

학습피드백으로서 보상과 처벌 관련 두뇌 활성화 연구*

김 진 희

강 은 주[†]

강원대학교 심리학과

본 연구의 목적은 학습상황에서 피드백으로 주어지는 금전적 획득/손실(학습 피드백)과 비학습적 상황에서 우연히 제시되는 의사 피드백(무선 피드백)을 비교하는 방법을 사용하여, 금전적 보상과 처벌의 학습 피드백으로서만 가지는 정보처리에 어느 두뇌 영역이 관여하는지를 규명하는 데 있다. 이를 위해 정상 성인($n = 22$)을 대상으로 fMRI scan 동안 단서 자극에 대한 범주 버튼 반응(좌/우)의 정확 여부에 따라 피드백이 제시되는 시행(학습시행)과 단서 자극의 위치판단 반응과 무관하게 피드백이 제시되는 시행(무선시행)을 사건 관련 fMRI 방략으로 제시하였다. 두 시행 간 보상과 처벌과 같은 동기적 사건에 대한 두뇌 반응이 변별적으로 나타나는지를 알아보기 위해 시행 유형(학습 vs. 무선)과 피드백 유형(보상 vs. 처벌)을 두 독립변인으로 한 반복측정 이원분산분석을 하였다(voxel-wise FWE $p < .001$). 그 결과, 좌측 배외측 전두피질(dorsolateral prefrontal cortex), 좌측 전측 도(anterior insular), 배내측 전두피질(dorsomedial prefrontal cortex) 등의 영역에서 유의한 상호작용 효과가 관찰되었는데, 이들 영역은 모두 학습-보상 피드백 및 무선-처벌 피드백보다 학습-처벌 피드백에 대해 증가한 두뇌 활성을 보였다. 본 연구 결과는 학습상황에서 주어지는 처벌 피드백에 대한 기존 전략의 변경이나 재평가를 위한 집행적 처리, 적절하지 못하거나 틀린 행동에 대한 오류처리 과정 그리고 실패 경험에 대한 부정적 정서처리가 위에서 언급한 피질신경망을 중심으로 이루어질 가능성을 보여준다. 따라서 학습의 처벌 피드백은 보상과 달리 위와 같은 추가적 정보처리 과정이 존재할 가능성을 시사한다.

주제어 : 강화학습, 부정 피드백, 오류처리, 전측도, 전전두피질

* 본 연구는 한국연구재단을 통해 미래창조과학부의 뇌과학원천기술개발사업으로부터 지원받아 수행됨 (NRF-2006-2005112).

[†] 교신저자: 강은주, 강원대학교 심리학과, (24341) 강원도 춘천시 강원대학길 1
연구분야: 인지신경심리학, 학습과 기억

Tel: 033-250-6850, E-mail: ekang@kangwon.ac.kr

음식, 금전적 획득 등의 긍정적 정서(valence)를 가지는 사건이나 통각, 금전적 손실 등과 같은 부정적인 정서가를 가지는 사건은 그 자체 경험이 우연에 의한 것일 수도 있고 또는 행동 결과로 얻어진 것일 수도 있다. 이러한 사건이 행동 반응에 대한 피드백 역할을 하는 경우 유기체는 그 결과를 얻거나 피하고자 해당 행동 가능성을 증가 또는 감소시킨다(Skinner, 1938). 조작적 조건화에 대한 동물 연구에서 주로 행동 결과로서 음식, 물, 성적 활동, 그리고 전기충격 등과 같은 자극을 사용하는데, 이러한 자극은 특별한 경험 없이 유기체의 행동을 증가 또는 감소시킬 수 있는 생득적인 요인이기에 일차적 강화인(primary reinforcer)이라 부른다. 강화 효과가 없는 자극이 일차적 강화인과 연합하여 행동을 증가시킬 수 있는 강화인이 된 경우에 이를 조건화된 강화인(conditioned reinforcer) 또는 이차적 강화인(secondary reinforcer)이라 부른다. 인간에게는 일차적 강화물뿐만 아니라 돈, 칭찬, 사회적 인정 등과 같은 좀 더 추상적인 형태인 사회적 강화물을 가질 수 있는데, 이런 이차적 강화물도 행동 수정과 학습을 일으키는 중요한 요인이 된다. 특히 돈은 현대사회에서 큰 사회적 가치를 가짐과 동시에 강한 동기적 요인으로 간주되기 때문에 학습 연구에서 보상(reward)과 처벌(penalty) 자극으로 널리 사용된다(Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Knutson, Adams, Fong, & Hommer, 2001).

보상 관련 두뇌 신경망은 인간 및 동물을 대상으로 신경생리학적, 약물학적, 손상, 기능 영상 기법을 이용한 다양한 연구를 통해 규명되어 왔다(Berridge & Robinson, 1998; Schultz, Dayan, & Montague, 1997). 보상 처리에 관여한다고 알려진 두뇌 영역으로 중뇌의 도파민 세포(dopaminergic neuron), 이와 신경해부학적으로 연결된 영역인 복측 선조체(ventral striatum) 중 측좌핵(nucleus accumbens), 흑질(substantia nigra), 편도체(amygdala), 안와전두영역(orbitofrontal cortex) 등이 알려져 있다. 우선, 동물을 대상으로 한 전기생리학적 연구들은 도파민 뉴런의 반응이 보상 예측 역할을 하며 보상의 발생을 예견하는 조건화된 자극으로 변화함을 밝혔는데 이 기제가 보상 학습의 기저가 된다(Montague, Dayan, & Sejnowski, 1996; Schultz, 1997). 인간을 대상으로 한 fMRI 연구에서는 중뇌 도파민 뉴런의 신호를 받는 선조체 영역에서 유사한 두뇌 활성화 패턴을 발견하였다. 예를 들어, 금전적 획득(Knutson et al., 2001), 매력적 얼굴(Bray & O'Doherty, 2007), 음악(Salimpoor et al., 2013) 등과 같은 다양한 보상을 예측하는 동안 측좌핵의 활성화 증가를 발견하여, 측좌핵의 보상예측오류(reward prediction error)의 역할을 나타내고 있다. 미상핵(caudate nucleus)과 피각(putamen)으로 구성된 배측 선조체(dorsal striatum)는 신체 운동을 조절하거나 정교화하는 데 관여한다고 알려져 있는데, 복측 선조체(ventral striatum)와는 기능적으로 구분된다. 예를 들어 복측 선조체의 경우 미래 보상을 예측하는 동안 활성이 관찰되는 반면, 배측 선조체는 보상을 주는 행동에 대한 정보를 유지하여 보상을 받기 위한 적절한 행동을 실행하는 동안 높은 활성을 보인다(O'Doherty et al., 2004; Tricomi, Delgado, & Fiez, 2004). 또한, 금전적 보상 피드백을 받은 후에 높은 활성화를 보이는 영역으로 안와전두피질이 보고된다. 안와전두피질의 활성화는 금전적 보상의 증가와 유의미한 상관관계를 보였다(Knutson, Fong, Bennett, Adams, & Hommer,

2003). 음식 자극(O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001) 또는 후각 자극(Anderson et al., 2003)과 같은 일차적 강화물을 사용한 연구들은 안와전두피질이 보상이 부호화에 관여하며 보상에 대한 평가적 역할을 함을 밝혔다(Hare, Camerer, & Rangel, 2009).

학습 상황에서 처벌 피드백은 자신의 실수 반응에 대한 모니터링을 위한 정보를 제공하기에 오류처리과정(error processing)을 수반한다. 이에 관여한다고 알려진 두뇌 영역으로 배측 전대상피질(dorsal anterior cingulate cortex; dACC)/배내측 전전두피질(dorsomedial prefrontal cortex; dmPFC)이 있다. 배측 전대상피질/배내측 전전두피질은 실수를 인지할 때 나타나는 오류 관련 부정 전위(error related negativity)의 진원지로 알려져 있으며(Debenner et al., 2005; Iannaccone et al., 2015), 내적 오류 신호(internal error signal)뿐만 아니라 금전적 손실과 같은 외적 오류 신호(external error signal) 처리에도 관여한다(Holroyd et al., 2004; Shackman et al., 2011). 복내측 전전두영역(ventromedial frontal lobe) 손상 환자들이 정상인과 달리 정적 피드백(금전적 획득)을 통한 학습은 정상적이거나, 부정적 피드백(금전적 손실)을 통한 학습 능력이 선택적으로 상실되었음이 보고되기도 하였다(Wheeler & Fellows, 2008). 금전적 손실과 같은 처벌 피드백에는 오류 처리 이외에도 부정적 피드백이 야기시킬 수 있는 부정적 정서 처리가 일어날 가능성도 고려해 볼 수 있는데(Kuhnen & Knutson, 2005; Xu, Liang, Wang, Li, & Jiang, 2009), 전측 도(anterior insula)가 선택적으로 관여할 가능성이 높다. 보상 피드백과 처벌 피드백을 비교 분석한 다수의 fMRI 연구에서도 처벌과 관련하여 전측 도의 활성화가 자주 보고되고 있다(Bischoff-Grethe, Hazeltine, Berggren, Ivry, & Grafton, 2009; Ullsperger & von Cramon, 2003). 전측 도는 시상 및 편도체의 광범위한 연결을 가지고 있으며(Stephani, Fernandez-Baca Vaca, Maciunas, Koubeissi, & Lüders, 2011), 내적 신체 상태뿐만 아니라 부정적 정서 및 각성을 표상하는 영역으로 알려져 있다(Gu, Hof, Friston, & Fan, 2013; Singer, Critchley, & Preuschoff, 2009). 또한, 전측 도는 정서처리의 역할 뿐만 아니라 부정적 사건 예측에 관여하는 것으로 알려져 있다(Kim, Shimojo, & O'Doherty, 2006; Samanez-Larkin, Hollon, Carstensen, & Knutson, 2008).

