되었다. 이 과제에서 성취/배려의 가치와 연관된 단어는 선행 연구(Maio et al., 2009; Bardi & Schwartz, 2003)를 토대로 선정하였다. 그 외 다른 마흔 개의 단어들은 이 두 조건에서 모두 동일했으며, 각 조건에서는 성취/배려의 가치와 연관된 단어들을 포함하여 총 오십 개의 단어가 사용되었다.

문장완성과제가 완료된 후, 성취집단과 배려집단은 각각 다시 두 집단으로 나누어져 두 종류의 다른 질문지(운전자 vs. 관찰자)에 응답하였다. 첫 번째 운전자 집단은 참가자 자신이 자율주행차 안에 앉아 있는 상황을 가정하는 집단이었고, 두 번째 관찰자집단은 참가자가 관찰자로서 자율주행차의 의사결정을 판단하는 상황을 가정하는 집단이었다. 이에 따라 본 연구에서 참가자들은 총 네 개의 상이한 집단(성취_운전자, 성취_관찰자, 배려_운전자, 배려_관찰자)에 무선적으로 배정되어, 자율주행차 사고상황에 대한 질문지에 응답하였다. 각 집단에서 연구참가자는 Fig. 2 (Bonneson et al, 2016)와 같이 세 가지 유형의 상황을 설명하는 글과 그림을 모두 동일하게 제시받았다(부록 2 참조). 첫 번째 상황은 자율주행차 맞은편에 갑자기 나타난 보행자가 열 명인 조건, 두 번째 상황은 운전자와 보행자가 각각 한 명인 조건이었다. 마지막으로 세 번째 유형은 운전자 옆에 동승자가 있었으며, 보행자가 열 명 이상이 있는 조건이었다. 각 집단에서 상황설명문의 제시순서는 역균형화되어서, 순서효과에 의해 결과가 영향을 받지 않도록 하였다. 연구 참가자는 모두 동일하게 세 가지 유형의 상황을 설명하

는 글을 읽은 후에 각 상황에 대한 네 개의 문항에 응답하였다. 문항은 다음과 같다: 첫째, 보행자를 무시하고, 계속 직진할 것인지 또는 옆 방향으로 차를 돌려 운전자가 피해를 입는 것이 더 적절한 것인가? 둘째, 사고상황에서 자율주행차가 결정을 한다면 보행자에게 가능한 피해를 적게 입히려는 방향으로 설계되는 것이 옳은가, 또는 운전자가 피해를 입지 않는 방향으로 설계되는 것이 옳은가? 셋째, 자율주행차가 운전자보다는 보행자의 피해를 가능한 줄이려는 방향으로 설계된다면, 자율주행차를 구입하려는 의향이 어느 정도인가? (7점 척도에서 응답함) 넷째, 방금 전에 제시된 상황에서 자율주행차 전방에 있는 보행자는 총 몇 명이었는가? 이 문항은 참가자가 상황설명문을 주의 깊게 읽었는지를 확인하는 문항이며, 참가자가 이 문항에 오답을 제시하면, 분석에서 제외되었다. (본 연구에서는 총 105명이 참가하였으며, 이 중 2명이 상황설명을 점검하는 문항에서 오답을 표기하여 자료분석에서 제외되었다. 2명의 탈락자는 각각 성취-운전집단에 1명, 배려-관찰집단에 1명이 속해있었다.)

자율주행자동차에 대한 의사결정과제가 완료되면, 마지막으로 본 연구에서 실시한 성취/배려의 가치점화를 확인하기 위해 참가자에게 성취목표를 어느 정도 추구하는지를 자기보고식 질문지에 따라 10점 척도 상에서 평정하도록 했다. 질문지의 문항은 Hart와 Albarrín(2009)에서 사용된 세 개의 문항(예: ‘나는 성공적으로 일을 잘 해 낼 때, 가장 만족스럽다’, ‘일의 성과가 좋지 않을 때, 나는 건디기 힘들다’) 번역하여 사용했다.

3.3. 연구 설계

본 연구는 4(집단: 성취-운전자, 성취-관찰자, 배려-운전자, 배려-관찰자) x 3(상황 시나리오 조건: 보행자 10명 vs. 보행자 1명 vs. 동승객 및 보행자 열명)의 혼합설계로 구성되었다. 첫 번째 변인은 집단간 변인이며, 두 번째 변인은 집단내 변인이었다. 측정변인은 공리주의 의사결정의 횟수 및 자율자동차 구매의향이었다.

3.4. 분석 방법

본 연구에서는 외부적으로 활성화된 가치와 다양한 사고 상황의 상호작용에 따라 자율자동차에 대한 윤리적 의사결정의 차이가 나타나는지를 검증하고자 하였다. 이를 위해 SPSS 18.0 통계분석프로그램을 사용하여 기술통계, 독립표본 t-검정, 상관분석, 교차분석 및 혼합변량분석을 수행하였다.

4. 연구 결과 및 논의

4.1. 조작점검

본 연구에서는 문장완성과제를 통해 성취/배려의 가치를 참가자에게 무의식적으로 접화시켰다. 이에 따라 실험조작이 참가자의 가치 활성화에 영향을 끼쳤는지를 분석하기 위해 각 조건에 따라 참가자의 성취동기를 확인하였다. 그 결과 성취집단(M = 6.499, SD = .763)에서는 배려집단(M = 5.201, SD = 1.021) 보다 참가자의 성취동기의 점수가 더 높게 나타났다, $t(101) = 3.35, p = .042$. 이 결과는 실험조작이 참가자의 가치접화에 차별화된 영향을 끼친 것을 보여준다.

