Neural and Computational Models of Gratitude and Pride

Ke Ding

München 2018

Neural and Computational Models of Gratitude and Pride

Inaugural-Dissertation

zur Erlangung des

Doktorgrades der Philosophie an der

Ludwig-Maximilians-Universität

München

vorgelegt von

Ke Ding

aus Henan, China

München, Oktober, 2018

Referent:	Dr. Klaus Wunderlich
Koreferent:	Prof. Dr. Paul Sauseng
Tag der mündlichen Prüfung:	11. 12. 2018

Acknowledgements

First of all, I would like to give thanks to God, who grants me the opportunity and strength to pursue my passion in science. Besides, my doctorate studies would not be possible without the help from many people.

In particular, I would also like to express my deepest gratitude to my supervisor, Dr. Klaus Wunderlich, my “Doctor Father”, who has always been very understanding and supportive whenever I have difficulties. He is the most talented scientist I have ever seen; and he always has the best advice for me. Even though we do not talk too much, I have learned a lot from him, from scientific insight to details in presentation and writing. He gave me all the freedom to do the studies I wanted to do and supported me along the way. In addition, he has encouraged and supported me to go to conferences to seek further training in scientific research as a postdoc researcher, which helps me establish many good connections with top scientists in the field.

Furthermore, I would like to thank my second supervisor Prof. Dr. Paul Sauseng, who has great personality and expertise in scientific research. He has also given me a lot of valuable and sincere advice in scientific career and life. I feel as if I worked with him for many years. I am so lucky to have him as my second supervisor and life tutor.

Besides, I also want to thank Prof. Dr. Virginia Flanagin, who was my fMRI tutor at the beginning of my doctorate studies. She always kindly answered all my questions and helped me with technical problems when I set the experiment up. Meanwhile, she is also my role model in scientific career. I have learned from her that a woman can be a good scientist as well as a good wife, being professional as well as interesting at the same time.

I would also like to thank my colleagues Dian Anggraini and Yan Feng, who have helped me with my experiment: Dian acting as the confederate in my experiment and Yan helping me record the videos. Besides, they also accompanied me when I conducted my fMRI experiment. Without them, my experiment would not work so well. We were a great team and I am so proud to have such nice colleagues and friends.

Apart from my colleagues in the psychology department, I would also like to thank the cool neuroscientists from the department of neurology, Neurologisches Forschungshaus of LMU. In particular, I would like to thank Prof. Dr. Virginia L. Flanagin, Prof. Dr. Thomas Stephan, and Prof. Dr. Peter zu Eulenburg who answered all my questions and helped me generously in their busy schedules. Besides, I would also like to thank Dr. Rainer Bögle, who taught me how to use ICA to reanalyze my fMRI data step by step. Dr. Matthias Ertl and him also participated in my pilot studies and gave me very practical advice for my experiments. I also need to thank Mr. Mathias Hübner, who helped arrange the scanning sessions, borrow the extra equipment, and assisted me in my fMRI experiment.

Personally, I would like to thank my family and friends who always give me love, strength, and support. Especially, it is really rare and valuable that my parents give me the freedom to pursue my dreams and make my own decisions. In collective Chinese culture, and as the only child of my family, I should have found a job and settled down somewhere near them as most of the girls from my hometown did. However, I rejected several high-paid job offers in China and left a young man who can do anything for me, took a scholarship to fly to Europe far away from home to continue my doctorate training. I can still remember clearly what my father said, when I was hesitating whether to take a job four years ago. “Whatever jobs people take, they make money. But you are a person who has dreams; you should go and realize your dreams.” These words have given me strength and courage to go through tough times and kept me away from temptations. I am so grateful to have such understanding and supportive parents as they are. Besides, I also would like to thank my other friends apart from the names which have been mentioned above: Cora, Chaonan, Caijiao, Saier and Beth who always saved me out of stress and helped me recover my passion for science; Yuxiang and Edison who helped me improve my programming skills; Phillipe who showed me how neuroscience could be applied to industry; Ruizhi and Shaobo who encouraged me to move to Munich and took the challenge to learn computational modeling; Prof. Kong, Prof. Kalenscher, Prof. Tobler, Prof. Ma who has given insightful advice on my studies and life in general, invited me to join the professors’ lunch or dinner to meet the top scientists from other fields during the conferences, and helped me whenever they can; Janice, who treated me as her own daughter and always took care of me when I felt exhausted from work and during my most difficult time in the last break-up. Besides, I would also like to thank my friends from the church: Amir, Beth, Anja, Marcel, Anne, Daniel, Tobias, Esther and Roberta, who prayed for me and helped me find peace and love in God again.

In addition, I have to thank all the participants of my study, who had to keep their heads fixed for over one hour in the fMRI scanner. Thanks to all their patience and cooperation that I am able to acquire the high quality data.

Last but not least, I want to thank for the funding from DFG and China Scholarship Council to support my research and living costs in Munich. Also I would like to thank LMU and MPI for offering me high quality workshops and courses to

improve my soft-skills and enrich my knowledge in psychology, statistics, and neuroscience.

Deutsche Zusammenfassung (German Summary)

Hintergründe:

Dankbarkeit und Stolz sind beides Gefühle der Zufriedenheit und Anerkennung, wobei der Erfolg beim Stolz sich selbst und bei Dankbarkeit jemand anderem zugeschrieben wird. Dankbarkeit und Stolz sind entscheidend für die Funktion einer Gesellschaft. Sie ermöglichen uns, zwischenmenschliche Beziehungen zu schaffen und Selbstvertrauen aufzubauen. Trotz des wachsenden Interesses an den neuronalen Grundlagen positiver Emotionen und subjektiver Gefühle wissen wir noch sehr wenig darüber, wie diese Emotionen im Gehirn verarbeitet und sich im Laufe der Zeit durch neue Erfahrungen in ihrer Intensität verändern.