앞서 살펴본 바와 같이 학습 상황에서 금전적 획득과 손실이라는 피드백은 인지적 요인은 물론이고 동기적인 요인(Pessiglione et al., 2007) 또는 정서적 요인(Ressler, 2004)을 포함하고 있다. 이는 금전적 보상이나 처벌이 적절한 행동으로 접근하게 도와주는 정보이자 동기적 사건임과 동시에 긍정적 또는 부정적인 감정을 일으키는 정서적 사건이 된다는 것이다. 행동 조형이 요구되는 조작적 학습상황에서 금전적 피드백은 정보가(informational value) 외에도 동기적/정서적 속성을 포함한다. 이로 인해 행동 수정을 위한 순수한 피드백 역할을 하는 금전적 획득과 손실을 처리하는 데 관여하는 두뇌 영역을 연구하는 것이 용이하지 않다. 이를 구별하기 위해 두 가지 접근법을 생각해 볼 수 있다. 첫 번째 방법은 금전적 획득과 손실을 포함한 동기적/정서적 피드백 조건과 이를 배제한 중성적 피드백 조건을 비교하는 방법으로, 금전적 획득과 손실이 유발하는 피드백의 동기적/정서적 속성을 규명할 수 있다. 두 번째 방법은 금전적 획득과 손실을 학습 피

피드백으로 제시하는 조건과 동일한 사건을 무선적으로 제시해 학습 피드백 역할이 배제된 조건을 비교하는 경우이다. 이를 통해 피드백의 정보적 속성을 연구할 수 있을 것이다.

본 연구에서는 두 가지 접근법 중 후자인 피드백의 정보적 속성 연구에 초점을 두었다. 기존 피드백학 연구에서 피드백으로 사용되는 사건은 주로 정서가(정적, 부정)를 가지는 사건들로, 학습 피드백 정보로서 처리되는 과정, 이와 독립적인 정서적/동기적 사건 자체로 처리되는 과정을 구별하기가 쉽지 않았다. 보상과 처벌 사건이 특정 행동을 조형하기 위해서는 그 행동과 특정 피드백 사이에 적어도 50% 이상의 확률적 관계가 요구된다. 예를 들어 확률학습 과제를 사용한 연구에서 특정 자극을 선택하였을 경우 보상이 제시될 확률과 처벌이 제시될 확률이 각각(예, 80% vs. 20%) 존재할 수 있다. 적어도 어느 한쪽으로 치우쳐진 확률계획으로 보상과 처벌을 받을 경우 해당 확률적 피드백은 여전히 특정 자극에 대한 적절한 행동에 관한 정보를 가지고 있어 학습의 피드백 역할을 한다고 볼 수 있다. 이런 보상 피드백 처리과정 동안 안와전두영역(O'Doherty, Critchley, Deichmann, & Dolan, 2003), 미상핵(Delgado, Miller, Inati, & Phelps, 2005)이 관여함이 보고되었다. 참가자가 피드백의 무선적 제시 사실을 알지 못하는 경우 학습 피드백과 마찬가지로 의사 피드백(pseudo feedback)에 대해 미상핵 활성이 관찰되었는데(Seger & Cincotta, 2005), 이는 참가자들이 행동 수정을 위해 피드백 정보를 이용하며 이 처리 과정이 미상핵을 통해 행해지고 있음을 의미한다.

본 연구에서는 피드백 정보에 차이가 나는 두 시행 유형을 포함하였다. 학습시행에는 금전적 보상이나 손실을 특정 단서에 대한 반응에 따라 유관적으로 제시하여, 해당 피드백이 선택 반응의 정/오 반응(correct/incorrect response)에 관한 정보를 가지도록 하였다. 이와 달리 무선시행에서는 과제 수행의 정/오 반응과 무관하게 피드백을 제시하고 참여자에게 무관성 특성을 사전에 알림으로써 피드백으로서의 정보처리가 보상이나 처벌에 대해 이루어지지 않도록 하였다. 두 시행의 금전적 보상과 처벌은 물리적으로 동일한 정서적/동기적 사건이지만, 학습시행에서만 이 사건들이 이후에 자신의 행동 수정을 위한 정보로써 활용되는 피드백 정보처리가 요구될 것으로 가정하였다.

본 연구는 fMRI 기법을 이용하여 금전적 보상 및 처벌이 학습 피드백으로 제시되는 경우(학습시행; learning trial)와 무선 피드백으로 제시되는 경우(무선시행; random trial)의 정보처리 과정의 두뇌 영상을 획득하였다. 첫째, 피드백 유형(보상, 처벌)과 상관없이 학습시행과 무선시행의 피드백 간 비교를 통해 무선 피드백에 비하여 학습 피드백에만 높은 활성화를 나타내는 학습 특정적 피드백 처리 두뇌 영역을 규명하고자 하였다. 둘째, 피드백 정보의 유무(학습시행, 무선시행)와 상관없이 보상과 처벌 피드백 유형에 차별적 반응을 보이는지를 통해 피드백으로 사용된 동기적/정서적 사건들의 정서가 처리에 관여하는 두뇌 영역을 규명하고자 하였다. 마지막으로, 피드백 유형과 시행 간의 상호작용을 보이는 두뇌 영역을 확인하여 동일한 정서가를 가진 피드백이라도 학습 관여 여부에 따라 달리 관여하는 두뇌 영역이 있는지를 조사하였다.

연구방법

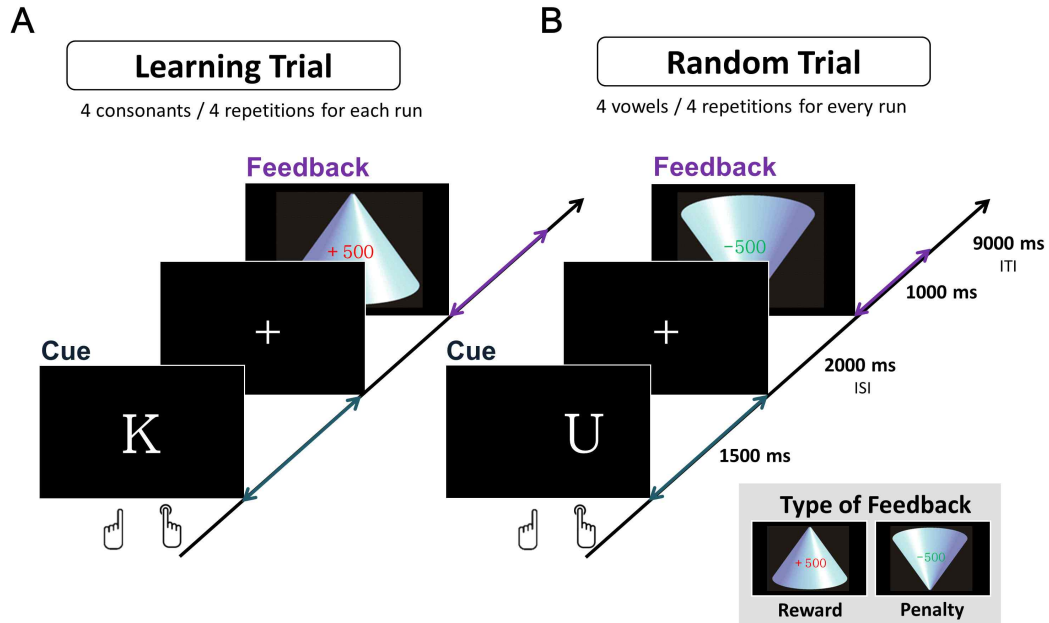
참가자

강원대학교에 재학 중인 학부생 22명(여자 18명/ 남자 4명, 평균연령 22 ± 3.2 세)이 본 연구에 참가하였다. 모든 참가자는 정상 시력을 가진 오른손 사용자로 과거 정신질환, 뇌수술경험 그리고 항정신성 약물 복용 경험이 없었다. 참가자는 과제 수행 전 실험참가 동의서에 서명하였고, 실험 종료 후 소정의 참가비와 과제 수행 중 획득한 보상 금액을 받았다. 본 연구에서 사용되는 임상시험계획서(protocol)는 삼성서울병원의 임상시험심사위원회(IRB)의 심의를 거쳐 승인되었다.

자극 및 행동과제

단서 자극(cue stimulus)으로 로마자 알파벳 16개의 자음과 4개의 모음(A, E, O, U)이 각각 학습 시행과 무선행에 사용되었다. 단서 자극은 16도의 시야각(visual angle)을 지닌 LCD 모니터에서 약 2.2도의 크기로 제시되었다. 본 연구는 과제 수행동안 두 시행 유형을 구별하는 데 참여자의 노력이 최소화되도록 설계되었다. 이를 위해 참여자가 사전에 알고 있는 범주 지식(예, 모음과 자음 구분)을 사용하여 두 시행 유형을 구분하였다. 즉, 모음과 자음이 단서로 제시되었을 때, 참가자가 어느 시행 유형인가를 구분하기 위해 실험 상황에 특정한 기억(가상 실험의 예, K, P, J = 무선행, O, T, A = 학습시행)을 인출하지 않아도 되게 하였다. 학습시행 자체가 실험 장면 특징적 기억 인출(단서-반응 연합 쌍 기억)이 요구되기 때문에 두뇌 처리과정을 영상으로 관찰함에 있어서 시행 간 구분을 위한 추가 인지과정이 요구되는 것을 최소화하였다. 피드백 자극은 원뿔 도형을 사용하였는데, 참가자들이 획득(+500)하거나 잃게 될 금액(-500)이 적힌 원뿔 도형의 제시 방향성을 다르게 하여 보상과 처벌 피드백을 구분하였다.

Figure 1은 fMRI 촬영 동안 참가자가 수행한 두 가지 과제 유형을 보여준다. 학습시행에서는 화면 정중앙에 자음 단서 자극이 제시되었다(Figure 1A). 참가자는 단서 자극이 제시(1.5초)되는 동안 왼손이나 오른손 검지 버튼 중 하나를 누르는 반응을 하도록 지시받았다. 각 자음 단서에 대해 좌 또는 우 반응이 인위적으로 정답 반응으로 할당되었다. 참가자는 반응의 정/오답을 나타내는 피드백 정보를 이용하여 단서에 짝지어진 정답 반응의 연합 쌍을 시행착오를 통해 습득하였다. 무선행의 경우 모음 단서 자극이 화면의 정중앙에서 좌 또는 우측으로 약간 치우치게(모니터의 정중앙에서 1.1도 외측) 제시되었다(Figure 1B). 참가자는 모음 자극의 나타나는 방향의 좌/우 반응 버튼을 누르도록 지시받았다. 참가자의 반응 정확 유무와 상관없이 무선행으로 피드백이 주어졌다. 두 시행에서 피드백은 금전적 획득(+500원)이나 손실(-500원)을 나타내는 시각 자극을 통해 1초 동안 제시되었다.