4.2. 자율주행차 기술 수용도 비교

본 연구의 네 집단이 자율주행차 기술에 대한 수용도에서 사전에 차이가 있었는지를 확인하기 위해 독립표본 t-검정을 수행하였다. 이를 위해 Table 1과 같이 자율주행차 기술에 대한 태도의 8가지 하위영역에서 각 영역별 평균점수를 산출한 후, 전체 영역의 평균을 산출하였다. 또한 Table 2와 같이 각 집단별 평균분석을 수행하였다. 그 결과 집단 간에 유의한 차이가 나타나지 않았다. 따라서 본 연구에서 자율주행차 사고상황에 대한 공리주의적 의사결정비율에서 조건 간 차이가 나타난다면, 이 결과는 연구참가자가 사전에 갖고 있었던 태도로 인한 차이라기보다는 본 연구에서 유도된 성취/배려의 가치 및 운전자/관찰자 관점의 차이로 인한 가능성이 더 높다.

Table 1. Means and standard deviations of self-driving car acceptance scale items across groups (AD: Achievement-driver, AO: Achievement-observer, BD: Benevolence-driver, BO: Benevolence-observer)

Sub-dimension	Group	M	SD
Perceived reliability	AD	4.730	.587
	AO	4.333	1.286
	BD	4.384	.778
	BO	5.000	1.215
Cost of automation	AD	4.000	1.292
	AO	3.345	.935
	BD	4.025	1.086
	BO	4.138	1.116
Appropriateness	AD	4.794	.816
	AO	5.111	.933
	BD	5.153	.900
	BO	4.694	1.044
Enjoyment	AD	3.615	1.371
	AO	3.814	1.353
	BD	3.512	1.174
	BO	4.111	1.399
Perceived usefulness	AD	4.589	1.076
	AO	4.074	1.372
	BD	4.410	1.405
	BO	5.055	1.136
Perceived ease of use	AD	4.538	1.363
	AO	3.925	1.838
	BD	4.769	1.274
	BO	4.416	1.348
Experience with automation	AD	5.923	.913
	AO	5.796	1.137
	BD	5.884	.778
	BO	5.541	.943
Intention to use	AD	5.000	1.442
	AO	4.444	2.118
	BD	4.769	1.704
	BO	4.833	1.711
Total	AD	4.649	.590.
	AO	4.355	.708
	BD	4.613	.622
	BO	4.724	.524

Table 2. Mean comparisons of self-driving car acceptance scale items between groups (AD: Achievement-driver, AO: Achievement-observer, BD: Benevolence-driver, BO: Benevolence-observer)

(I) Group	(J) Group	I-J	t	p
AD	AO	.293	1.634	.108
	BD	.035	.210	.835
	BO	-.074	-.473	.638
AO	AD	-.293	-1.634	.108
	BD	-.258	-1.406	.166
	BO	-.368	-1.517	.097
BD	AD	-.035	-.210	.835
	AO	.258	1.406	.166
	BO	-.110	-.674	.504
BO	AD	.074	.473	.638
	AO	.368	1.517	.097
	BD	.110	.674	.504

4.3. 성취/배려의 가치와 운전자/관찰자 집단에 따른 공리주의적 의사결정 비교

본 연구의 결과를 네 가지 집단(성취-운전자, 성취-관찰자, 배려-운전자, 배려-관찰자)과 세 가지 유형의 상황 시나리오 조건(보행자 10명 vs. 보행자 1명 vs. 동승자 및 보행자 10명 이상)에 따라 구분하여 빈도분석을 수행하였다.

우선, 첫 번째 상황(보행자 10명 vs. 운전자 1명)에서는 전체 참가자의 78.6%가 차를 돌려, 운전자를 희생시키더라도 보행자 열 명을 구하는 것이 도덕적으로 더 적절한 것으로 응답하였다. 이에 비해, 두 번째 상황(보행자가 1명 vs. 운전자 1명)에서는 운전자를 희생시켜 보행자를 구해야 한다는 응답이 전체 참가자의 29.1%로 나타났다. 반면, 세 번째 상황(보행자가 열 명 이상 vs. 운전자 및 동승자)에서는 전체 참가자의 55.4%가 차를 돌려 보행자를 구해야 한다고 응답했다.

Table 3과 같이 교차분석을 수행한 결과, 첫 번째 상황에서 카이제곱 검증값이 유의하였다, $\chi^2 = 22.442, p < .001$. 성취-운전자 집단에서는 열 명의 보행자보다는 현재 운전자를 살려야 한다는 응답이 전체 참가자의 13.6%를 차지하여 다른 세 집단에 비해 가장 많이 나타났다. 이에 비해, Table 4와 같이 보행자가 한 명인 상황에서는 카이제곱 검증값이 유의하지 않았다, $\chi^2 = 1.511, p = .682$. 따라서 네 집단의 연구참가자가 모두 보행자가 한 명인 경우에는 보행자보다는 운전자를 구하는 것이 더 적절한 것으로 판단하는 것으로 나타났다. 반면, Table 5와 같이 보행자

Table 3. Observed and expected frequencies of making utilitarian decisions for the greater good across conditions regarding AV's traffic accidents with ten pedestrians and its own passenger

Group	To save ten pedestrians		To save its own passenger	
	Observed	Expected	Observed	Expected
AD	12 (11.7%)	20.4	14 (13.6%)	5.6
AO	25 (24.3%)	21.2	2 (1.9%)	5.8
BD	22 (21.4)	20.4	4 (3.9%)	5.6
BO	22 (21.4%)	18.9	2 (1.9%)	5.1

Table 4. Observed and expected frequencies of making moral decisions for the greater good across conditions regarding AV's traffic accidents with one pedestrian and its own passenger