Ziele der Studie:

In meiner Studie untersuche ich die differentielle neuronale Verarbeitung der Emotionen Dankbarkeit und Stolz, und die algorithmischen und neuronalen Mechanismen, die dafür verantwortlich sind, dass sich die Expression der beiden Emotionen durch neue Erfahrung oder veränderte Umweltfaktoren anpasst.

Methoden:

Wir haben eine neuartige Aufgabe entwickelt, die auf der Quizshow "Wer wird Millionär" basiert. Während Probanden diese Aufgabe bearbeiten, messen wir ihre Verhaltensdaten und Gehirnaktivität mit funktionaler Kernspintomographie. Wir untersuchen, welche Hirnregionen an der Verarbeitung von Dankbarkeit und Stolz beteiligt sind, wie das menschliche Gehirn diese Emotionen im Laufe der Zeit repräsentiert und wie es Sie aktualisiert, wenn neue Informationen verfügbar sind. Um die Veränderungen über die Zeit zu erklären, erstellen wir computationale Modelle, die die Prozesse algorithmisch darstellen.

Ergebnisse:

Unsere Ergebnisse deuten darauf hin, dass Dankbarkeit mehr mit neuronalen Aktivitäten in der bilateralen temporoparietal Junction (TPJ) verbunden ist, ein Areal, dass bereits in früherer Forschung mit der Theorie des Geistes in Verbindung gebracht wurde. Im Gegensatz dazu war Stolz mehr mit neuronalen Aktivitäten im nucleus caudatus, der Teil des Belohnungssystems ist, und mit dem Hippocampus verbunden. Zwischen der Aktivität im Motorkortex (präzentraler Gyrus), im Belohnungssystem (ventrales Striatum und Putamen) und den Polen im Temporalkortex lässt sich ein direkter Zusammenhang mit der von den Probanden berichteten Stärke der Emotion nachweisen. Im Gegensatz dazu fand sich beidseitig im Putamen neuronale Aktivität, die direkt mit der berichteten Stärke des Stolzgefühls zusammenhängt. Darüber hinaus war die Aktivität in ventromedialem präfrontalen Kortex (vmPFC) mit einem emotionalen Vorhersagefehler-Signal verbunden, was darauf hindeutet, dass diese Region in den Prozess der Aktualisierung der empfundenen Stärke von Dankbarkeit und Stolz Gefühlen einbezogen ist. Regressionsmodelle deuten darauf hin, dass die empfundene Stärke von Dankbarkeit und Stolz durch unterschiedliche Faktoren beinflusst wird. Bei Dankbarkeit spielt hier insbesondere das erwartete Verhalten des Gegenübers eine Rolle, während die erhaltene Belohnung den grössten Einfluss auf Stolz hatte.

Bedeutung und Auswirkungen:

Unsere Ergebnisse beschreiben die Mechanismen und neuronalen Grundlagen für die positiven Emotionen Stolz und Dankbarkeit, die die Zuordnung der Belohnung begleiten, sei es aufgrund der eigenen Anstrengung oder der Hilfe anderer. Unsere Studien trägt zum Verständnis der beiden Emotionen bei. Die gewonnen Erkenntnisse könnten in Zukunft auch für die Entwicklung neuer Psychotherapien für Patienten mit emotionalen Störungen hilfreich sein.

Schlagworte:

Dankbarkeit, Stolz, Emotionen, Belohnung, fMRI, Computational Modeling

Abstract

Backgrounds:

Gratitude and pride are both benefit-related emotions, whereby the pride attributes success to oneself and gratitude to another. Gratitude and pride are vital to the function of a society, allowing one to create interpersonal relationships and build self-confidence. Despite growing interest in the neural underpinnings of positive emotions and subjective feelings, we know very little about how these emotions are represented in the brain and computationally updated over time by new experience.

Aims of the study:

We aimed to fill the gap by finding the specific neural representations of the dynamic emotional experience of gratitude and pride, and the functional neural substrates for updating positive emotions in general. Furthermore, we also aimed to find the best computational models to give the best explanations how these two emotions are updated as the environmental factors change.

Methods:

We developed a novel behavioral task based on the gameshow "Who Wants to be a Millionaire", which we used together with functional MRI, and computational modeling. We investigated which brain regions are involved in representing gratitude and pride, how the human brain keeps track of these emotions over time and how it updates them when new information is available.

Results:

We found that gratitude was more associated with neural activities in the bilateral temporoparietal junction (TPJ), which has previously been implicated in Theory of Mind. In contrast, pride was more associated with neural activities in the caudate nucleus, which is part of the reward system, and hippocampus. Importantly, when we look for neural activity parametrically modulated with the reported magnitude of gratitude feelings we found correlations mainly in the motor cortex (precentral gyrus), reward system (ventral striatum, putamen) and Theory of Mind network (temporal pole). In contrast, neural activity pertaining to the strength of the feeling of pride was found in the bilateral putamen. Moreover, activity in ventromedial prefrontal cortex (vmPFC) was related to an emotional prediction error signal, suggesting that this region might be involved in the process of updating our level of gratitude and pride feelings. Computational modeling revealed different models for gratitude and pride. Gratitude model uniquely involved the prediction of others' behavior, while pride model involved mainly the reward.

Implications:

Our findings delineate the computational mechanisms and neural circuitry for positive emotions that accompany the attribution of getting reward whether it is due to one's own effort or help of others. Besides, our studies contribute to theories of emotions in several different aspects, especially to the newest theory of constructed emotion. Our findings have clinical implications for developing new psychotherapies for patients with emotional disorders.