(Figure 1) Experimental design. For each condition, a cue stimulus was presented for 1.5 s, followed by a feedback stimulus (1 s) with a 2 s ISI (inter-stimulus interval). (A) In learning trials, participants learned by trial and error to make left or right responses to consonant cues, guided by performance-dependent feedback. (B) In random trials, participants were instructed to respond according to the left-right position of the vowel cues, and were previously informed that feedback would be given with random probability. ITI = inter-trial interval. Insert shows two kinds of feedback, indicating monetary reward or penalty (500 KW, ~ .50 USD).

학습시행에서는 run마다 4개의 새로운 자음이 4회 반복 제시되어 한 run에서 학습이 이루어질 수 있게 하였다. 무선행에서는 연습 효과 및 학습 필요성을 배제하기 위하여 4개의 모음 자극이 모든 run에 4번씩 반복 제시되었다. 이미 학습된 단서-반응 연합 쌍에 관한 재인 여부를 측정하기 위해, 후속 run에서 동일한 연합 쌍에 대한 측정 시행(기억시행)이 run마다 4시행씩 포함되었다. 첫 번째 run의 경우 scan 전의 연습과제에서 제시되었던 4개의 연합 쌍이 사용되었으며, 마지막 run에서 학습된 단서의 재인 검사는 측정에서 제외되었다. 위의 기억시행은 소수 시행으로 두뇌 영상 분석에서는 제외되었다.

본 fMRI 연구에서는 모두 4회의 run에 걸쳐 fMRI data가 측정되었다. 각 run은 총 36개의 시행(학습시행 16시행, 무선행 16시행, 기억시행 4시행)으로 이루어져 있고, 3가지 시행이 사건 관련 fMRI 방법(rapid event-related fMRI)을 따라 무작위 순서로 제시되었다. 매 run의 소요시간은 8분 24초이며 총 252개의 부피 영상(volume images)이 획득되었다. 각 시행의 혈액학적 반응 함수

(hemodynamic responses function: HRF)의 추정 효율성을 증가시키기 위해 시행 간 간격(inter-trial interval: ITI)의 최적화(Optseq2, Dale, 1999) 과정을 거쳐 획득된 평균 9초(range=1-27s)의 시행 간 간격을 두었다. IFIS/SA system (Integrated Functional Imaging System, MRI Devices, Inc., Gainesville, FL)과 연동된 E-prime (version 1.1, Psychology Software Tools, Inc., Pittsburgh, PA)을 사용하여 MRI scanner 내부 LCD 모니터에 실험 자극을 제시하고, 참가자의 양손 버튼 반응 자료를 수집하였다.

fMRI 촬영 전, 본 실험과 동일한 시행 수를 가진 연습과제를 통해서 각 시행 유형에서의 반응과 피드백의 유관성의 유무에 대해 사전에 숙지할 수 있게 하였다. 참가자들은 fMRI 회기가 끝난 후 과제 수행 동안 획득한 합계금을 참여에 대한 보수 이외에 추가로 받거나 잃을 수 있음을 사전 지시로 알고 있었다. 실험 종료 후, fMRI 스캐너 밖에서 참가자들에게 학습한 단어 자극에 대한 재인 검사를 실시하였다.

기능영상 획득 및 분석

본 연구의 두뇌 영상은 삼성서울병원의 3.0 T MR 기기(Intera Achieva; Philips Medical Systems, the Netherlands)에서 획득되었다. T2-weighted single-shot EPI 기능영상 획득에 사용된 변수는 다음과 같다: repetition time(TR) = 2000ms, echo time(TE) = 31ms, flip angle = 90°, number of volumes per run = 252 volumes, number of slice = 30, interleaved, 4mm slices with no gap, field of view(FOV) = 22 cm, matrix size = 64 × 64, voxel size = 3.43 × 3.43 × 4 mm³. 영상 절편들은 AC-PC(anterior commissure-posterior commissure)선에 평행하게 위치되게 하였다. 기능영상 촬영 후에 공간 정규화 과정을 위해 고해상도 해부학 영상은 T1-weighted 3D fast-field echo sequence (TR = shortest, TE = 4.60 ms, flip angle = 8°, FOV = 240 × 240 mm², matrix size = 240 × 240 × 180, voxel size = 1.0 × 1.0 × 1.0 mm³)를 사용하여 획득되었다.

모든 통계적 영상분석은 Matlab 7.6(Mathworks, USA)에서 구현한 SPM8(Statistical Parametric Mapping; Wellcome Department of Cognitive Neurology, London, UK)을 사용하여 이루어졌다. 개인의 사전 영상분석(preprocessing)은 부피 영상의 시간 보정(slice timing), 머리 움직임 효과를 보정(realignment) 및 표준 공간(Montreal Neurological Institute template)으로의 이미지에 정합(normalization), 이어서 공간적 편평화(smoothing, full width at half maximum=6mm) 순으로 이루어졌다. 이렇게 사전처리가 된 개인의 영상은 각 피드백 제시 시점의 BOLD 신호 변화를 보기 위하여 시행 유형별로 일반선형모델을 사용하여 분석되었다.

시행 유형은 참가자들이 각 시행에서 받은 보상/처벌 피드백에 따라서 이분되므로 개인 분석에서 최종 6가지 regressors 가 모델링 되었다. 기억시행의 경우 낮은 정확 반응률로 기억-보상 피드백 개수가 적어 추후 집단 분석에서는 제외되었다. 최종 집단 분석에는 4개의 피드백 유형(학

습-보상, 학습-처벌, 무선-보상, 무선-처벌 시행)만이 사용되었다. 이 단계에서 획득된 개인 분석 결과 영상은 참가자를 무선변인으로 하는 변량분석에 사용되었다. 시행요인(2수준: 학습 vs. 무선)과 피드백 정서가 요인(2수준: 보상 vs. 처벌)의 주효과와 두 요인 간 상호작용 효과가 분석되었다.

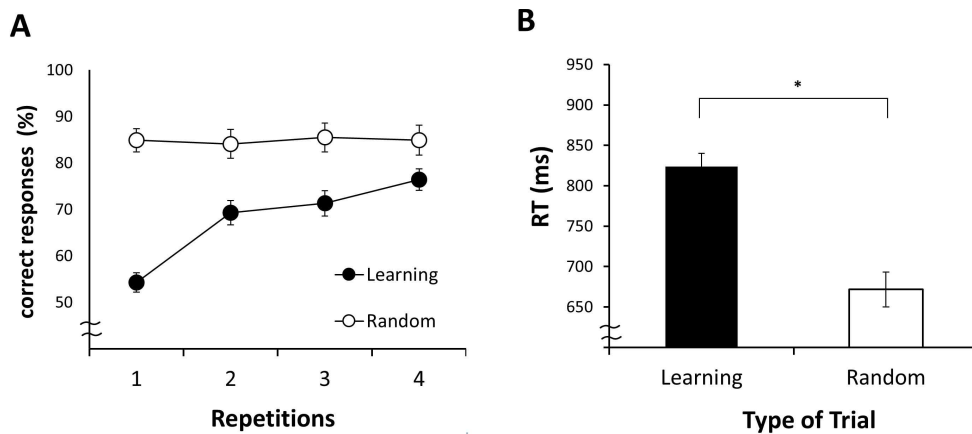
집단 분석의 통계적 유의성은 전체 두뇌 부피소 단위(whole brain voxel-wise)에서 다중비교검정으로 인한 1종 오류 증가를 보정하기 위한 기법으로 사용되는 family-wise error rate $p < .001$ (이하 FWE)이라는 매우 엄격한 통계적 준거를 사용하여 이루어졌다. 이와 더불어 voxel 수준에서 해당 준거를 만족하면서 부피소가 연속적으로 5개 이상으로 이뤄진 덩이소(cluster)만을 유의미한 두뇌 영역으로 가정하고 본 연구 결과에서 보고하였다. 유의한 활성화 차이를 보인 두뇌 영역에 관해 MarsBaR software(<http://marsbar.sourceforge.net>)를 이용하여 관심영역 분석(Region of Interest; ROI analysis)을 하였다. ROI 영역에 대해 피드백 제시와 관련된 평균 BOLD 신호 변화량(mean percent signal change), 자극 제시 시점을 기준으로 BOLD 신호 함수를 추정한 시간-경과 자료(time-series data)를 획득하였다. 변량분석에서 유의한 상호작용을 보이는 두뇌 영역에 관해서는 각 시행 유형의 시간-경과 자료를 산출하여 피드백 관련 두뇌 반응의 단순 주효과 검증(simple main effect analysis)을 실시하였다. 단서 자극의 제시 시점에서부터 이들 영역의 활성화 차이를 나타내는가를 알아보기 위해 BOLD 신호의 시간-경과의 시작점을 단서 제시 시점으로 값을 산출하였다. 피드백 관련 두뇌 반응은 혈액학적 반응 함수(canonical HRF)를 고려하여 피드백 자극 제시 후 5초가 경과된 시점(단서 자극 제시 후 8초, Figure 4에서 보라색 box로 표시)의 BOLD 신호로 가정하였다.

결 과

행동 결과

단순한 시각 판단만 요구되는 무선시행과 달리 학습시행에서는 반복에 따른 정확 반응률(percent of correct response) 증가가 나타났는데, 이는 시각 단서에 대한 좌/우 범주 반응에 이어지는 피드백에 대한 반복 경험으로 시각범주 반응의 학습이 이루어졌음을 시사한다. 이런 차이는 시행 유형(type of trial: 학습, 무선)과 반복 제시 횟수(repetitions: 1st, 2nd, 3rd, 4th) 요인의 유의한 상호작용 효과로 지지가 되었다, $F(3, 63)=14.17$, $p < .0001$, (Figure 2A). 학습이 요구되지 않는 무선시행은 반복 제시와 상관없이 높은 정확 반응률을 보여 평균 정확율의 시행 유형 간 유의미가 나타났다, $F(1, 21)=33.73$, $p < .0001$ (학습시행 $M=67.8\%$, $SD=7.3$; 무선시행 $M=84.9\%$, $SD=12.2$). 반응시간에서 시행 유형의 주효과가 유의하게 나타났는데, 무선시행의 단서 방향판단

에 걸리는 반응시간($M=671.7\text{ms}$, $SD=101.0$)이 학습시행의 반응시간($M=824.9\text{ms}$, $SD=71.6$)보다 빨랐다, $F(1, 21)=66.1$, $p < .0001$, (Figure 2B). 반복 제시 횟수에 따른 주효과도 유의하게 나타났는데, 학습시행에서 학습의 진행으로 반응시간이 점차 빨라지며 또한 무선시행에서도 시행을 반복할수록 반응시간이 감소되었다, $F(3, 63)=3.9$, $p < .05$. 이전 run에서 학습한 단어에 대한 인출을 요구하는 기억시행의 평균 정확 반응률은 58.5% ($SD=10.3$)로 우연 수준(50%)보다는 유의하게 높았다, $t(21)=3.8$, $p < .001$. 이는 마지막 반복 제시의 평균 정확 반응률보다 낮지만 이전 run에서 학습한 단서의 정답 반응에 관한 장기 기억 전환이 이루어졌음을 보여준다.