Group	To save one pedestrian		To save its own passenger	
	Observed	Expected	Observed	Expected
AD	6 (5.8%)	7.6	20 (19.4%)	18.4
AO	10 (9.7%)	7.9	17 (16.5%)	19.1
BD	8 (7.8%)	7.6	18 (17.5%)	18.4
BO	6 (5.8%)	7.0	18 (17.5%)	17.0

Table 5. Observed and expected frequencies of making utilitarian decisions across groups regarding AV's traffic accidents with more than ten pedestrians and its own two passengers

Group	To save more than ten pedestrians		To save its own two passengers	
	Observed	Expected	Observed	Expected
AD	4 (3.9%)	14.4	22 (21.4%)	11.6
AO	19 (18.4%)	14.9	8 (7.8%)	12.1
BD	20 (19.4%)	14.4	6 (5.8%)	11.6
BO	14 (13.6%)	13.3	10 (9.7%)	10.7

가 열 명 이상인 상황에서 운전자 옆에 앉은 동승자(친구)가 있는 경우, 집단에 따라 유의한 차이가 나타났다, $\chi^2 = 24.250, p < .001$. 성취-운전자 집단에서 차를 돌려 보행자를 구해야 한다는 응답비율(3.9%)이 가장 낮았다. 이에 비해 성취-관찰자, 배려-운전자 조건 및 배려-관찰자 집단에서는 각각 18.4%, 19.4% 및 13.6%로 운전자와 동승자를 희생시키더라도 열 명의 보행자를 보호하는 것이 도덕적으로 더 적절한하다고 판단했다.

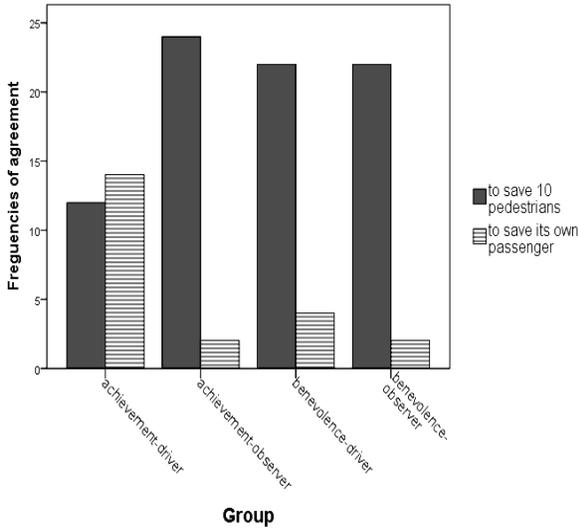


Fig. 3. Frequencies of utilitarian decisions for the greater good in traffic accidents with 10 pedestrians and its own passenger across conditions

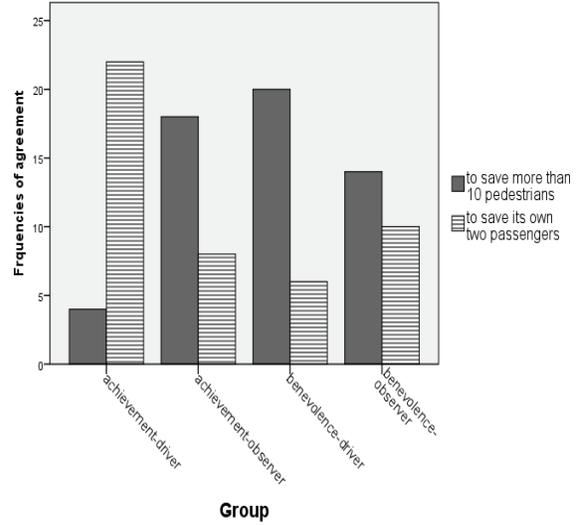


Fig. 4. Frequencies of utilitarian decisions for the greater good in traffic accidents with more than 10 pedestrians and its own two passengers across conditions

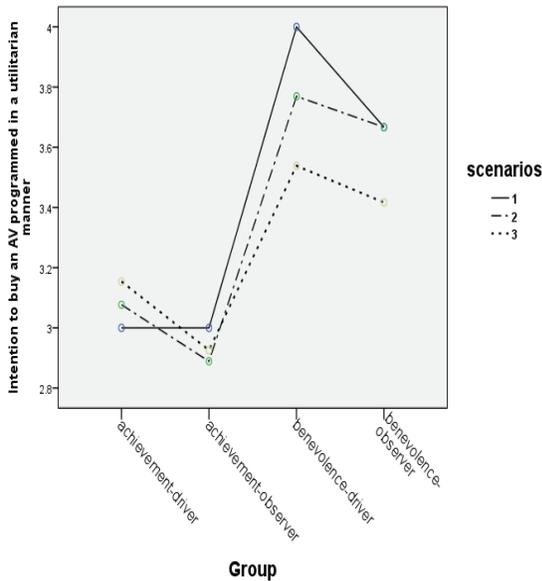


Fig. 5. Graphed interactions for intention to buy an AV programmed in a utilitarian manner across groups (scenario 1: its own passenger vs. 10 pedestrians, scenario 2: its own passenger vs. one pedestrian, scenario 3: its own two passengers vs. more than ten pedestrians)

요약하면, 보행자가 한 명인 경우에는 집단 간 차이 없이 운전자를 구해야 한다는 결정을 내렸지만, 보행자가 다수인 경우에는 성취-운전자 집단에서 다른 집단에 비해 보행자를 구해야 한다는 응답이 가장 낮았다. 따라서 보행자가 한 명인 상황을 제외한 다른 상황에서 모두 성취-운전자 관점이 다수를 위해 소수를 희생해야 한다는 공리주의 결정을 감소시키는데 영향

을 끼친 것으로 보인다.