Key words: gratitude, pride, emotions, reward, fMRI, computational modeling

Table of Contents

1. Introduction

In the past decade, there has been great interest in the neural and computational mechanisms of human learning and decision making (Glimcher & Fehr, 2013). Numerous behavioral and functional imaging studies have extended our knowledge of how the human brain learns values of available actions or outcomes and on how a decision can be rendered between such multiple options (O'Doherty, 2004; Rangel, Camerer, & Montague, 2008). However, most of the experiments have so far focused on decisions that subjects made based on their own learning and experience. In such cases, any success or failure in the task could be attributed directly to the subjects, possibly leading to emotions of pride after successfully solving the task and gaining rewards. Yet, in daily life, many outcomes in our society are influenced by the help of others. For example, a collaborator in a scientific experiment provides samples or access to a machine, and thereby, facilitates success in the study. Or a loved one who cares for us while we are sick and help us with housework during that time. In those cases, any rewards we gain could—at least partially—be attributed to the help of another person. Most of us would then experience an emotion of gratitude in response. On the emotional level, both gratitude and pride could be seen as signals of accomplishment, and while complementary, the attribution of the pride is to oneself, gratitude is to another. Acting as an emotional currency for the achievement of reward, gratitude and pride are vital to the function of a society, allowing one to create interpersonal relationships (Algoe, Fredrickson, & Gable, 2013; Algoe, Haidt, & Gable, 2008). and build self-confidence (Tracy & Robins, 2007b; Weiner, Russell, & Lerman, 1979). However, less is known how these positive emotions are represented in the brain, and how the magnitude of their expressions are computationally and neutrally updated over time based on new experience. Therefore, in the current study, we aim to investigate the neural substrates and computational processes underlying gratitude and pride.

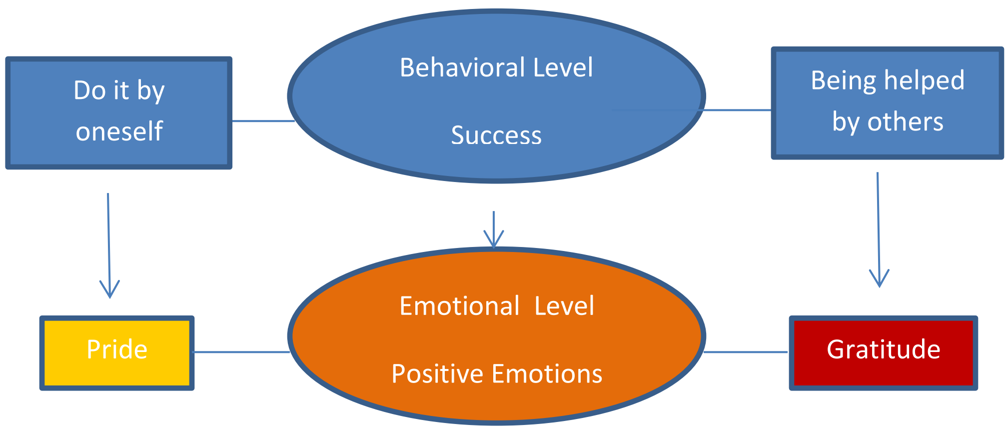

Figure 1. Attribution of success to oneself leads to emotional response of pride and attribution of success to others leads to emotional response of gratitude

1.1 Gratitude

1.1.1 The importance of studying gratitude

"A thankful heart is not only the greatest virtue, but the parent of all the other virtues."

— Cicero

The word gratitude originated from the Latin root "gratia", meaning grace, thankfulness and pleasure. The words linked with "gratia" are always positive in nature, which is in line with modern psychological studies that gratitude belongs to positive affects (McCullough, Emmons, & Tsang, 2002; Watkins, Woodward, Stone, & Kolts, 2003).

Unlike simple positive emotion like joy or happiness, gratitude has a unique conceptualization as a moral affect, and it plays three major roles in social interactions (McCullough, Kilpatrick, Emmons, & Larson, 2001). First, it serves as a moral barometer, a positive emotional response to indicate we have recognized the benefits from others' good behavior. Second, it motivates us to act prosocially towards the benefactor and people around us. Third, the expression of gratitude reinforces the benefactors to behave prosocially in the future. Therefore, gratitude not only benefits the well-being of both the beneficiaries and benefactors, but also creates a virtuous circle of prosocial behaviors in the society.

Empirical studies have revealed that gratitude promotes more prosocial behavior than neutral and even other positive conditions. In Bartlett and DeSteno (2006)'s study, the confederates and participants were asked to complete a series of cognitive tasks. In the gratitude condition, there was a malfunction with the participants' computer, the confederates helped the participants to fix their computers so that their work on the computer was restored and they did not have to redo the previous test. In the neutral condition, confederates and participants discussed about where the experimenter might be. While in the amusement condition, an amusing video clip was played to both the confederates and participants, and the confederates talked with the participants that they enjoyed the video and started a small talk with them. After the whole tests, the confederates asked the participants if they could help them with a long boring and mentally demanding survey. The time participants spent on the survey was calculated as the measure of the quantity of prosocial behavior. Results showed that, in the gratitude condition, participants spent six more minutes on the survey than those in the neutral condition. What's more, grateful participants spent eight more minutes than those in the amusement condition. Similarly, another study (Tsang, 2006) also demonstrates that gratitude condition promotes more prosocial behavior than a general positive condition: people receiving favors from the partners would like to give out more money than receiving good outcomes by chance. Even though generally positive emotions can increase helping behavior (Carlson, Charlin, & Miller, 1988), when the help is hedonically costly, people in positive emotions are less likely to engage in such behaviors (Isen & Simmonds, 1978). These studies show that gratitude has a unique prosocial action tendency than general positive emotions to promote more prosocial behavior even when the help is costly.