(Figure 2) Behavioral results. (A) Percent of correct responses for learning trials vs. random trials as a function of repetitions. The percent of correct responses increased from the 1st to the 4th repetition in the learning trials only. (B) Response times (RT) of random and learning trials differs significantly. Error bars represent SEs. * $p < .001$.

fMRI 결과

피드백의 학습 요인

피드백 효과가 시행 유형(학습시행과 무선시행)에 따라서 어떻게 달라지는지에 대하여 알아보기 위해 실시한, 시행과 피드백 유형을 변인으로 한 이원변량분석 결과는 다음과 같다. 보상이나 처벌 사건의 정서가와 상관없이 피드백이 자신의 행동 반응에 대해 연관성을 가지고 제시되는 학습시행과 관련성 없이 제시되는 무선시행 간에 활성화 차이를 보이는 두뇌 영역을 확인하였다(Table 1, Trial Types). 그 결과, 좌측 하두정피질(inferior parietal cortex), 양측 소뇌(cerebellum), 양측 배외측 전전두피질(dorsolateral prefrontal cortex), 양측 전운동피질(premotor cortex), 배내측 전전두피질/배측 전대상피질(dmPFC/dACC), 양측 도/하전두회(insula/inferior frontal gyrus), 우측 미상핵 머리(head of caudate nucleus), 우측 전췌기소엽(precuneus), 그리고 외측 안와전두영역(lateral

〈Table 1〉 Brain regions showing effects of trial and feedback types, and interactions between trial and feedback types

Regions	L/R	BA	MNI coordinate			T-value [†]	Cluster size
			x	y	z		
Trial Types							
Learning > Random							
inferior parietal cortex*	L	40/7	−36	−48	36	10.0	574
cerebellum	R	-	8	−76	−24	8.18	90
	L	-	−8	−80	−30	7.93	95
*	R	-	30	−64	−30	8.13	59
DLPFC	R	46	44	32	28	6.89	119
*	L	46	−42	28	24	6.80	44
premotorcortex	L	6	−40	10	52	7.60	91
	L	9	−46	8	32	6.89	43
	R	8	30	16	48	7.22	20
dmFC/dACC*	R	32	2	22	46	7.84	352
insula/inferiorfrontalgyrus*	R	47	38	20	−2	7.03	20
	L	47	−32	22	−6	6.63	11
internal capsule	L	-	−8	2	2	7.01	10
caudate nucleus (Head)	R	-	8	16	2	6.88	10
precuneus	R	7	10	−62	42	6.72	13
lateral orbitofrontal cortex	R	10	32	50	4	6.58	10
Random > Learning							
fusiform gyrus	L	36	−28	−34	−20	8.16	37
posterior cingulate gyrus	L	-	−6	−52	22	7.89	40
angular gyrus	L	39	−48	−68	20	7.68	57
parahippocampal gyrus	L	28	−20	−14	−22	7.29	32
	R	28	24	−16	−24	6.75	11
superior temporal gyrus	R	22	60	−8	4	6.84	44
precentral gyrus	R	4	12	−36	62	6.80	27
inferior frontal gyrus	L	47	−30	34	−12	6.66	9
Feedback Types							
Reward > Penalty							
cuneus	L	17/18	−12	−94	6	6.97	28
	R	17/18	14	−92	6	6.84	11
Penalty > Reward							
lingual gyrus	L	18	−10	−82	−10	7.48	18
insula/inferior frontal gyrus*	L	47	−42	18	−2	7.17	25
	L	47	−30	24	−8	5.66	5
Interaction between Trial and Feedback types*							
DLPFC	L	48	−40	16	24	58.81	27
dmPFC/dACC	L	32/8	−2	24	42	44.87	24
insula/inferior frontal gyrus	L	47	−30	26	−6	42.71	5

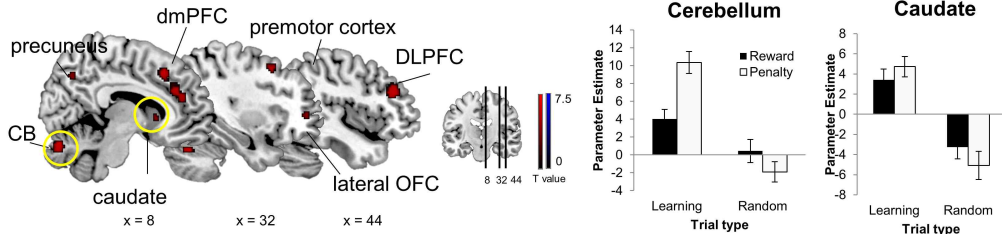
[†] voxel-wise FWE $p < .001$ ($T > 6.18$, $k \geq 5$). * brain regions also showing significant interaction of learning and feedback types.

[‡] F value for 2×2 ANOVA. BA, Brodmann's area; L, left; R, right; DLPFC, dorsolateral prefrontal cortex; dACC, dorsal anterior cingulate cortex.

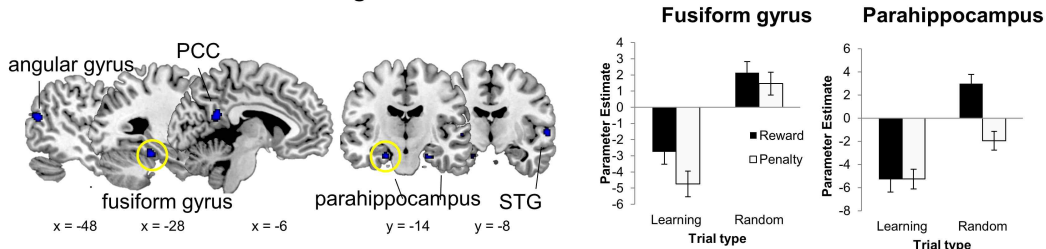
orbitofrontal cortex) 등에서 무선행의 피드백에 비해 학습시행의 피드백에 대해 높은 활성화를 보였다(Figure 3A). 좌측 하두정피질, 좌측 배외측 전전두피질, 배내측 전두피질/배측 전대상피질, 양측 도/하전두회는 유의한 상호작용 효과를 보이는 영역이었다. 양측 소뇌, 우측 배외측 전전두피질, 우측 미상핵 머리, 우측 전좌기소엽, 그리고 중전두회 영역은 학습시행의 피드백에 대해 높은 활성화를 보이는 반면 무선행의 피드백에 관해서는 반응을 보이지 않았다.

반면 학습 피드백에 비해 무선행 피드백에 대해 높은 활성화를 보이는 영역에 대한 분석 결과, 좌측 방추상회(fusiform gyrus)와 양측 해마방회(parahippocampus) 등을 비롯한 여러 두뇌 영역에서 활성화가 관찰되었다(Figure 3B). 이 영역들에 대한 추가 ROI 분석 결과, 무선행에서의 피드백 자극에 높은 반응을 하는 영역이 아니라 학습시행에서 피드백을 받을 때 오히려 비활성화(deactivation)를 보이는 영역임이 밝혀졌다.

A Learning feedbacks > Random feedbacks



B Random feedbacks > Learning feedbacks

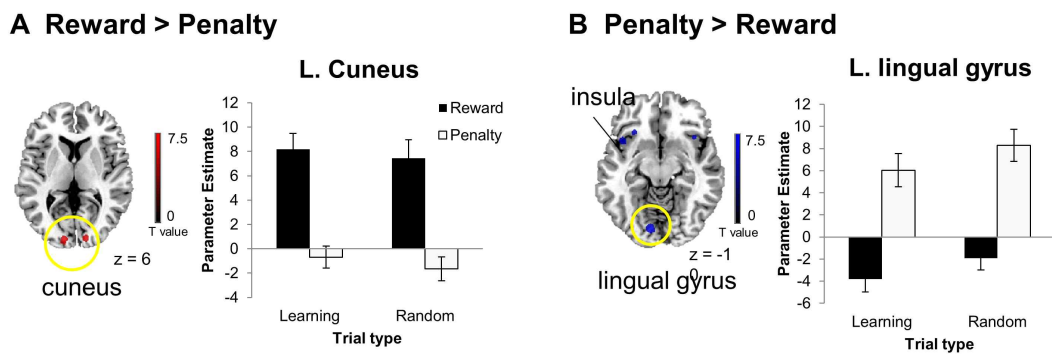


(Figure 3) Learning Effects. (A) Brain regions showing significantly greater activations for learning feedbacks relative to random feedbacks (left panel), as indicated by bar graphs for the cerebellum (CB) and caudate nucleus (right panels). (B) Brain regions showing greater activations for random feedbacks than learning feedbacks (left panel), as indicated by bar graphs for the left fusiform gyrus (right panels) and the parahippocampal gyrus.