4.4. 자율주행차 의사결정에 대한 도덕적 판단 비교

본 연구에서는 세 가지 유형의 다양한 상황에서 자율주행차가 의사결정을 하게 된다면, 운전자와 보행자 중에서 누구를 우선시하도록 설계되는 것이 도덕적으로 적절할 것인지에 대한 질문을 실시하였다. 본 연구의 결과를 네 가지 집단(성취-운전자, 성취-관찰자, 배려-운전자, 배려-관찰자)과 세 가지 유형의 상황 시나리오 조건(보행자 10명 vs. 보행자 1명 vs. 동승자 및 보행자 10명)에 따라 구분하여 빈도분석을 수행하였다.

우선, Table 6과 같이 첫 번째 상황조건(보행자 10명 vs. 운전자 1명)에서 자율주행차가 결정을 내린다면, 전체 참가자의 56.3%가 차를 돌려, 운전자를 희생시키더라도 보행자 열 명을 구하도록 설계되는 것이 도덕적으로 더 적절한 것으로 응답하였다. 이에 비해, 보행자가 한 명인 상황에서는 운전자를 희생시켜 보행자를 구하도록 자율자동차가 설계되는 것이 도덕적으로 더 적절하다는 응답이 전체 참가자의 46.6%로 나타났다. 반면 Table 7과 같이 보행자가 열 명 이상이며, 운전자 옆에 타고 있는 친구가 있는 상황에서 자율주행차가 결정을 내린다면, 전체 참가자의 49.6%가 자율주행차가 보행자를 구하는 결정을 내리도록 설계되어야 한다고 응답했다.

Table 6. Observed and expected frequencies of making utilitarian decisions of autonomous vehicles across groups regarding AV's traffic accidents with ten pedestrians and its own passenger

Group	To save ten pedestrians		To save its own passenger	
	Observed	Expected	Observed	Expected
AD	10 (9.7%)	14.6	16 (15.5%)	11.4
AO	16 (15.5%)	15.2	11 (10.7%)	11.8
BD	14 (13.6%)	14.6	12 (11.7%)	11.4
BO	18 (17.5%)	13.5	6 (5.8%)	10.5

Table 7. Observed and expected frequencies of making utilitarian decisions of autonomous vehicles across groups regarding AV's traffic accidents with more than ten pedestrians and its own two passengers

Group	To save more than ten pedestrians		To save its own two passengers	
	Observed	Expected	Observed	Expected
AD	4 (3.9%)	12.9	22 (21.4%)	13.1
AO	15 (14.6%)	13.4	12 (11.7%)	13.6
BD	18 (17.5%)	12.9	8 (7.8%)	13.1
BO	14 (13.6%)	11.9	10 (9.7%)	12.1

Table 6과 같이 교차분석을 수행한 결과, 카이제곱 검증값이 유의하지 않았지만, 경향성이 나타났다, $\chi^2 = 6.934, p = .078$. 성취-운전자 집단에서 가장 적은 연구참가자(9.7%)가 자율주행차가 결정을 내린다면, 현재의 운전자 한 명보다는 열 명의 보행자를 살리도록 프로그램이 설계되어야 한다는 응답을 하였다. 이에 비해, 자율주행차가 결정을 내리는 경우에 보행자가 한 명인 상황에서는 카이제곱 검증값이 유의하지 않았다, $\chi^2 = 2.072, p = .558$. 따라서 보행자가 한 명인 경우에는 운전자를 구하는 것이 더 적절하다는 응답이 네 가지 집단의 연구참가자에게서 유의한 차이 없이 나타났다. 반면, 보행자가 열 명 이상일 때, 운전자 옆에 타고 있는 친구가 있는 상황에서 자율주행차가 결정을 내린다면, 보행자와 차안에 앉은 승객 중에서 누구를 더 우선시하도록 설계되어야 할 것인가에 대해 집단에 따라 유의한 차이가 나타났다, $\chi^2 = 17.300, p < .001$. Table 7과 같이 각 집단에 따라 살펴

보면, 성취-운전자 집단에서 차를 돌려 보행자를 구해야 한다는 응답비율(3.9%)이 가장 낮았다. 이에 비해 성취-관찰자, 배려-운전자 집단 및 배려-관찰자 집단에서는 각각 14.6%, 17.5% 및 13.6%로 자율주행차가 열 명이상의 보행자를 우선적으로 보호하도록 설계되는 것이 도덕적으로 더 적절하다고 판단했다.

요약하면, 자율주행차 의사결정이 공리주의 원칙에 따라 다수의 보행자를 구하는 것이 도덕적으로 더 적절하다고 판단하는 것은 성취-운전자 집단에서 가장 낮았다. 그 외 다른 집단(성취-관찰자, 배려-운전자 및 배려-관찰자)에서는 집단 간에 유의한 차이가 없었다. 따라서 자율주행차가 의사결정을 내리도록 프로그램을 설계한다고 가정해도 성취의 가치가 활성화되고, 운전자 관점에서 상황과의 직접적 연관성을 경험할 때, 공리주의 원칙을 받아들이는데 가장 어려움이 있는 것으로 나타났다.

4.5. 다수의 보행자를 우선시하는 공리주의 원칙의 자율자동차 구매의향 비교

자율자동차가 공리주의 원칙에 따라 운전자 피해보다는 다수의 보행자 피해를 최소화하는 원칙으로 설계된다고 가정할 때, 자율자동차 구매의향이 달라지는 정도를 분석하기 위해, 4(집단: 성취-운전자, 성취-관찰자, 배려-운전자, 배려-관찰자) x 3(상황 시나리오: 보행자 1명 vs. 보행자 10명 vs. 동승자 및 보행자 10명)의 혼합변량분석을 수행하였다. 이 중 첫 번째 변인은 집단간 변인이었고, 두 번째 변인은 집단내 변인이었다. 측정변인은 자율자동차의 구매의향이었다.