Evidence also shows that expressing gratitude reinforces the benefactors to be more likely to engage in prosocial behaviors. The most remarkable evidence is that participants who simply received a "Thank you" from the confederate would continue to voluntarily take more electrical shocks for her (McGovern, Ditzian, & Taylor, 1975). Another study found that writing a "Thank you" on the back of the check additionally by the server in the restaurant increases tips significantly (Rind & Bordia, 1995). Similarly, other studies have found that a simple appreciation call to former customers could increase sales (Carey, Clicque, Leighton, & Milton, 1976), and in health care field, receiving thank-you letters improved the case manager's visiting rate to residential clients (Clark, Northrop, & Barkshire, 1988). Furthermore, one study (Grant & Gino, 2010) explored two possible mechanisms how gratitude expression may influence prosocial behavior: the benefactors may have a stronger feeling of self-efficacy or social worth when being thanked, which could reinforce them to act prosocially. Results show that only social worth mediated gratitude expression and prosocial behavior. This study indicates that, expressing thankfulness motivates the benefactor to engage in further prosocial activities by making the benefactor feel socially valued.

What's more, gratitude can improve interpersonal relationship in the long run. A study of married couples (mean relationship length over 20 years) shows that feeling of gratitude predicted both one's own and the spouse's relationship satisfaction (Gordon, Arnette, & Smith, 2011). Not only in long term relationships gratitude was found to predict relationship satisfaction, but also in short term relationships such as cohabiting couples, increased gratitude was predictive of subsequent enhanced relationship quality (Algoe, Gable, & Maisel, 2010). The motivating and rewarding nature of gratitude emotion makes it have a prominent effect on relationship building than any other positive emotions. People who feel grateful are more likely to approach the benefactor, spend time with them, express their feelings and do something to create, maintain and deepen the relationship rather than simply conducting a repay behavior. Algoe and Haidt (2009) have found that, compared to recalling joy and admiration experiences, people recalling grateful memories are more willing to associate with and spend more time with the person in the future. Participants who recalled grateful memories reported the highest positive relationship focus than those recalling elevation, joy and admiration memories. To be more specific, gratitude elicitation increases participants' feeling of interpersonal closeness and eager to build relationships. In another study (Bartlett, Condon, Cruz, Baumann, & Desteno, 2012), people who were helped in one experiment by a benefactor were twice likely to come back for another experiment working with the same person than those in the neutral condition. Apart from the motivation to approach and affiliate with the benefactor, gratitude also increases trust in others, which is an essential part in forming a relationship. A series studies that people who wrote about their past grateful situations gave significantly higher trust rating scores in a subsequent trust judgement task towards unfamiliar others than people who wrote on pride, guilt and anger (Dunn & Schweitzer, 2005). While feeling grateful makes people more likely to trust others, some other studies suggested that expressing gratitude could make the beneficiaries appear more trustworthy (Bartlett et al., 2012; Gordon et al., 2011). Another important function of gratitude in interpersonal relationship is that it enhances the communal strength in relationship, which is to what extend one feels the responsibility to meet the needs of the other partner rather than act according to the social norm of reciprocity. This communal orientation in relationship is the basis of stable and mature relationships. In a series of studies by Lambert and his colleagues (2010), expressing gratitude was highly correlated with the perceived communal strength of the relationship, even after controlling for relationship satisfaction and social desirability. Furthermore, a longitudinal study showed that the increase in gratitude expression could predict the increase in communal strength in 6 weeks. Finally, an experimental manipulation directly tested the causal relationship between gratitude expression and the communal strength of the relationship. Participants were randomly assigned to four intervention conditions for three weeks. In gratitude expression condition, participants were asked to increase their expression of gratitude to their friends. In the other three control conditions, one is to ask participants to think about their daily activities, another is to think about grateful things about their friend but not to express them, the last is to think and talk about positive memories with their friend. After three weeks, the perceived communal strength in participants in gratitude expression condition was significantly higher than any other control condition. Taken together, the studies clearly demonstrated that gratitude expression can promote communal strength of a relationship, which is important for the development of a close relationship.

Gratitude not only benefits people in their social lives, but also in their subjective well-being. First, grateful people are generally happier, less likely to suffer from mental diseases, and have higher level of life satisfaction. Two early studies showed that, even using different questionnaires as measurements, trait gratitude was significantly and positively correlated with global happiness, positive affect, and life satisfaction (McCullough et al., 2002; Watkins et al., 2003). Subsequent studies replicated such findings in different populations and cultures (Chen & Kee, 2008; Froh, Kashdan, Ozimkowski, & Miller, 2009; Toussaint & Friedman, 2009). Furthermore, gratitude is proven to be a better predictor of subjective well-being than any of the big five personality traits across different studies (McCullough et al., 2002; Wood, Maltby, Gillett, Linley, & Joseph, 2008). On the other hand, trait gratitude was negatively correlated with depression and anxiety across two studies and four populations (Watkins et al., 2003), and increased trait gratitude can predict a decrease in depression and anxiety level (McCullough et al., 2002) . Besides, a study of older adults (N = 818) shows that, the influence of financial difficulties on depression was especially prominent for the elderly who were less grateful; however, it failed to aggregate depressive symptoms on those who were more grateful (Krause, 2009). Apart from depression and anxiety, a large sample study (N = 2621) found that religious thankfulness was also predictive of a lower risk of lifelong phobias, bulimia nervosa, and substance dependence such as nicotine, alcohol and cocaine (Kendler et al., 2003). Second, a lot of gratitude intervention studies have demonstrated a reliable effect on improving one's subjective well-being. Different forms of gratitude exercises were given in previous studies, such as listing three to five things to be grateful for (Emmons & McCullough, 2003; Geraghty, Wood, & Hyland, 2010a, 2010b; Seligman, Steen, Park, & Peterson, 2005), contemplating on things or people to be grateful for (Watkins et al., 2003), along with writing and delivering a gratitude letter to their benefactor (Froh, Kashdan, et al., 2009; Froh, Yurkewicz, & Kashdan, 2009; Seligman et al., 2005). Yet, no matter what form of exercises were given, the results similarly demonstrate a significant increase in positive affect and life satisfaction, as well as a decrease in negative affect (for a review, see Wood, Froh, & Geraghty, 2010). Gratitude intervention even outperforms other positive emotion interventions such as pride intervention in enhancing subjective well-being (Watkins, Uhder, & Pichinevskiy, 2015). As for short and long-term effects, gratitude intervention can immediately boost positive mood (Froh, Kashdan, et al., 2009; Watkins et al., 2003), and gratitude intervention's positive effects have been shown to last up to six months (Seligman et al., 2005). Gratitude intervention is also proven effective in different groups, such as college students (Study 1 & Study 2, Emmons & McCullough, 2003; Watkins et al., 2003), patients with neuromuscular diseases (Study 3, Emmons & McCullough, 2003), patients with body-image dissatisfaction (Geraghty et al., 2010a), patients with anxiety disorder (Geraghty et al., 2010a, 2010b), adolescents in school (Froh, Kashdan, et al., 2009; Froh, Yurkewicz, et al., 2009), and children in school (Froh, Kashdan, et al., 2009).

