Abstract of thesis entitled:

A social reinforcement learning mechanism underlying conformity and its effect on outcome evaluation

Submitted by WANG, Yang

for the degree of Doctor of Philosophy in Psychology

at the Chinese University of Hong Kong in April 2016

Conformity refers to changing one's behavior to match with the others, which has been considered as the basis for communication, coordination, and even culture. Previous studies about conformity mainly focused on examining different modulators of conformity and the cognitive process underlying it. However, little is known about how conformity is learnt. To fill this gap, I focused on the social and non-social reinforcement learning mechanisms underlying conformity.

In study 1, I tested whether reward or punishment is the stronger motivator underlying learning, while reward and punishment can be either social or non-social. A series of probabilistic selection tasks (PS task) was adopted. Traditional PS task includes a training and a testing phase. During training phase three different stimulus pairs (AB, CD, EF) are presented randomly, and participants have to choose one from each pair. Feedback of whether it is a correct choice is then provided. Choosing A leads to correct feedback in 80% of AB trials, while B is incorrect in those trials. Similarly, C is correct in 70% of CD trials, and E is correct in 60% of EF trials. Subsequent testing phase includes novel combinations of all the training stimuli (e.g.,

AE, BD, etc). In all the novel pairs including A or B, their performance of approaching A and avoiding B represents reward learning and punishment learning, respectively. On the other hand, the performance of selecting between two stimuli with similar reinforcement history (e.g., A-C: 80%-70%) was expected to be worse than pairs with apparently different reinforcement history (e.g., A-E: 80% -30%), which refers to the conflict monitoring effect.

Experiment 1a serves as a baseline experiment by using the original PS task with non-social feedback (incorrect/ correct) provided. It was used to exclude the possibility that any preference in either approaching A or avoiding B, if there was any, in later social reinforcement learning situation was due to the general learning bias. The results demonstrated that there was no bias in either approaching A or avoiding B in non-social reinforcement learning. Only conflict monitoring effect was found. Experiment 1b used a modified PS task with social reward and punishment (being congruent/ incongruent with another person) as feedback and an antique value evaluation as the cover story. The results showed that participants were more susceptible to social punishment. They performed better in avoiding social punishment tagged antique than approaching social reward tagged one. Experiment 1c used a gambling game as the cover story and Japanese characters as materials. Results showed that even upon knowing others' information did not have any value, participants still performed better in avoiding social punishment. A conflict monitoring effect was also found.

In study 2, I answered the following question: is our susceptibility to social

reward and punishment in social reinforcement learning related to conformity behavior measured in a separate task? Random shaped polygons with six different colors were used as materials, and social reinforcement ratio was tagged on the colors. Due to the task difficulty, experiment 2a did not find any effect when only different colored board lines were used to form polygons. In experiment 2b, by filling the polygons with different colors, participants showed a negative social learning bias with and without self-preference controlled. More importantly, their performance in avoiding social punishment was correlated with normative conformity after self-preference was controlled. The results suggested that normative conformity was related to our susceptibility to the basic social punishment signal during social reinforcement learning.

Utilizing event-related potential (ERP) technique, Study 3 investigated how and when the conforming or non-conforming behavior might affect outcome evaluation process. In a perceptual judgment task, participants were given the choice between conforming and not conforming with incongruent group opinion followed by an outcome (correct/ incorrect). The FRN effect (incorrect minus correct), which reflected error monitoring, following the non-conforming decision was stronger than the effect following the conforming decision. It suggested that conformity influenced our outcome evaluation process.

In conclusion, participants were more susceptible to social punishment than social reward during social reinforcement learning. Their susceptibility to social punishment in implicit social reinforcement learning was related to normative

conformity. Conformity and non-conformity behavior affected outcome evaluation process.

摘要

改變自己的行為去和他人一致叫做從眾。從眾被認為是交流，合作甚至是文化的基礎。以前關於從眾的研究著重在檢驗從眾的不同影響因子以及認知加工過程。但是人們很少研究從眾是如何學來的。因此，本論文關注從眾的社會和非社會性強化學習機制。

研究一檢驗了獎賞和懲罰哪一個是我們社會和非社會學習的更強的動力。一系列概率選擇任務 （PS 任務）被當前研究所採用。 傳統的 PS 任務包括一個訓練和一個測驗階段。 在訓練階段中，三隊刺激（AB, CD, EF）會隨機呈現，被試需要進行從兩個刺激中選擇一個。接下來，選擇是否正確的反饋會提供給被試。在 AB 中選擇 A 有 80％可能性是正確的，而 B 在這些試次中則會是錯的。CD 中選擇 C 會有 70％正確率，EF 中選擇 E 會有 60％正確率。 接下來的測驗階段會包含所有刺激的兩兩組合。 所有包含 A 的新刺激組合中選擇 A 以及所有含有 B 的新刺激組合中不選 B 分別代表了獎勵和懲罰學習。 另外一方面，從擁有相似強化歷史的兩個刺激 (例如:A-C: 80%-70%)中進行選擇的正確率會低於擁有不同強化歷史的刺激組合（例如: A-E: 80%-30%），這被稱之為衝突監督效應。

實驗 1a 是基線實驗，使用 PS 任務以及非社會刺激（正確/錯誤）。實驗 1a 是為了保證如果之後在社會學習情況下發現任何對於學習偏好並不是因為我們基本的偏好所導致的。結果顯示， 在非社會強化學習中，被試沒有偏愛趨近獎勵或者逃避錯誤。只有衝突監督效應存在。 實驗 1b 使用 PS 任務的變式，用古董評估作為故事框架並且提供社會獎賞和懲罰的反饋（和另外一個人是否一

致）。 結果發現被試對社會懲罰更加敏感。他們在避免選擇大家都不喜歡的古董上比在選擇大家都喜歡的古董上表現要好。實驗 1c 使用賭博作為故事框架以及日本字作為實驗材料。結果顯示，即使被試知道他人的信息無用，仍然在避免社會懲罰上表現的更好。衝突監督效應也有出現。

實驗二回答了問題：我們在學習過程中對於社會獎賞和懲罰的敏感性是否和從眾行為有關？六種顏色的隨機形狀的幾何圖形被用為實驗材料，社會強化概率只和顏色有關與形狀無關。實驗 2a 使用不同顏色的線條圍成的圖形作為材料，由於任務過難，並沒有發現任何效應。實驗 2b 使用充滿不同顏色的幾何圖形作為實驗材料，結果顯示了負性社會學習偏好。 更重要的是，他們避免選擇由大家都不喜歡的顏色所組成的圖案的表現，在去除了自我對顏色的偏好效應之後與社會規範性從眾相關。結果顯示了社會規範性從眾與我們在社會學習過程中對社會懲罰的敏感性有關。

使用腦電技術（ERP），實驗三研究了從眾和不從眾如何影響我們對於行為結果的評估。使用知覺判斷任務，被試在看到自己與群體意見不同時可以選擇從眾或者不從眾，接著會出現結果反饋（正確/不正確）。 FRN 效應（對和錯之間的差異波）反應了對錯誤監控的過程。不從眾決策之後的結果相對於從眾決策之後的結果引發了較大的 FRN 效應。這表明從眾影響了結果評估過程。

綜上，在社會強化學習過程中，被試對社會懲罰比對社會獎賞更敏感。內隱社會強化學習中對於社會懲罰的敏感性和社會規範性從眾有關。另外，從眾或者不從眾的行為會影響我們對於結果的評估。