우선 상황별 분석을 살펴보면, Table 8과 같이 공리주의 원칙의 자율자동차 구매의향은 첫 번째 상황(보행자 10명 vs. 운전자 1명)에서 가장 높았다. 평균분석에 의하면, 공리주의 원칙으로 설계된 자율자동차 구매의향은 운전자 한 명을 희생시켜서 보행자 열 명을 구할 수 있는 시나리오(M = 3.412, SD = .144)에서 운전자 및 동승자 모두를 희생시키는 조건에서 운전자 및 동승자를 희생시키고, 보행자 열 명 이상을 구하는 조건(M = 3.256, SD = .134)에서보다 더 높았다, $t(102) = 2.247, p = .015$. 다른 상황조건 간에는 구매의향에서 유의한 차이가 나타나지 않았다.

Table 8. Means and standard deviations of intention to buy an AV programmed in a utilitarian manner across groups and scenarios

Scenario	Group	M	SD
Its own passenger vs. 10 pedestrians	AD	3.001	1.442
	AO	3.008	1.776
	BD	4.012	.980
	BO	3.678	1.523
	Total	3.412	1.501
Its own passenger vs. one pedestrian	AD	3.087	1.164
	AO	2.897	1.672
	BD	3.778	1.070
	BO	3.671	1.465
	Total	3.341	1.397
Its own two passengers vs. more than ten pedestrians	AD	3.157	1.317
	AO	2.931	1.708
	BD	3.542	.948
	BO	3.421	1.349
	Total	3.256	1.363

한편 변량분석결과에 의하면, 상황조건의 주효과가 유의하였다, $F(2,198) = 3.950, p = .021, \eta^2 = .038$. 집단 간의 주효과는 유의하지 않았지만, $F(3, 198) = 2.238, p = .009, \eta^2 = .064$, 집단과 상황의 상호작용이 유의했다, $F(6, 198) = 3.055, p = .007, \eta^2 = .085$. Fig. 4와 같이 성취-운전자 집단과 성취-관찰자 집단에서는 세 가지 유형의 상황에서 모두 일관적으로 다른 집단보다 공리주의 원칙으로 설계된 자동차의 구매의향이 낮게 나타났다. 반면 배려-운전자 집단에서는 공리주의 원칙에 따라 설계된 자율자동차 구매의향이 상황에 따라 다르게 나타났다. 이 집단에서는 운전자와 동승자를 모두 포기해야 하는 상황($M = 3.54, SD = .948$)에서 운전자만을 포기하는 상황($M = 4.00, SD = .980$)보다 공리주의 원칙에 기반한 자율자동차 구매의향이 유의하게 낮았다, $t(25) = -3.094, p = .005$. 또한 배려-관찰자 집단에서도 운전자와 동승자를 모두 포기해야 하는 상황($M = 3.42, SD = 1.349$)에서 공리주의에 따라 설계된 자율주행차 구매의향이 운전자만을 포기하는 상황($M = 3.67, SD = 1.523$)에서의 구매의향보다 유의하게 낮았다, $t(23) = -2.769, p = .011$. 이에 비해 성취-운전자 집단 및 성취-관찰자 집단에서는 다양한 상황조건에 따라 구매의향이 유의하게 달라지지 않았다.

요약하면, 공리주의 원칙에 기반한 자율자동차 구매의향은 배려의 가치가 활성화될 때, 성취의 가치가 활성화될 때보다 돌발 상황에서 운전자 이외 누구를 포기해야 하는지에 의해 더 많은 영향을 받는 것으로

보인다. 반면 성취의 가치가 활성화되면, 운전자 이외 포기해야 되는 대상의 상황적 특성에 의해 많은 영향을 받지 않고, 일관적으로 공리주의 결정에 대한 저항이 있는 것으로 보인다.

4.6. 자율주행차 기술에 대한 수용도와 공리주의 기반의 자율자동차 구매의향의 상관관계

Table 9와 같이 자율주행차 기술에 대한 참가자의 태도에서 참가자의 점수가 높을수록(태도가 긍정적일수록) 다수의 보행자 우선원칙의 자율자동차 구매의향이 전반적으로 더 높게 나타났다, $r = .311, p = .001$. 또한 이 상관관계는 세 가지 유형의 상황(보행자 10명 vs. 보행자 1명 vs. 동승자 및 보행자 10명이상)에서 모두 유의하게 나타났다. 예를 들어, 운전자와 동승자가 있는 조건에서 다수의 보행자를 구하겠다는 응답은 연구참가자의 자동화기술에 대한 긍정적 태도가 높을수록 더 높게 나타났다, $r = .227, p = .011$. 이 결과는 한 명의 운전자만을 위하기보다는 보행자 우선 원칙으로 다수의 피해를 줄이는 방향으로 자율자동차가 설계된다고 하더라도 연구참가자가 사전에 자율주행차의 자동화 기술에 대한 긍정적 태도가 형성되어 있을수록 돌발상황에서 다수를 위해 운전자가 희생되는 결정을 내리는, 공리주의 원칙기반 자율자동차를 수용하고, 구매할 의향이 높은 것으로 해석할 수 있다.