Gratitude also benefits one's physical health. In the gratitude intervention studies by Emmons & McCullough (2003), several health related outcomes were reported. First, 201 college students went through a 10-week intervention in one of the three conditions: (1) the gratitude condition. Participants write about five things they were grateful for during the week; (2) the major events condition. Participants write about five major events happened during the week; and (3) the hassles condition. Participants write about five hassles happened during the week. Results showed that people in gratitude condition reported fewer physical symptoms such as headaches, chest pain, skin irritation and sore muscles compared to the other two conditions. Meanwhile, people in gratitude condition also spent significant more time doing physical exercise (over one hour more per week) than those writing the hassles. In a subsequent study employing patients with neuromuscular diseases instead, gratitude writing intervention significantly improved patients' sleep quality, including increasing hours of sleep and a higher rate of waking up freshly in the morning. Later a cross-sectional study (Wood, Joseph, Lloyd, & Atkins, 2009) with big sample (186 males, 215 females) examined the relationship of gratitude and sleep in details, and found that higher trait gratitude can predict better sleep quality, longer sleep duration, less sleep latency and also less daytime dysfunction. The effect of gratitude on all the sleep related variables was mediated by pre-sleep cognitions: gratitude was related with high positive pre-sleep cognition and low negative pre-sleep cognition. Apart from positive effects on sleep, trait gratitude was found to be associated with lower risk in nicotine dependence, alcohol dependence and drug abuse (Kendler et al., 2003), which are all known to be detrimental to health and extremely hard to quit. In addition, gratitude is also found to be a protective factor of health against stress in older adults (Krause, 2006). Although the evidence of gratitude on physical health is relatively less than in social relationships and subjective well-being and more experimental work needs to be done, it is promising that gratitude can benefit a person physically, socially and emotionally.

In summary, gratitude's motivating and rewarding nature enables it to make a lot of unique contributions to society as a whole, interpersonal relationships and personal well-being than many simple positive emotions. It motivates and reinforces people to conduct more prosocial behaviors, creating a virtuous circle in the society. Furthermore, it also motivates people to build meaningful relationships, improves relationship quality, increase interpersonal trust and communal strength. Besides, on a personal level, gratitude also benefits people's subjective well-being: it increases positive emotions, decreases negative emotions and improves life satisfaction. In addition, gratitude can also benefit people's physical health by increasing physical exercises, improving sleep quality and protecting against stress and substance dependence.

1.1.2 The definition of gratitude

Gratitude has been defined as an emotional state, a personal trait, a habit, a moral virtue, a coping strategy in different contexts (Emmons & McCullough, 2003). It could be towards a particular benefactor (Tesser, Gatewood, & Driver, 1968) or towards life in general (McCullough et al., 2002; Wood et al., 2010). Therefore, it is hard to give a unifying definition.

Gratitude is a positive emotional response acknowledging and appreciating blessings in life. It can shift our attention to discover more blessings and become a habit of appreciating people and things that release us from sufferings and contribute to our well-being. The feeling of gratitude also fosters the habit to express thankfulness, such as saying thank you or acting for the benefits for the benefactor or other people, which in many cultures and religions is considered to be a moral virtue which promotes prosocial behaviors. Therefore, as mentioned previously, some researchers also define gratitude as a moral affect (McCullough et al., 2001).

As a personal trait or disposition, gratitude is defined as “a generalized tendency to recognize and respond with grateful emotion to the roles of other people’s benevolence in the positive experiences and outcomes that one obtains” (McCullough et al., 2002), and a “life orientation towards noticing and appreciating the positive in the world” (Wood et al., 2010). A grateful person should experience grateful emotions with greater intensity, more frequently, across a wider range of life circumstances, and attribute the good outcomes to a wider range of sources (McCullough et al., 2002).

In our study, we only concerns gratitude in terms of an emotional response towards a benefactor. From this perspective, gratitude is commonly triggered to people when “something good has happened to them, and they recognize that someone else is largely responsible for this benefit” (Watkins, 2014). In such situations, we conclude that two important factors are essential: (1) a reward is given, and (2) the credit is attributed to others. In the following paragraphs, we are going to discuss in details the attributes of reward which triggers gratitude and one important cognitive prerequisite which allows us to attribute the reward to other people.

1.1.3 Reward and gratitude

“The sentiment which most immediately and directly prompts us to reward, is gratitude.”