피드백의 정서적 요인: 금전적 보상과 처벌에 대한 두뇌 반응

시행 유형과 상관없이 버튼 반응 후에 받은 피드백이 보상인가 또는 처벌인가에 따라 활성화 차이를 보이는 두뇌 영역을 살펴보았을 때(Table 1, Feedback Types), 좌측 설상회(lingual gyrus), 양측 설소엽(cuneus), 양측 도(insula)에서 정서가가 다른 두 피드백(보상 vs. 처벌) 간 활성 차이가 발

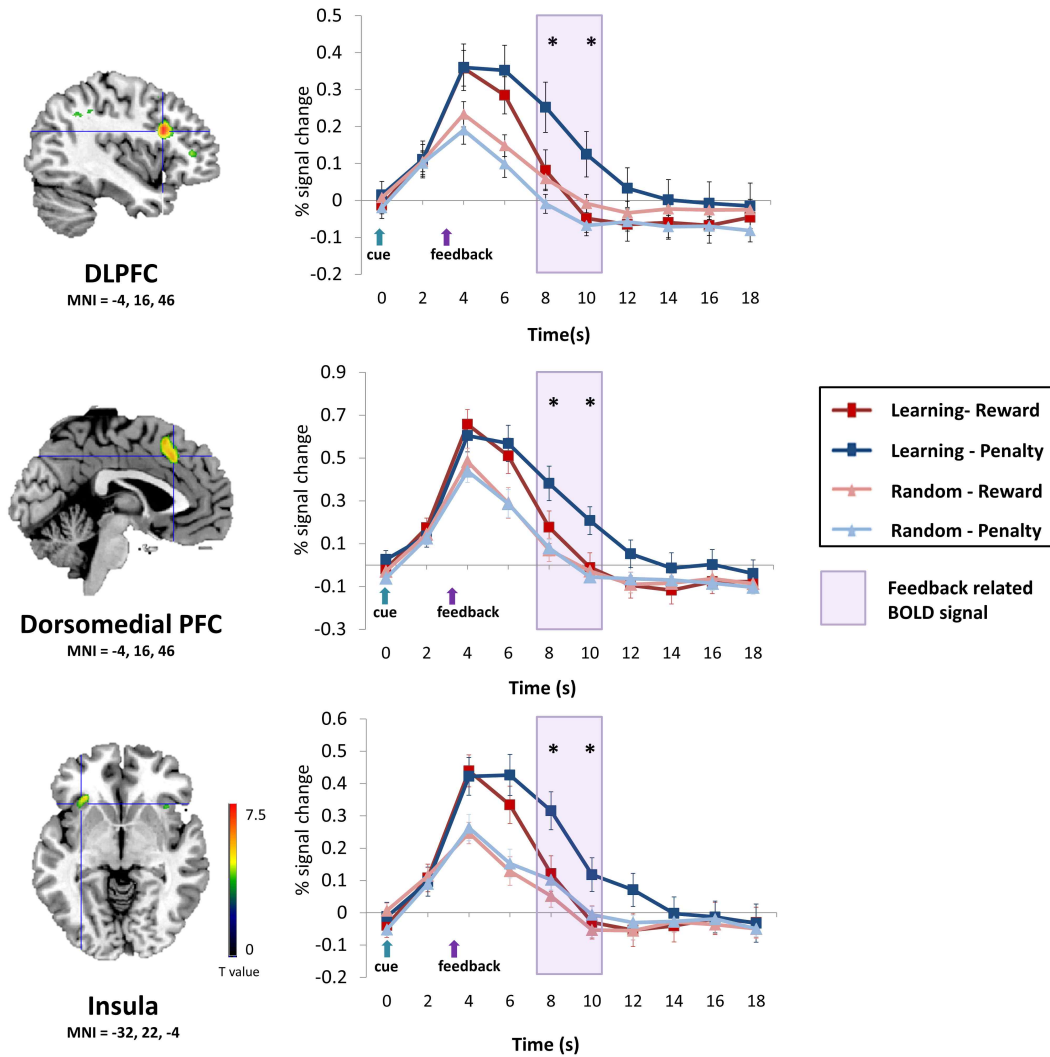
견되었다(Figure 4). 좌측 설상회와 좌측 전측 도는 금전적 처벌에 대하여 높은 활성화를 보였다. 양측 도의 경우 사실상 학습시행의 처벌에 대하여 무선행의 처벌 피드백보다 월등히 더 큰 정도의 활성화를 보여 피드백 정서가에 대한 주효과 뿐 아니라 피드백 정서가와 시행 유형 간의 상호작용을 보임이 추후 분석에서 발견되었다. 반면 양측 설소엽의 경우 처벌 피드백에 비해 보상 피드백에서 증가된 활성이 관찰되었고(Figure 4A), 좌측 설상회는 시행유형과 상관없이 처벌 피드백에 대해 높은 활성화 양상을 보였다(Figure 4B).



(Figure 4) Differences in activations between monetary reward and penalty. (A) Activations were significantly higher for reward in the bilateral cuneus. (B) The right lingual gyrus and left insula showed significantly greater activations for penalty.

시행유형 요인과 피드백유형 요인의 상호작용

좌측 배외측 전전두피질(DLPFC), 배내측 전전두피질(dmPFC), 그리고 좌측 전측 도(anterior insular) 등의 영역에서 피드백에 반응하는 BOLD 신호의 유의한 시행×피드백 상호작용 효과가 관찰되었다(Table 1, Interaction between Trial and Feedback Types, voxel-wise FWE $p < .001$). 이들 ROI 영역의 시간-경과 분석 결과는 Figure 5에 제시되었다. 피드백 관련 BOLD 신호(시간-경과 그래프에서 보라색 상자로 표시)를 이용하여 상호작용 효과에 관한 단순 주효과 분석을 한 결과, 모든 ROI에서 1) 학습-처벌 피드백에 대한 두뇌 활성 정도가 무선행-처벌 피드백에 비해 유의하게 높았으며, $p < .001$, 2) 학습-처벌 피드백의 활성화 정도가 학습-보상 피드백보다 유의하게 높았다, $p < .001$.



(Figure 5) Brain regions showing interaction between trial types (learning vs. random) and feedback types (reward vs. penalty) (left panel). An increase in percent signal change occurred mainly for penalty in the learning trials, as shown by the time-series in the right panel.

논 의

인간을 대상으로 한 피드백 학습 연구에서 자주 사용되는 금전적 피드백은 적절한 행동 선택을 위한 인지적 정보처리 반응뿐만 아니라 정서적 또는 동기적 반응을 일으킬 가능성이 있다. 본 연구의 목적은 피드백의 동기적/정서적 처리과정을 구별하여 학습과 관련된 피드백의 인지적

정보 처리과정의 신경 실질(neural substrates)을 규명하는 데 있다. 금전적 획득 또는 손실이라는 사건과 관련한 두뇌 활동을 사건 관련 fMRI 방략을 이용하여 분석하였는데 모든 분석은 변량분석(ANOVA)의 통계적 논리에 근거하여 이루어졌다. 이를 통해 학습시행과 무선행을 비교하여 학습 관련 특성에 관여하는 영역과 금전적 획득과 손실 간 비교를 통해 피드백의 정서적 특성에 관여하는 영역을 살펴보았다. 특히 학습 관련 피드백 유형 간의 차이를 보이는 영역을 규명하였다.

본 연구의 두뇌 영상 결과를 이해함에 있어서 학습시행과 비교하고자 하는 무선행의 몇 가지 특성을 고려해보면 다음과 같다. 첫째, 무선행에서는 단서 자극의 제시 위치에 따라 좌/우 버튼 반응을 하는 지각판단과제를 실시함으로써 직전 시행의 피드백 정보가 후속 시행의 단서 자극에 대한 버튼 반응에 전혀 영향을 주지 않도록 하였다. 무선행에서 단서에 대한 버튼 반응을 유지하되 참가자의 학습 가능성을 배제하게 하여 감각 및 운동 차원에서 학습시행과 동일하게 하였다. 특히 참가자가 단서 자극을 보는 즉시 지각판단과제를 수행하면 되는 것을 알 수 있도록 단서 자극을 학습시행과 구분되는 4개의 모음만을 사용하였다. 둘째, 과제 수행 전 참가자에게 무선행에서 제시되는 피드백의 무관성을 숙지하게 하여 피드백을 통한 직전 행동에 관한 확인과 행동 수정을 위한 추가 처리과정을 일어나지 않도록 하였다. 셋째, 전체 무선행의 50%에 해당하는 시행에서는 금전적 보상 피드백에 제시되고 나머지 50%의 시행에서는 금전적 처벌 피드백이 따른다는 것을 참가자가 미리 알게 하여 이후 시행의 피드백 예측 과정이 일어나지 않게 하였다. 무선행은 금전적 획득이나 손실이 사전에 예측되거나 관련된 후속 인지 과정이 요구되지 않는다는 점에서 학습시행과 구분된다. 따라서 무선행에서는 동일한 피드백 사건에 정서적 또는 동기적인 반응과 같은, 더 기본적인 정보처리 과정만 관여할 것이라고 가정하였다.

학습 효과를 보이는 두뇌 영역: 미상핵과 내측 측두영역

본 연구의 두 시행 유형을 구분 짓는 특징은 피드백이 행동과 연관성을 가지고 제시되는지 여부이다. 구체적으로, 무선행의 피드백은 단서 자극에 대한 목표 반응의 정확 유무와 무관하게 제시되어 수행에 관한 정보를 담고 있지 않았지만, 학습시행의 피드백은 동일 단서 자극에 대한 목표 반응을 유지·수정하는데 중요한 정보를 포함하고 있었다. 무선행의 피드백보다 학습시행에 제시되는 피드백에 선택적으로 높은 활성화를 보이는 영역으로 다수의 두뇌 영역이 발견되었으나 대부분은 시행 유형 x 피드백 요인(보상 vs. 처벌)의 상호작용 효과에 기인하였다. 미상핵과 소뇌 영역이 학습시행과 무선행 간 차이를 보이는 대표적인 영역으로 발견되었다.

강화학습에서 기저핵(basal ganglion)의 하위 구조물인 미상핵의 역할에 대해서는 많은 연구가 행해져 왔다. 특히 미상핵은 모든 피드백에 반응하기보다는 피드백이 행위의 결과로 제시될 때

에 더 증가된 반응을 보이는데, 이는 미상핵이 행동 유관성 정보처리에 관여함을 나타낸다(Balleine, Delgado, & Hikosaka, 2007; O'Doherty et al., 2004; Seger & Cincotta, 2005; Tricomi et al., 2004). 본 연구의 학습-피드백만이 직전 반응의 정답 여부에 따라 제시되었다는 점에서 예측 가능한 발견이다. 시행 유형 간 차이를 보인 두뇌 영역으로 기저핵의 활성화가 관찰된 것은 운동 관련 학습에 중요한 역할을 보고한 선행연구(Middleton & Strick, 2002; Williams & Eskandar, 2006)와도 일치하는 결과라고 볼 수 있다. 학습-피드백 처리과정 동안 소뇌의 두뇌 반응이 증가됨을 발견하였는데, 두 시행 모두에서 단서에 대한 운동 반응 산출이 요구되었다는 점에서 학습 시행의 소뇌 활성 증가는 기존에 잘 알려진 소뇌의 운동 통제 기능(Lotze et al., 1999; Stoodley, Valera, & Schmahmann, 2012)에서 확장된, 다양한 인지 과정에서의 소뇌 역할에 대한 최근 견해를 지지하는 증거가 될 수 있을 것이다(Peterburs & Desmond, 2016; Rustemeier, Koch, Schwarz, & Bellebaum, 2016).