Table 9. Means and Correlations between acceptance of AV's technology and intention to buy an AV programmed in a utilitarian manner for the greater good

Variables	Mean (SD)	Correlations		
		1	2	3
1. Acceptance of AV's technology	4.581 (.624)			
2. Intention to buy an AV programmed to sacrifice its own passenger to save 10 passengers	3.412 (1.501)	.311*		
3. Intention to buy an AV programmed to sacrifice its own passenger to save one passenger	3.341 (1.397)	.295*	.922**	
4. Intention to buy an AV programmed to sacrifice its own two passengers to save more than ten passengers	3.256 (1.363)	.227*	.906**	.923**

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

5. 결론

본 연구에서는 참가자에게 무의식적으로 성취 또는 배려와 연관된 가치가 연상되는 문장완성과제를 수행하도록 했을 때, 연구참가자의 운전자 또는 관찰자 관점에 따라 자율주행차의 사고상황에서 공리주의적 의사결정의 비율이 달라지는지를 분석하였다. 그 결과 연구참가자들은 보행자가 한 명일 경우 조건에 따라 차이가 없이 운전자를 구하는 방향으로 결정을 내리는 것이 더 적절하다는 판단을 내렸다. 이에 비해 보행자가 열 명인 경우에는 운전자를 희생시켜서라도 다수의 보행자를 구해야 한다는 응답이 전반적으로 나타났다. 이 결과는 선행연구(Bonnefon et al., 2016)에서 보고한 결과와도 일치한다. 따라서 한 대의 자율자동차로 보행자가 열 명이상을 해치게 될 사고상황이 실제 현실에서 발생할 가능성은 낮지만, 이 상황에서는 운전자 자신을 희생시켜서라도 다수의 보행자를 구하는 것이 더 적절하다는 판단이 보편적임을 알 수 있다.

또한 여기에서 중요한 것은 보행자가 다수일 경우 성취-운전자 관점에서는 다른 관점에 비해 다수의 보행자를 선택하는 비율이 유의하게 낮았던 것이다. 이 결과는 본 연구에서 예측했던 방향과 일치한다. 흥미로운 것은 배려의 가치가 활성화될 경우, 희생자가 자신의 옆에 앉은 동승자일 경우 다수의 보행자를 희생시켜서라도 동승자를 구해야 한다는 반응이 운전자 혼자 있는 상황보다 유의하게 더 높았던 것이다. 이 결과는 동승자가 있는 상황은 윤리적 의사결정에서 공리주의적 결정을 내리는데 부정적 영향을 끼칠 가능성을 보여준다. Greene 등(2004)의 도덕판단 이중처리모형(Dual Process Theory)에 따르면, 윤리적 의사결정과정의 과정에는 심사숙고하는 사고와 직관적 정서가 중요한 역할을 한다. 이 모형에 의하면 사고가 주된 역할을 할수록 공리주의결정이 증가하고, 정서/직관이 주도적이면, 공리주의 결정이 감소한다(Choi, Ahn., & Na, 2011). 배려의 가치가 활성화되어 있는 상황에서도 자신에게 중요한 동승자가 옆에 앉아있는 상황에서는 의사결정의 과정에 정서가 개입되어, 다수의 보행자를 우선시하는 공리주의적 의사결정을 감소시킬 가능성이 있다. 향후 후속연구에서는 동승자를 정서적 친밀감에 따라 다양한 집단으로 세분화하

여(예: 친구, 가족, 직장 동료 등), 의사결정과정의 차이를 분석해볼 필요가 있다.

본 연구의 네 가지 집단 중에서 공리주의 결정을 내리는데 사고상황에서 가장 큰 저항을 보였던 집단은 성취-운전자 집단조건이었다. 이 조건에서는 향후 자율자동차가 공리주의 원칙으로 설계된다면, 자율자동차를 구매할 의향이 가장 낮게 나타났다. Bonnefon 등(2016)의 일련의 연구에서 연구참가자들은 자율자동차가 공리주의 원칙으로 설계되는 것이 적절하다고 응답했지만, 자신이 실제 구매할 의향은 매우 낮게 응답했다. 성취의 가치가 현대 사회에서 사회적으로 자주 부각됨을 감안할 때, 성취-운전자 집단은 실제 현장에서 가장 많이 나타날 수 있는 집단이다. 따라서 본 연구에서 나타난 공리주의 기반으로 설계된 자율자동차에 대한 낮은 구매의향은 향후 자율자동차가 대중화되는 실제 상황에서 다양한 입장 간의 심각한 충돌로 발전될 가능성이 있음을 보여준다.

본 연구의 한계점은 다음과 같다. 첫째, 본 연구는 돌발사고 상황에서 보행자 수에 따른 공리주의적 의사결정의 관계에 초점을 맞추었다. 따라서 실제 현실 상황의 다양한 변수를 적용하지 못한 가상의 단순화된 사고상황이기 때문에 결과의 해석에 한계가 있다. 사고원인은 보행자 수뿐만 아니라 핸들을 쥐는 각도, 브레이크 작동방법, 노면 또는 타이어상태 등에 따라서도 달라지며, 운전자의 반응분포도 이에 의해 영향을 받을 것이다. 또한 인공지능이 사고결과를 어느 정도 정확하게 예측하게 될 것인지도 중요한 변인이 될 수 있다. 가상의 상황이라는 부자연스러움은 참가자의 솔직한 반응에 영향을 끼쳐, 연구결과의 생태학적 타당도를 떨어뜨릴 가능성이 있다. 후속연구에서는 실제 사고상황에 가장 근접한 상황시나리오를 다양하게 도입하는 것이 필요할 것으로 보인다. 둘째, 본 연구에서는 연구참가자 개인적 변인으로서 자율자동차 자동화 기술에 대한 수용도를 측정하였다. 연구참가자가 자동화기술에 대한 태도가 긍정적일수록 공리주의 원칙 기반의 자율자동차 구매의향이 높게 나타났다. 그러나 이 결과는 본 연구에 참가했던 대학생이 연령대가 20대 초반으로 집중되어 있기 때문에 자동화 혁신 기술에 대한 긍정적 태도가 이미 형성되어 있을 가능성이 높아서 연구결과를 일반화하는데 한계가