—Adam Smith

Even though gratitude is a response to acquiring reward or benefit from others, and also a motivator to get more reward by expressing thankfulness or conducting prosocial behaviors towards others, surprisingly, few studies have worked on the relationship between reward and gratitude. Gratitude is found to be a protective factor of substance dependence (Kendler et al., 2003), which is associated with dysfunctions in the brain’s reward system, but the mechanism behind it is yet unknown. In Wood’s social-cognitive theory of gratitude (Wood, Maltby, Stewart, Linley, & Joseph, 2008), benefit appraisal is the key cause to generate state level gratitude, and the benefits given should be perceived as valuable, costly, and of genuine motivation. Of these three attributes of benefits, the first two are related to reward. First, the benefit must be of something valuable to the recipient, either the person is in need of or desire to have. People would not be thankful for getting things they don't need or want to have. Second, the benefit related with high cost (time, effort, money) is correlated with higher level of gratitude. People are less likely to be grateful for the air they breathe, because it links with no cost. Therefore, the magnitude of the reward positively influences gratitude. Moreover, even though more help generally leads to more gratitude, one study (Wood, Brown, & Maltby, 2011) has found that the level of gratitude elicited by certain amount of help (a loan of 35 euro or 30 min of help) depends on the relative rank of all the help they can get from their friends, indicating the social comparison is involved. As for the form of the reward, it could be either material, like gifts, financial support, and helping behavior or non-material, e.g. love, emotional support or even spiritual experiences (Emmons & McCullough, 2003). Taken together, the reward which could trigger gratitude is influenced greatly by people's subjective evaluation and interpretation.

However, previous studies did not investigate the fluctuation of gratitude over a long time scale, during which the reward could increase or decrease all the time, and more importantly, past reward experience may influence the anticipation and evaluation of the current reward, too. For example, if you get one apple on the first day, two apples on the second day, and three apples on the third day, you may expect to get more than three apples on the fourth day and you may even think about making an apple pie or giving out to some friends when you receive them. If it turns out that on the fourth day you get five apples, you may be very grateful, but if you get only one apple, you may get disappointed and ungrateful, even though it is still better than nothing. Therefore, we speculate that the expectation of the reward, and the prediction error of the reward, which is the difference between what we actually get and what we expect, may influence our feeling of gratitude irrespective of the actual reward.

Besides, previous research didn't separate and quantify the contributions of the objective value and subjective value of the benefit to gratitude. For example, the objective value of one euro is always the same and can be quantified easily, and 10 euro is definitely a bigger reward than one euro. Nevertheless, depending on how much people need it, the subjective value changes from situation to situation and could play an independent role separated from the objective value. For instance, when people are in the supermarket and find out they don't have a one euro coin to unlock the shopping cart, they will appreciate a lot if someone can lend them a one euro coin rather than a 5 euro paper bill.

1.1.4 Theory of Mind and gratitude

To be able to feel grateful to another person, one must realize the good intentions behind the person's behavior. According to Wood's social-cognitive theory of gratitude (Wood, Maltby, Stewart, et al., 2008), realizing that the benefactor' help is out of genuine motivation for the welfare of the recipient is one of the three key components in benefit appraisal, which is the cause of the feeling of gratitude. In fact, in everyday social interaction, people infer others' intentions and motivations all the time. The ability to attribute mental states to others, represent others' thoughts and beliefs, identify their action intentions is called the Theory of Mind (Frith & Frith, 1999; Premack & Woodruff, 1978) . Theory of Mind is a fundamental social cognitive ability for successfully understanding other people and responding to others' actions, and therefore facilitating to establish and maintain interpersonal relationships. From a developmental perspective, this ability crystalizes earlier than gratitude (Emmons & Shelton, 2002; Froh, Miller, & Snyder, 2007) and it keeps developing with age (Happé, Winner, & Brownell, 1998).

Children before four years old cannot understand that others' mental representation of the world could be different from the reality (Wimmer & Perner, 1983), yet when they get older, more mature and complex mental presentations have been developed to better support attributing and understanding others' intentions and motivations.

McAdams and Bauer (2004) first proposed the hypothesis that Theory of Mind is a necessary cognitive prerequisite for developing gratitude, yet they did not provide empirical evidence to support this hypothesis. So far there is only one study (Freitas, O'Brien, Nelson, & Marcovitch, 2012) which examined the relationship between understanding gratitude and Theory of Mind in a group of 5-year-old children. Gratitude understanding was tested through two verbal stories, one story is about a girl who lost her cat, and an aunt stopped making a cake and spent a long time to help her find the cat. Another story is about a boy who felt cold at school on a winter day, and a girl lent a warm sweater to him. After the stories were told, the children were asked a series of questions including how the person being helped would feel, why they felt so and whether they should help the benefactors in another situation when they need help. Theory of Mind was measured through 3 kinds of tests. The first test is to see if children can take a different visual perspective. A card with different animal drawings on each side was presented between the children and the interviewer and questions were asked such as what kind of animal the interviewer was seeing. The second test is a classic false belief test, in which children were asked to predict a doll's behavior when she had a false belief of where a hidden object should be. In this scenario, one doll put an object in a box and left the room, and then another doll came in and moved the object to another box. Children were asked what the first doll would think where the object was and where she would look for it. Children must be able to represent the knowledge and beliefs of the situation which the first doll held, which were different from their own, to pass the test. The third test is a second-order false belief test, which is harder than the previous one. Children were told a story involving two interacting characters, e.g. Patty and Sally. In the story, Sally tries to play a trick on Patty by hiding her toy in the closet. However, Sally does not know that Patty has seen her hiding the toy from the outside of the window. Children were asked a series of questions about the mental states of the two characters. This test assesses the children's ability to represent and infer the thoughts of one person about another person. Results showed that, the scores of the two tests on understanding gratitude were both significantly correlated with all the three Theory of Mind test scores. Children who performed well on Theory of Mind tests have better understanding of gratitude, suggesting the important role of Theory of Mind on understanding gratitude. However, the hypothesis that Theory of Mind is a necessary requisite to understand gratitude is not well supported by subsequent analyses. Although children who did not have complete evidence of Theory of Mind (not correctly answered the questions in the second-order false belief task) were mostly those who did not understand gratitude, still around 5% of the children in both gratitude tests who did not have a complete evidence of Theory of Mind showed a understanding of gratitude. The results are conflicting with the previous one, and it could be due to two reasons: (1) the author only used the scores of the second-order false belief task as an indicator of the Theory of Mind ability, and children who did not pass the second-order false belief task could perform reasonably well on the other two Theory of Mind tasks, who was fairly enough to understand gratitude; and (2) gratitude was narrowly defined as whether repaying the favor was included in children's answers of justification for helping the benefactor. However, repaying the favor was more like to follow the social norm of reciprocity which does not necessarily require understanding the mental states of others. Given the conflict findings and the limitations of this study, more studies are necessary in order to fully understand the relationship between Theory of Mind and gratitude.