본 연구는 방추상회(fusiform gyrus), 후대상회(posterior cingulate gyrus), 각회(angular gyrus), 해마방회(parahippocampal gyrus), 상측두이랑(superior temporal gyrus) 등의 다수의 두뇌 영역에서 학습-피드백에 오히려 비활성화(deactivation) 반응을 보임을 발견하였다. 후대상회, 각회, 중심전이랑 등의 영역은 휴식 상태(resting state)동안 뇌 활성 증가가 나타나는 디폴트 모드 네트워크(default mode network; DMN)(Raichle et al., 2001)의 hub로 제안되는 영역들이다. 인지 과제를 이용한 기능영상 연구들은 과제 수행동안 DMN가 기저상태보다 낮은 비활성화를 보임을 관찰하였으며, 그 비활성화의 정도가 과제 난이도/요구와 상관이 있음을 보고하였다(Buckner, Andrews Hanna, & Schacter, 2008; Fox et al., 2005; Mckiernan, Kaufman, Kucera-Thompson, & Binder, 2003; Spreng, 2012). 본 연구에서 무선시행의 피드백 처리 과정에 비해 학습시행의 과제 요구가 높음을 고려하면 학습시행의 피드백 제시 시점에서 여러 DMN 영역의 비활성화 반응증가가 관찰된 것을 같은 현상으로 이해할 수 있을 것이다. 특히 양측 방추상회와 해마방회에서 관찰된 학습-피드백에 대한 비활성화 결과는 본 연구에서 사용된 학습 과제의 특성과 관련하여 논의될 필요가 있다. 이들 두뇌 영역은 내측 측두엽(medial temporal lobe)의 하위구조물로서 절차기억(procedural memory)을 요구하는 과제동안 비활성화를 보이는 것으로 알려진 영역이다(Luking & Barch, 2013; Seger & Cincotta, 2006). 본 연구의 학습시행에서 사용된 시각범주판단 과제는 앞서 언급한 절차적 기억에 의존하는 학습으로 이런 기억 형성에 기저핵(미상핵 및 피각)의 활동이 증가하는 반면, 서술적 기억(episodic memory) 형성에 관여하는 것으로 알려진 내측 측두영역의 활동이 저하되는 경쟁적인 관계를 보인다는 것이 보고된 바 있다(Poldrack et al., 2001; Seger & Cincotta, 2005). 해당 과제 동안 두 영역의 활성화 양상을 살펴보면 두 영역 간에 부적 상관관계가 보고되기도 한다. 본 연구에서 해마방회가 학습 관련 피드백에 높은 정도의 비활성화를 보인 본 연구 결과는 이전 연구 결과와의 맥락에서 서술기억 체계의 억제로 해석할 수 있을 것이다.

금전적 보상과 처벌에 대한 두뇌 반응: 고차 시각 영역, 내측 안와 전두피질과 측좌핵

보상과 처벌 피드백을 비교한 결과에서 일차시각영역의 인접 영역인 양측 설소엽과 좌측 설상회가 발견된 것은 보상과 처벌에서 사용된 원뿔 도형의 방향과 글자색(보상-녹색, 처벌-빨강)의 차이에 의한 시각적 처리과정이 반영된 것으로 볼 수 있다. 유사하게 금전적 피드백의 두 유형(보상/처벌)과 그 크기(magnitude)를 화살표의 방향/색과 크기로 구분하여 사용한 연구(Delgado, Locke, Stenger, & Fiez, 2003)에서도 처벌 피드백에 비해 보상 피드백에 높은 좌측 설소엽 활성화가 발견되었는데, 연구자들은 이를 피드백 시각 자극의 모양과 크기에 대한 시각적 차이의 영향으로 보았다. 하지만 부정적인 그림(Taylor et al., 1998)이나 부적정서 상황에서 학습된 단어(Maratos, Dolan, Morris, Henson, & Rugg, 2001)를 인출하는 동안 설상회의 활성화가 보고된 바 있어, 본 연구에서 발견된 처벌 피드백에 관한 증가한 설상회 활성화가 처벌의 부적의미 정서처리와 관련되었을 가능성 또한 배제할 수 없다.

본 연구에서 탐색적 분석을 위한 낮은 통계 검증 기준(uncorrected $p < .001$, $T > 3.19$)을 적용하였을 때, 보상처리에 관여한다고 알려진 좌측 측좌핵(MNI 좌표 L: -8, 16, -10, $T=4.37$, $k=45$, R: 10, 10, -12, $T=4.24$, $k=18$)과 좌측 내측 안와전두영역(MNI 좌표 L: -4, 56, -4, $T=4.15$, $k=95$)에서 보상 피드백 특정적 반응(보상 > 처벌)이 관찰되었다. 이는 기존 연구 결과(Liu, Hairston, Schrier, & Fan, 2011)와 일치하는 발견이다. 이들 보상 관련 영역에서 보상 피드백에 낮은 활성화가 발견된 것은 학습-보상 피드백이 확률이 아닌 결정적 수반성으로 제시된 것과 관련이 있을 것이다. 본 연구에선 성공적 수행 후에 주어지는 보상 피드백을 100%의 확률로 예측할 수 있어, 기존의 연구와 달리 보상의 유인가 역할보다는 정답 반응에 대한 확인적 역할이 더 크며 보상 예측의 불확실성은 낮았다. 유사한 학습 절차에서 학습 진행에 따라 보상이 가지는 현저성이나 보상에측오류가 감소한다는 것이 이미 보고되었다(Kim, Kim, & Kang, 2015; Tricomi & Fiez, 2012). 본 연구의 두 시행 유형 모두에서 금전적 보상 피드백에 관한 불확실성이 높지 않았고 과제 전략을 수정할 필요가 없었다. 복측 선조체 같은 보상 영역이 보상에 대한 불확실성에 민감하게 반응하고 그에 따른 수행 전략 변경에 관여한다는 과거 연구들(Aron et al., 2004; Buzzell et al., 2016)을 고려할 때 본 연구에서 보상 관련 두뇌 영역에 통계적으로 유의한 활성화(보상 > 처벌)를 발견하기 어려웠을 가능성이 있다.

학습의 처벌 피드백: 배외측 전두영역-하두정영역의 집행회로, 배내측 전두영역, 전측 도

본 연구의 주 관심은 피드백 정서가에 따라 학습시행과 무선행에 다른 특성을 보이는 두뇌 신경 실질의 규명에 있었다. 시행(학습 vs 무선행)과 피드백 유형(보상 vs 처벌)을 요인으로 한 이원변량분석의 상호작용을 살펴본 결과, 좌측 배외측 전전두피질(DLPFC), 배내측 전두피질

(dmPFC/dACC), 그리고 좌측 전측 도(anterior insular)에서 학습-처벌에 특정적 활성화를 보이는 것으로 밝혀졌다. 이는 보상피드백 처리 동안에는 시행유형에 따른 활성화 정도에 차이를 보이는 두뇌 영역이 발견되지 않은 것과는 대조적이었다.

배외측 전전두피질이 작업 기억 및 지속적 주의 같은 집행기능에 중요한 역할을 한다는 것은 잘 알려져 있는데(Curtis & D'Esposito, 2003; Krawczyk, 2002), 이는 피드백 학습에 영향을 주는 중요한 인지 속성 중 하나이다(Peters, Van Duijvenvoorde, Koolschijn, & Crone, 2016; Zanolie, Van Leijenhorst, Rombouts, & Crone, 2008). 예를 들어, Seger(2008)는 시행착오를 통한 피드백학습 동안 피드백이 집행 기능에 관여하는 피질선조회로(corticostriatal loop)를 조절하는 방법으로 학습에 영향을 준다고 제안한 바 있다. 부정 피드백은 전략을 변경하거나 작업기억의 표상을 수정해야 한다는 신호 역할을 하는데, 이 과정이 선조체와, 배외측 전전두피질 그리고 후측 두정엽(posterior parietal lobe) 등의 영역과의 연결인 집행회로(executive loop)를 통해 이루어진다고 보았다. 본 연구 결과에서 배외측 전전두피질과 좌측 하두정엽이 무선피드백에 비해 학습피드백에 높은 활성화(학습>무선)를 보이는 동시에 학습-처벌 피드백에 관해 선택적인 반응(학습-처벌>학습-보상)을 보였다는 점은 학습-처벌 피드백이 집행 회로에 의해 처리되고 있다는 증거라고 볼 수 있다(Collins & Frank, 2012; Seger, 2008).

본 연구에서 처벌 피드백에 대해서만 활성화된 영역으로 전대상회를 포함한 배내측 전전두피질(dmPFC/dACC)이 관찰되었다. 배내측 전두영역이 오류처리에 관여하는 두뇌 영역임(Holroyd et al., 2004; Iannaccone et al., 2015; Steele et al., 2016)을 고려한다면 예측가능한 발견이다. 이들 영역은 본인의 수행에 대한 예측과 다른 피드백에 대해 증가된 반응을 보이는데, 강화학습이론에 따르면 전대상회/배내측 전전두영역의 인지 통제 및 오류처리과정은 중뇌 도파민 신경세포에서 전달되는 보상예측오류 신호에 의해 일어난다고 보았다(Amiez, Joseph, & Procyk, 2006; Holroyd & Coles, 2002; Shackman et al., 2011). 본 연구의 탐색적 분석(voxel-wise FWE $p < .05$)에서 흑질(substantia nigra, MNI 좌표 L: -8, -14, -12, $F=42.61$, $k=31$; R: 8, -14, -1, $F=37.01$, $k=16$) 및 중뇌(midbrain, MNI 좌표 6, -28, -8, $F=37.5$, $k=13$)에서도 전대상회/배내측 전전두영역과 유사한 반응 패턴이 발견되었다는 점에서 이전 연구와 일치된 결과라 볼 수 있다. 본 연구의 무선 시행에 사용된 처벌 피드백의 경우는 자신의 행동과 무관하게 제시된다는 것을 이미 참가자에게 사전에 숙지시켜주었기에 무선-처벌 피드백에 대해서는 오류 예측 신호뿐만 아니라 후속 행동 수정을 위한 오류처리 과정이 요구되지 않기에 전대상회/배내측 전전두영역이 관여하지 않았을 것이다. 이에 비해 학습-처벌 피드백은 반응이 틀렸음을 알려주는 학습 피드백으로 충분히 오류 처리를 야기했을 것으로 보인다.