있다. Nees(2016)가 지적했던 것과 같이 새로운 혁신 기술에 대한 이상화는 긍정적 태도만을 형성하여, 향후 예측하지 못한 갈등상황이 발생했을 때, 이 기술에 대한 불신이 형성되어, 장기적으로 이 기술이 보급되어 정착되는데 오히려 부정적 영향을 끼칠 수 있다. 또한 본 연구에서는 참가자가 대학생이었기 때문에 자신이 운전자로서 실제 상황에서 돌발사고를 경험하고, 복잡한 의사결정의 과정을 거쳤을 가능성이 높지 않다. 이것 역시 자율자동차의 자동화기술에 대한 추상적이고 긍정적 태도만을 갖게 할 수 있기 때문에, 본 연구결과의 해석에 한계가 있다. 향후 후속연구에서는 운전면허증을 소지한 숙련된 운전자로 연구참가자를 제한하여, 실제 운전상황의 사전 지식을 기반으로 자율주행차 의사결정의 과정이 달라지는지를 검증해볼 필요가 있다. 이와 더불어 후속연구에서는 다양한 연령대 및 직업군의 참가자 집단과 실제 상황에 근접한 상황을 세분화하여 포함시켜 연구결과의 일반화 가능성을 높일 필요가 있다. 셋째, 본 연구에서는 상황 시나리오를 활용하여 돌발상황에서의 공리주의적 의사결정비율을 측정하였다. 상황 시나리오에 대한 이해를 보행자 인원을 묻는 문항으로 확인하였지만, 참가자별로 글의 내용을 어떻게 이해했는지를 점검하지 않은 것을 연구의 제한점으로 볼 수 있다. 예를 들어 보행자가 열 명인 상황에서 참가자가 열 명 모두가 죽었는지 또는 이 중에서 한 명만 피해를 입었는지를 추가질문하지 않았기 때문에 연구참가자의 해석에 따라 의사결정이 달라질 가능성이 있다. 후속연구에서는 상황설명문의 이해를 정확하게 점검하는 것이 필요할 것으로 보인다. 넷째, 본 연구에서는 특정한 가치의 활성화가 자율주행차의 사고상황에서 의사결정의 과정에 영향을 끼칠 가능성을 보여주었다. 본 연구에서 활용한 문장완성과제는 외부의 작은 자극이 무의식적으로 운전자와 보행자 중에서 보호해야 할 대상을 선택하는데 영향을 끼칠 수 있음을 보여주었다. 그러나 문장완성과제는 일상의 상황에 많이 근접하지 않기 때문에 아쉬움이 남는다. 후속연구에서는 생태학적으로 더 자연스러운 실험과제를 활용하여, 연구결과의 외적 타당도를 높일 필요가 있다.

REFERENCES

- Bardi, A., & Schwartz, S. H. (2003). Values and behavior: Strength and structure of relations. *Personality and Social Psychology Bulletin*, 29(10), 1207-1221.
- Bonnefon, J., Shariff A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science*, 352(6293), 1573-1576. DOI: 10.1126/science.aaf2654
- Byun, S. (2017). An approach to ethical guidelines of autonomous vehicle. *Journal of Ethics*, 112, 199-216. DOI: 10.15801/je.1.112.201703.199
- Chae, J. (2017). The effects of shopping value, ease of use, and usefulness on mobile purchase intention. *Science of Emotion & Sensibility*, 20(2), 73-86. DOI: <https://doi.org/10.14695/KJSOS.2017.20.2.73>
- Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclination in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology*, 104(2), 216-235. DOI: 10.1037/a0031021.
- Choi, N., Ahn, R., & Na, G. (2011). A Study of thinking style and consumption behavior in consumer's decision making. *Science of Emotion & Sensibility*, 14(2), 279-292.
- Fitzsimons, G. M., & Shah, J. Y. (2008). How goal instrumentality shapes relationship evaluations. *Journal of Personality and Social Psychology*, 95(2), 319-337. DOI: 10.1037/0022-3514.95.2.319
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44(2), 389-400. DOI: 10.1016/j.neuron.2004.09.027
- Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, 107(3), 1144-1154. DOI: 10.1016/j.cognition.2007.11.004
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science*, 293(5537), 2105-2108.

- DOI: 10.1126/science.1062872
- Greene, J. D. (2009). Dual-Process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie. *Journal of Experimental Social Psychology, 45*(3), 581-584. DOI: 10.1016/j.jesp.2009.01.003
- Greene, J. D. (2016). Our driverless dilemma: When should your car be willing to kill you? *Science, 352*(6393), 1514-1515.
DOI: 10.1126/science.aaf9534
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814-834.
DOI: 10.1037/0033-295x.108.4.814
- Hart, W., & Albarracín, D. (2009). The effects of chronic motivation and achievement primes on the activation of achievement and fun goals. *Journal of Personality and Social Psychology, 97*(6), 1129-1141. DOI: 10.1037/a0017146
- Joo, M., & Lee, J. (2013). Differential effects of self-relevance levels on framing effects in decision making. *Science of Emotion & Sensibility, 16*(2), 177-186.
- Karnoukos, S. (2018). Self-driving car acceptance and the role of ethics. *IEEE Transactions on Engineering Management, 99*, 1-14.
DOI: 10.1109/TEM.2018.2877307
- Koop, G. J. (2013). An assessment of the temporal dynamics of moral decisions. *Judgement and Decision Making, 8*(5), 527-539.
- Lee, J. (2016). How the technology of autonomous driving affects the scope and level of driver's duty of care and the necessity for embedding ethical ability in autonomous vehicles. *Hongik Law Review, 17*(4), 445-472.
DOI: 10.16960/jhrl.17.4.201612.443
- Maior, G. R., Pakizeh, A., Cheung, W.-Y., & Rees, K. J. (2009). Changing, priming, acting on values: Effects via motivational relations in a circular Model. *Journal of Personality and Social Psychology, 97*(4), 699-715. DOI: 10.1037/a0016420
- Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., & Cusimano, C. (2015). Sacrifice one for the good of many? People apply different moral norms to human and robot agents. *Proceedings of the Tenth Annual ACM/IEEE International conference on Human-Robot Interaction, 117-124*.
DOI: 10.1145/2696454.2696458
- Nees, M. A. (2016). Acceptance of self-driving cars: An examination of idealized versus realistic portrayals with a self-driving car acceptance scale. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 60*(1), 1449-1453.
DOI: 10.1177/1541931213601332
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology, 25*, 1-65.
- Schwartz, S. H. (1996). *Value priorities and behavior: Applying the theory of integrated value systems* (Vol. 8). Mahwah, NJ: Erlbaum.
- Sütfield, R., Gast, R., König, P., & Gordon, P. (2017). Using virtual reality to assess ethical decisions in road traffic scenarios: Applicability of value-of-life-based models and influences of time pressure. *Frontiers in Behavioral Neuroscience, 11*, 1-13.
DOI: 10.3389/fnbeh.2017.00122.
- Waldrop, M. M. (2015). Autonomous vehicles: No drivers required. *Nature, 518*(7357), 20-23.
DOI: 10.1038/518020a.