1.1.5 Neuroimaging studies of gratitude

Given the significant social, psychological and physical benefits of gratitude, people are curious to understand the neural mechanism how and why gratitude can have such great effects on our well-being. Neuroimaging is a powerful tool to understand how the brain works. It helps us to see what's going on in the "black box" of human mind to provide more clues to answer the questions arising from psychological theories and hypotheses. So far, several neuroimaging studies have made important attempts to uncover the neural mechanism of gratitude.

In an early study of neural correlates of human social values (Zahn et al., 2009), participants were presented with one sentence containing their actions towards their best friend either confirm or counter to a social value, or their best friend's action towards them. For example, "Tom (participant's best friend's name) acts generously towards Sam (participant)" was counted as a gratitude condition, and "Tom acts stingily towards Sam" was counted as an indignation/anger condition. In the functional MRI experiment, subjects were simply asked to judge their feelings as pleasant or unpleasant. Results showed that gratitude condition was specially associated with neural activities in the hypothalamus. Using the same paradigm, the subsequent structural MRI study (Zahn, Garrido, Moll, & Grafman, 2014) found that grey matter volumes in the right inferior temporal gyrus were positively correlated with individual differences in gratitude. However, in this paradigm, there is no guarantee that the participants really experienced a specific gratitude feeling during the task. Later researchers such as Fox (2015) thinks it is a moral judgement task which could involve gratitude, and Kini (2016) thinks the paradigm is more about identifying emotions in a social narrative, rather than experiencing emotions.

The first study specially investigating the neural correlates of the experience of gratitude is from Fox and his colleagues (2015). In this fMRI study, verbal narratives were applied to induce gratitude. Participants were asked to imagine that they were the survivors of the Holocaust, for each trial they read a short story depicting a scenario in which they were being helped by some kind-hearted people, and then reflected on how they would feel if they were in that situation. After the reflection period, they were asked to rate how grateful they feel from 1(a little) to 4(a lot). The results showed that the gratitude rating correlated with brain activities in a cluster covering many regions in mPFC, including ACC and OFC.

The neural correlates of gratitude expression and how gratitude expression intervention influenced neural activity were studied (Kini et al., 2016) on patients with anxiety and depression. Two groups of patients went through the following treatments for 3 months before the fMRI scan: one group performed a gratitude writing intervention (writing a gratitude letter to express gratitude) additionally to psychotherapy, and the other group did only the psychotherapy without additional intervention. The fMRI experiment used a 'pay it forward' task, in which participants were given some money at the beginning of each trial, and then were asked to donate some money to a charity cause according to how grateful they felt. Participants were also asked to rate how much the decision was influenced by gratitude, desire to help and guilt respectively. First, the study found that across the whole population, gratitude modulated brain activities during donation decision (expression of gratitude) in the left superior parietal lobule, left superior frontal gyrus, left inferior frontal gyrus, and right middle occipital gyrus after controlling for desire to help and guilt. Second, compared to the therapy-as-usual group, gratitude intervention group had greater gratitude modulated brain activity on pregenual ACC. Similarly, a recent study (Karns, Moore III, & Mayr, 2017) found a change in fMRI BOOD signals in vmPFC after a gratitude writing intervention in healthy population for 3 weeks.

The role of intentionality in help in influencing gratitude was recently investigated by Yu and his colleagues (2017) using a pain sharing task. In the fMRI task, the decision of whether to share the pain stimuli for the participants was made either by a human partner (intentional) or by a computer (unintentional). After the pain delivery, participants were asked to transfer an amount of money to the partner. Results showed that, at the decision stage, intentional help elicited more brain activities in vmPFC, SMA, precentral gyrus and postcentral gyrus. At the pain delivery stage, the activities in bilateral insula, precentral gyrus and postcentral gyrus were suppressed in intentional help condition. Moreover, the activation in vmPFC can predict the individual differences in reciprocity (the amount of money transfer), and the activation in PCC can predict self-reported gratitude. Further multivariate pattern analysis showed that the septum/hypothalamus's neural activity pattern can dissociate intentional help from unintentional help conditions.

Taken together, previous neuroimaging studies have highlighted brain regions which were involved in Theory of Mind (dorsal mPFC) (Fox et al., 2015), reward processing (vmPFC, OFC, VTA, and caudate) (Fox et al., 2015; Karns et al., 2017; Kini et al., 2016; Yu et al., 2017), moral cognition (mPFC, ACC) (Fox et al., 2015; Kini et al., 2016), predicting the effects of others' action (ACC) (Kini et al., 2016) and basic emotions and pain (Insula) (Fox et al., 2015; Yu et al., 2017).