전측 도는 정서 감정 및 각성 표상에 관여하는 것으로 알려져 있다(Gu et al., 2013; Singer et al., 2009). 특히, 통증, 금전적 손실과 같은 부정적 사건에 대한 전측 도의 높은 활성화는 부정 정서처리와의 관련성을 나타낸다(Kuhnen & Knutson, 2005; Xu et al., 2009). 본 연구에서 두 시행의

금전적 손실 모두에 높은 활성을 보일 것이라는 예상과 달리, 무선-처벌 피드백보다 학습-처벌 피드백에 증가된 반응이 나타났다. 이러한 결과는 개인들이 자신의 행동과 무관한 금전적 손실에 관해서는 낮은 부정적 정서 반응을 보였을 가능성을 나타낸다. 다른 설명은 현저성 네트워크(salient network)의 hub로서의 전측 도의 기능(Menon & Uddin, 2010)에 바탕을 둔다. 전측 도는 외/내부 환경에서 주어진 자극이 후속 정보처리에 얼마나 관련되어 있는지와 같은 현저성(salience)에 민감하게 반응한다고 알려져 있다(Metereau & Dreher, 2013). 특히 오류 신호는 높은 현저성을 가지는 사건이므로 전측 도가 해당 오류 신호를 재빨리 탐지하여 전대상회로 전달함으로써 학습 및 적응적 행동을 매개하는 역할을 하는 것으로 제안된다(Bastin et al., 2016; Ham, Leff, de Boissezon, Joffe, & Sharp, 2013). 학습-처벌 피드백에 선택적으로 높은 반응을 보인 영역으로 전측 도와 배내측 전전두엽 모두가 본 연구와 관찰된 것도 선행 연구와 맥을 같이 한다. 즉, 금전적 손실이라는 동일한 사건일지라도, 해당 사건이 후속 행동 수정을 위한 정보를 담고 있는 학습시행의 처벌인 경우만 높은 현저성을 가진 사건으로 처리되었을 것이다. 이런 현저성 처리과정을 통해 배내측 전전두엽과 배외측 전전두엽에 의한 효율적인 오류처리 및 집행 기능 정보처리가 이루어졌을 것으로 생각된다.

본 연구는 피드백의 동기적/정서적 요인과 인지적 요인을 구분할 수 있는 실험패러다임을 통해 학습상황에서의 피드백이라도 보상과 처벌에 따라 어떻게 다른 처리가 일어나는지를 구별하고자 하였다. 본 연구의 발견을 요약하자면 다음과 같다. 첫째, 학습-보상 피드백과 무선-보상 피드백에 대해 변별적인 반응을 보이는 두뇌영역이 발견되지 않았다. 이는 학습이 이미 형성된 “단서 자극-반응”에 대한 보상 피드백은 행동 강화를 위한 추가적인 정보를 제공한다기보다는 보상의 긍정적인 정서처리과정만 일으켰을 가능성을 의미한다. 앞서 학습이 진행됨에 따라 학습시행의 보상은 이미 습득된 연합기억을 확인하고 성공적 수행에 대한 보상의 의미를 지니게 되었을 가능성을 논의한 바 있다. 따라서 보상 피드백에 대해 긍정적 정서를 일으킴으로써 적절한 행동을 유지하게 하는 역할을 하였을 것이다. 둘째, 보상과는 반대로 처벌은 무선적으로 제시되는 경우에 비해 학습이 요구되는 시행에서 제시된 경우 다양한 두뇌 신경망의 활성화를 야기하는 것으로 나타났다. 처벌 피드백이 적절한 행동수정을 위한 추가적인 정보처리과정을 유발했을 것이다. 학습상황의 처벌은 학습이 진행되는 동안 반응의 오류를 지시하는 학습 효과를 유지하며, 기존 전략의 변경, 재평가를 필요로 하는 과정을 야기할 수 있을 것이다. 이는 이후 행동선택의 방략수정을 가능하게 하는 정보를 가지게 되었을 것이고 이런 정보처리 과정에 관여하는 두뇌 영역이 전측 도-배내측 전전두엽/전대상회-배외측 전전두피질-하두정피질로 구성된 신경망을 보여주고 있다.

본 연구 결과 해석에 있어 몇 가지 방법론적 제한점을 논의할 필요가 있다. 먼저, 두 시행의 단서 자극은 세부적인 시각 속성(자음과 모음의 차이, 화면상 제시 위치)에서 차이가 있었다. 자

음과 모음이 구분되는 두뇌 영역에서 처리됨을 보인 과거 연구들(Carreiras & Price, 2008; Obleser, Leaver, VanMeter, & Rauschecker, 2010)과 달리, 본 연구는 음소나 의미 판단을 요구하지 않았다. 따라서, 단서로 주어진 모음과 자음의 비모국어 철자의 정보처리의 차이가 본 피드백 분석 결과에 영향을 미칠 만큼 크지 않았을 것이라 가정하였다. 오히려 방법론적으로 고려될 필요가 있는 절차 차이는 학습시행과 달리, 최소한의 지각적 판단 정보처리를 요구하기 위하여 무선행에서 단서가 정중앙을 약간 벗어난 시야장에서 제시되었다는 점이다. 무선행에서 자극이 제시된 시야장의 위치 차이가 초기 시각정보처리에 관여하는 두뇌 영역에 활성화 패턴의 위치 차이를 야기할 것으로 가정할 수 있다. 그러나 본 연구 기법의 해상도에서는 유의미한 결과로 관찰되지 않았다. 둘째, 단서와 피드백 자극 간의 제시 시간을 일정하게 유지하였기 때문에 중첩(overlap)된 단서 관련 BOLD 신호와 피드백 관련 BOLD 신호 간 분리(dissociation)가 어렵다. 학습시행의 단서 처리과정 동안 학습한 연합 쌍 정보를 기억에서 인출하는 과정이 요구되며, 학습을 위해 단서에 더 많은 양의 주의가 할당되었을 것이다. 이런 단서 처리과정에서 두뇌 활성 차이가 피드백 관련 활성 차이에 영향을 주었을 가능성을 배제할 수 없다. 하지만 과거 메타분석을 이용한 선행 연구(Liu et al., 2011)에 의하면 보상 두뇌 신경망에 포함된 다수의 영역이 정도 차이를 보이지만 피드백의 예측(anticipation), 결과(outcome), 평가(evaluation) 과정에 모두 관여한다. 본 결과에서도 상호작용 효과를 보인 배외측 전두영역-하두정영역의 집행회로, 배내측 전두영역 그리고 전측도의 시간-경과 그래프를 살펴봤을 때 단서 제시 관련 BOLD 신호(단서 자극 제시 후 5초)에서 시행 간의 차이가 관찰되었다. 이들 영역은 무선행에 비해 학습시행의 단서 자극에 대해 높은 활성을 보였다. 동시에, 학습시행의 처벌 피드백이 제시되었던 시행에 선택적으로 피드백 관련 BOLD 신호가 이후에 더 증가하고 있는 것이 관찰된다. 본 연구에서 단서와 피드백에 각각 시간 차이를 두고 나타나는 BOLD 신호 크기를 고려했을 때 피드백 관련 두뇌 활성 신호를 탐지(detection)하고 그 신경 실질 규명을 목적으로 한 본 연구 결과 해석에 해당 한계점이 문제가 되지 않음을 알 수 있다. 셋째, 본 연구에서 학습 피드백로서 금전적 처벌의 정보 처리에 관여하는 두뇌 영역을 규명하였음에도 처벌 피드백이 특정 단서의 학습 수행 정도에 미치는 영향을 알아보지 못하였다. 이는 해당 단서에 대하여 목표 반응을 습득해가는 학습 동안 보상과 처벌 피드백을 받는 경험의 축적으로 행동이 나타나기 때문이다. 후속 연구에서 특정 단서에 대해 한 유형의 피드백 경험(보상 또는 처벌)만 짝지어지게 하는 방법론적인 절차(Bischoff-Grethe et al., 2009)를 사용하여 이를 구분할 수 있을 것이다. 마지막으로 본 연구의 학습-처벌 피드백 처리에 관여하는 양측 도-배측 전두/전대상회-좌측 배외측 전전두피질-좌측 하두정피질 신경망의 상호관련성에 관한 해석은 선행 연구에 기반을 두고 있다. 본 연구의 발견을 바탕으로 DCM(dynamic causal model), SEM(structural equation model)과 같은 네트워크 분석을 적용하여 영역 간 인과적 관련성을 밝히는 후속 연구가 요구된다.

참고문헌

- Amiez, C., Joseph, J. -P., & Procyk, E. (2006). Reward encoding in the monkey anterior cingulate cortex. *Cerebral Cortex*, 16, 1040-1055.
- Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., Gabrieli, J. D., & Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, 6, 196-202.
- Aron, A. R., Shohamy, D., Clark, J., Myers, C., Gluck, M. A., & Poldrack, R. A. (2004). Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *Journal of Neurophysiology*, 92, 1144-1152.
- Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *The Journal of Neuroscience*, 27, 8161-8165.
- Bastin, J., Deman, P., David, O., Gueguen, M., Benis, D., Minotti, L., Hoffman, D., Combrisson, E., Kujala, J., & Perrone-Bertolotti, M. (2016). Direct recordings from human anterior insula reveal its leading role within the error-monitoring network. *Cerebral Cortex*, bhv352.
- Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience?. *Brain Research Reviews*, 28, 309-369.
- Bischoff-Grethe, A., Hazeltine, E., Bergren, L., Ivry, R. B., & Grafton, S. T. (2009). The influence of feedback valence in associative learning. *Neuroimage*, 44, 243-251.
- Bray, S., & O'Doherty, J. (2007). Neural coding of reward-prediction error signals during classical conditioning with attractive faces. *Journal of Neurophysiology*, 97, 3036-3045.
- Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, 30, 619-639.
- Buckner, R. L., Andrews Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences*, 1124, 1-38.
- Buzzell, G. A., Roberts, D. M., Fedota, J. R., Thompson, J. C., Parasuraman, R., & McDonald, C. G. (2016). Uncertainty-dependent activity within the ventral striatum predicts task-related changes in response strategy. *Cognitive, Affective & Behavioral Neuroscience*, 16, 219-233.
- Carreiras, M., & Price, C. J. (2008). Brain activation for consonants and vowels. *Cerebral Cortex*, 18, 1727-1735.
- Collins, A. G., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience*, 35, 1024-1035.

- Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, 7, 415-423.
- Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, 8, 109-114.
- Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *The Journal of Neuroscience*, 25, 11730-11737.
- Delgado, M. R., Locke, H. M., Stenger, V. A., & Fiez, J. A. (2003). Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cognitive, Affective & Behavioral Neuroscience*, 3, 27-38.
- Delgado, M. R., Miller, M. M., Inati, S., & Phelps, E. A. (2005). An fMRI study of reward-related probability learning. *Neuroimage*, 24, 862-873.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, 102, 9673-9678.
- Gu, X., Hof, P. R., Friston, K. J., & Fan, J. (2013). Anterior insular cortex and emotional awareness. *The Journal of comparative neurology*, 521, 3371-3388.
- Ham, T., Leff, A., de Boissezon, X., Joffe, A., & Sharp, D. J. (2013). Cognitive control and the salience network: an investigation of error processing and effective connectivity. *The Journal of Neuroscience*, 33, 7091-7098.
- Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*, 324, 646-648.
- Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109, 679-709.
- Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G., & Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, 7, 497-498.
- Iannaccone, R., Hauser, T. U., Staempfli, P., Walitza, S., Brandeis, D., & Brem, S. (2015). Conflict monitoring and error processing: New insights from simultaneous EEG - fMRI. *Neuroimage*, 105, 395-407.
- Kim, H., Shimojo, S., & O'Doherty, J. P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biology*, 4, e233.
- Kim, S., Kim, J., & Kang, E. (2015). Dynamic changes in feedback processing as learning progresses. *The*

- Korean Journal of Cognitive and Biological Psychology*, 27, 419-450.
- Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of Neuroscience*, 21, RC159.
- Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., & Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *Neuroimage*, 18, 263-272.
- Krawczyk, D. C. (2002). Contributions of the prefrontal cortex to the neural basis of human decision making. *Neuroscience and Biobehavioral Reviews*, 26, 631-664.
- Kuhnen, C. M., & Knutson, B. (2005). The neural basis of financial risk taking. *Neuron*, 47, 763-770.
- Liu, X., Hairston, J., Schrier, M., & Fan, J. (2011). Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, 35, 1219-1236.
- Lotze, M., Montoya, P., Erb, M., Hülsmann, E., Flor, H., Klose, U., Birbaumer, N., & Grodd, W. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *Journal of Cognitive Neuroscience*, 11, 491-501.
- Luking, K. R., & Barch, D. M. (2013). Candy and the brain: neural response to candy gains and losses. *Cognitive, Affective & Behavioral Neuroscience*, 13, 437-451.
- Maratos, E. J., Dolan, R. J., Morris, J. S., Henson, R. N., & Rugg, M. D. (2001). Neural activity associated with episodic memory for emotional context. *Neuropsychologia*, 39, 910-920.
- Mckiernan, K. A., Kaufman, J. N., Kucera-Thompson, J., & Binder, J. R. (2003). A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *Journal of Cognitive Neuroscience*, 15, 394-408.
- Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure & Function*, 214, 655-667.
- Metereau, E., & Dreher, J. -C. (2013). Cerebral correlates of salient prediction error for different rewards and punishments. *Cerebral Cortex*, 23, 477-487.
- Middleton, F. A., & Strick, P. L. (2002). Basal-ganglia 'projections' to the prefrontal cortex of the primate. *Cerebral Cortex*, 12, 926-935.
- Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience*, 16, 1936-1947.
- O'Doherty, J., Critchley, H., Deichmann, R., & Dolan, R. J. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *The Journal of Neuroscience*, 23, 7931-7939.

- O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, 304, 452-454.
- O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, 4, 95-102.
- Obleser, J., Leaver, A., VanMeter, J., & Rauschecker, J. (2010). Segregation of Vowels and Consonants in Human Auditory Cortex: Evidence for Distributed Hierarchical Organization. *Frontiers in Psychology*, 1, 232.
- Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., & Frith, C. D. (2007). How the brain translates money into force: a neuroimaging study of subliminal motivation. *Science*, 316, 904-906.
- Peterburs, J., & Desmond, J. E. (2016). The role of the human cerebellum in performance monitoring. *Current Opinion in Neurobiology*, 40, 38-44.
- Peters, S., Van Duijvenvoorde, A. C. K., Koolschijn, P. C. M. P., & Crone, E. A. (2016). Longitudinal development of frontoparietal activity during feedback learning: Contributions of age, performance, working memory and cortical thickness. *Developmental Cognitive Neuroscience*, 19, 211-222.
- Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature*, 414, 546-550.
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, 98, 676-682.
- Ressler, N. (2004). Rewards and punishments, goal-directed behavior and consciousness. *Neuroscience and Biobehavioral Reviews*, 28, 27-39.
- Rustemeier, M., Koch, B., Schwarz, M., & Bellebaum, C. (2016). Processing of positive and negative feedback in patients with cerebellar lesions. *Cerebellum*, 15, 425-438.
- Salimpoor, V. N., van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., & Zatorre, R. J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science*, 340, 216-219.
- Samanez-Larkin, G. R., Hollon, N. G., Carstensen, L. L., & Knutson, B. (2008). Individual differences in insular sensitivity during loss anticipation predict avoidance learning. *Psychological Science*, 19, 320-323.
- Schultz, W. (1997). Dopamine neurons and their role in reward mechanisms. *Current Opinion in Neurobiology*, 7, 191-197.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275, 1593-1599.
- Seger, C. A. (2008). How do the basal ganglia contribute to categorization? Their roles in generalization,

- response selection, and learning via feedback. *Neuroscience and Biobehavioral Reviews*, 32, 265-278.
- Seger, C. A., & Cincotta, C. M. (2005). The roles of the caudate nucleus in human classification learning. *The Journal of Neuroscience*, 25, 2941-2951.
- Seger, C. A., & Cincotta, C. M. (2006). Dynamics of frontal, striatal, and hippocampal systems during rule learning. *Cerebral Cortex*, 16, 1546-1555.
- Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, 12, 154-167.
- Singer, T., Critchley, H. D., & Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences*, 13, 334-340.
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York, NY: D. Appleton-Century Company.
- Spreng, R. N. (2012). The fallacy of a “task-negative” network. *Frontiers in Psychology*, 3, 145.
- Steele, V. R., Anderson, N. E., Claus, E. D., Bernat, E. M., Rao, V., Assaf, M., Pearlson, G. D., Calhoun, V. D., & Kiehl, K. A. (2016). Neuroimaging measures of error-processing: Extracting reliable signals from event-related potentials and functional magnetic resonance imaging. *Neuroimage*, 132, 247-260.
- Stephani, C., Fernandez-Baca Vaca, G., Maciunas, R., Koubeissi, M., & Lüders, H. O. (2011). Functional neuroanatomy of the insular lobe. *Brain Structure & Function*, 216, 137-149.
- Stoodley, C. J., Valera, E. M., & Schmahmann, J. D. (2012). Functional topography of the cerebellum for motor and cognitive tasks: an fMRI study. *Neuroimage*, 59, 1560-1570.
- Taylor, S. F., Liberzon, I., Fig, L. M., Decker, L. R., Minoshima, S., & Koeppe, R. A. (1998). The effect of emotional content on visual recognition memory: a PET activation study. *Neuroimage*, 8, 188-197.
- Tricomi, E. M., Delgado, M. R., & Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron*, 41, 281-292.
- Tricomi, E. M., & Fiez, J. A. (2012). Information content and reward processing in the human striatum during performance of a declarative memory task. *Cognitive, Affective & Behavioral Neuroscience*, 12, 361-372.
- Ullsperger, M., & von Cramon, D. Y. (2003). Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *The Journal of Neuroscience*, 23, 4308-4314.
- Wheeler, E. Z., & Fellows, L. K. (2008). The human ventromedial frontal lobe is critical for learning from negative feedback. *Brain*, 131, 1323-1331.

- Williams, Z. M., & Eskandar, E. N. (2006). Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nature Neuroscience*, 9, 562-568.
- Xu, L., Liang, Z. -Y., Wang, K., Li, S., & Jiang, T. (2009). Neural mechanism of intertemporal choice: from discounting future gains to future losses. *Brain Research*, 1261, 65-74.
- Zanolie, K., Van Leijenhorst, L., Rombouts, S. A., & Crone, E. A. (2008). Separable neural mechanisms contribute to feedback processing in a rule-learning task. *Neuropsychologia*, 46, 117-126.

1차원고접수 : 2016. 12. 26
1차심사완료 : 2017. 01. 17
2차원고접수 : 2017. 02. 25
최종게재승인 : 2017. 03. 13

(Abstract)

Learning-associated Reward and Penalty in Feedback Learning: an fMRI activation study

Jinhee Kim

Eunjoo Kang

Department of Psychology, Kangwon National University

Rewards or penalties become informative only when contingent on an immediately preceding response. Our goal was to determine if the brain responds differently to motivational events depending on whether they provide feedback with the contingencies effective for learning. Event-related fMRI data were obtained from 22 volunteers performing a visuomotor categorical task. In learning-condition trials, participants learned by trial and error to make left or right responses to letter cues (16 consonants). Monetary rewards (+500) or penalties (-500) were given as feedback (learning feedback). In random-condition trials, cues (4 vowels) appeared right or left of the display center, and participants were instructed to respond with the appropriate hand. However, rewards or penalties (random feedback) were given randomly (50/50%) regardless of the correctness of response. Feedback-associated BOLD responses were analyzed with ANOVA [trial type (learning vs. random) x feedback type (reward vs. penalty)] using SPM8 (voxel-wise FWE $p < .001$). The right caudate nucleus and right cerebellum showed activation, whereas the left parahippocampus and other regions as the default mode network showed deactivation, both greater for learning trials than random trials. Activations associated with reward feedback did not differ between the two trial types for any brain region. For penalty, both learning-penalty and random-penalty enhanced activity in the left insular cortex, but not the right. The left insula, however, as well as the left dorsolateral prefrontal cortex and dorsomedial prefrontal cortex/dorsal anterior cingulate cortex, showed much greater responses for learning-penalty than for random-penalty. These findings suggest that learning-penalty plays a critical role in learning, unlike rewards or random-penalty, probably not only due to its evoking of aversive emotional responses, but also because of error-detection processing, either of which might lead to changes in planning or strategy.

Key words : reinforcement learning, negative feedback, error processing, anterior insula, prefrontal cortex