원고접수: 2019.08.22

수정접수: 1차 2019.09.27

2차 2019.10.21

게재확정: 2019.10.21

부록 1. 자율주행자동차 수용도 검사

연번	영역	문항내용
1	자동화에 대한 신뢰	자율주행차는 안전하다.
2		자율주행차가 목적지까지 나를 데려다줄 수 있을 것이다.
3		사람들은 자율주행차를 감시해야 할 필요가 있다.
4	비용	자율주행차가 기존의 차보다 더 비싸도 구매할 의향이 있다.
5		자율주행차의 이득은 가격대비 더 많다고 생각한다.
6		자율주행차를 구매하기 전에 비용측면을 제일 중요하게 따져보아야 한다.
7	자동화의 적절성 여부	나는 컴퓨터가 자동차를 몰아서는 안 된다고 생각한다.
8		자율주행차를 통제하는 것이 중요하다.
9		자율주행차가 처리하기에는 너무 어려운 상황이 있다.
10	자동화된 기계를 좋아하는 정도	나는 운전과 같은 기계조작을 좋아한다.
11		차안의 승객보다는 운전자가 되는 것이 나는 더 좋다.
12		나는 운전을 좋아하고 즐긴다.
13	자동화의 유용성을 인식하는 정도	자율주행차는 업무효율성을 증가시킬 것이다.
14		자율주행차는 안전성을 증가시킬 것이다.
15		자율주행차는 교통사고를 감소시킬 것이다.
16	자동화의 편리함을 인식하는 정도	자율주행차는 사용하기 쉬울 것이다.
17		자율주행차를 조작하려면 많이 배워야 할 것이다.
18		자율주행차 조작을 이해하는데 까지 시간이 많이 걸릴 것이다.
19	자동화에 대한 경험	기계는 내 업무를 수월하게 하는데 도움이 된다.
20		나는 새로운 기술을 시도해서, 안 좋은 경험이 있었다.
21		나는 컴퓨터를 통해 많은 업무가 편리해지는 것을 경험했다.
22	자동화를 적용하려는 의지	나는 자율주행차를 갖고 싶다.
23		자율주행차를 갖게 된다면, 나는 대체로 내가 운전하고 싶다.
24		운전자 스스로 자율주행차의 기능을 조절할 수 있는 것이 중요하다.

부록 2. 상황설명문 예시

1. 아래의 글과 그림을 같이 보시고, 질문에 V표로 답해주십시오.

당신은 지금 고속도로에서 혼자 운전하고 있습니다. 그런데 갑자기 열 명의 보행자가 차 바로 앞에 나타났습니다. 이 돌발 상황에서 당신은 두 가지의 대안 중에서 하나만을 선택할 수 있습니다. 첫 번째 선택은 차의 방향을 도로 옆으로 돌려, 열 명의 보행자를 살리고, 당신 자신은 도로 옆의 장애물로 인해 목숨을 잃게 되는 것입니다. 두 번째 선택은 차를 계속 직진시켜 열 명의 보행자를 사망에 이르게 하지만, 당신은 살게 되는 것입니다.

1) 위의 글에서 제시한 두 가지 선택안중에서 어느 것을 선택하시겠습니까? (둘 중에 하나만 선택하십시오)
 차의 방향을 돌린다 _____ 직진한다 _____

2) 사고상황에서 당신이 아니라 자율주행차가 결정을 한다면, 보행자에게 가능한 피해를 적게 입히려는 방향으로 설계되는 것이 옳습니까, 또는 운전자가 피해를 입지 않는 방향으로 설계되는 것이 옳습니까? (둘 중에 하나만 선택하십시오)
 보행자 피해를 가능한 최소화 _____
 운전자 피해를 가능한 최소화 _____

3) 자율주행차가 운전자보다는 보행자의 피해를 가능한 줄이려는 방향으로 설계된다면, 자율주행차를 구입하려는 의향이 어느 정도입니까? 자신의 의견과 가장 일치하는 정도에 V표로 응답해주십시오(1: 전혀 구매의향이 없다, 7: 매우 구매하고 싶다)

1	2	3	4	5	6	7
전혀						매우
구매의향이						구매하고
없다						싶다

4) 지금 읽으셨던 글에 제시된 상황에서 보행자는 총 몇 명이었습니까?

_____ 명