However, there were several limitations with previous studies. First, the study claimed Theory of Mind involvement (Fox et al., 2015) used verbal narratives to elicit gratitude, which already inherently required participants to take on a third person perspective (as if they were survivors of the Holocaust).Therefore, it is not known whether Theory of Mind is involved because of understanding the social narratives, or from the feeling of gratitude. Interestingly, even though understanding intention is an important part of Theory of Mind, in the paradigm which intentionality in gratitude was studied (Yu et al., 2017), no Theory of Mind brain regions were reported and discussed. Second, for the social interactive tasks, participants were forced to repeatedly receive money or pain sharing in each trial. It was very unnatural, and can be very different from the gratitude feeling generated in real social interactions. Third, the role of expectation is not taken into consideration. Participants may raise their level of expected reward or help during the course of experiment. Fourth, we observed small effect sizes in previous studies, and some of brain activities were likely to be task-specific. Therefore, more research is needed to shed light on the neural basis of gratitude.

1.2 Pride

1.2.1 The importance of studying pride

"If you believe in yourself and have dedication and pride - and never quit, you'll be a winner. The price of victory is high but so are the rewards."

-Paul Bryant

Pride is a self-consciousness emotion that is psychologically and evolutionally important to us (Tracy, Shariff, & Cheng, 2010). Unlike basic emotions, self-consciousness emotions are closely linked with self-representation, and help us to attain complex social goals (Tracy & Robins, 2004a). Among the five different kinds of self-consciousness emotions: shame, pride, guilt, envy and embarrassment, pride is the only positive emotion that makes us feel good about ourselves. The positive feeling of our global 'self' reinforces us to repeat the behaviors which lead us to feel proud and motivates us to pursue higher achievements, in the long term it increases our self-worth, and while in social interaction, the nonverbal expression of pride signifies the success to other people in the society and promotes social status (Tracy & Robins, 2004b; Tracy et al., 2010).

First, feeling proud is not only the results of achievement and good social conduct, but also the motivator of more pride-eliciting behaviors in the future. Though much theoretical work (Tracy & Robins, 2004a, 2004b, 2007a) had predicted the motivational role of pride in achievement and socially valued behaviors, the first empirical evidence comes from Williams and DeSteno (2008) 's studies. In the first experiment, participants were required to perform two cognitive tasks, of which the latter is tedious and mentally taxing. After finishing the first task, participants were given three kinds of feedback: (1) pride-eliciting, the experimenter told the participants that they did well on the task with appreciating voice and smiling to indicate she was impressed; (2) control condition without releasing information on their task performance; and (3) control + condition, the experimenter informed the participants of their good task performance, but without any clue that the she was impressed. This condition was to generate self-efficacy, which is a cognitive estimation on how well one can perform on prospective tasks (Bandura, 1977), different from the affective experience of pride. Afterwards, all the participants went through the second task, and the amount of time they could continue to work on the task was recorded as a measure of perseverance. Results found that, participants in pride condition spent significantly more time on the second task compared to those in the two control conditions, indicating that the emotion of feeling proud motivates individuals to be more persistent to pursue higher achievement, and this effect cannot be explained by self-efficacy. The second experiment was similar to the first one, except that it used positive emotion as another control condition instead of self-efficacy: after the first cognitive task, participants in the pride-eliciting and control conditions viewed a set of affectively neutral pictures, while those in the positive emotion condition viewed positive emotion-eliciting pictures. Result replicated that participants in pride condition worked significantly longer than the two control conditions. Moreover, the effect on perseverance on subsequence task cannot be explained by general positive emotion. From these two studies, we can see that pride has a unique motivational effect on perseverance of achieving further success. Besides, another study (Pekrun, Elliot, & Maier, 2009) also has similar findings that high levels of pride can predict good academic performance in 218 undergraduate students. On the other hand, some researchers investigated the motivational effect of low pride on achievement (Weidman, Tracy, & Elliot, 2016). They found that low level of pride also promotes achievement, by motivating changes on achievement-related behaviors. More specifically, in study 1 they tested a group of athletes under training for long distance running programs and found that participants with low level of pride in their training progress reported higher intention to change their following training plans, adjust training habits in order to achieve better training effects. Furthermore, in a large sample of students (N = 1024) the researched tested the model addressing the relationship between pride, achievement-related behaviors and academic achievement. Results showed that there was a significant indirect effect for exam-specific pride on subsequent exam scores. Students who did not perform well on the initial exam had lower level of pride, which predicted a higher intention to change their study plan and habits as well as longer hours per day to prepare for the second exam. These achievement-related behaviors predicted significant higher subsequent exam scores. Taken together, the motivational effect of pride on achievement is seen through two paths: one with high level of pride, which reinforces people to maintain or persevere on similar behaviors, the other with low level of pride, which drives people to change their old behavior patterns for more adaptive ones (Weidman et al., 2016).

Second, pride is closely related with self-esteem. Brown and Marshall (2001) first found that pride, rather than a lot of any other emotions, such as shame, inspiration and enthusiasm which were all measured by the Positive and Negative Affect Scale (PANAS) (Watson, Clark, & Tellegen, 1988), has the highest correlation with self-esteem, which was measured by the Rosenberg Self-Esteem Scale (RSE) (Rosenberg, 1965). Tracy and her colleagues further distinguished two kinds of pride, and found that only achievement-based authentic pride was positively correlated with self-esteem, and associated with successful social relationships and mental health (Tracy, Cheng, Robins, & Trzesniewski, 2009; Tracy & Robins, 2007b). Subsequent studies replicated the positive correlation between pride and self-esteem, and pointed out the mediating effect of pride between self-esteem and positive affect (Stanculescu, 2012). The influence of pride on self-esteem is evolutionally adaptive. From a functionalist view, after certain achievements one would feel proud as a response; and this positive feeling informs one's self-worth and social value, which promotes positive feelings and thoughts about one's global self, and results in high self-esteem (Tracy et al., 2010).

Third, pride could convey and enhance one's social status. Studies have shown that people have certain stereotypes of emotions associated with high social status: for a negative outcome, the person described as angry was believed to have higher social status than the one described as sad and guilty; for a positive outcome, the person described as proud was believed to have higher social status than the one described as appreciative (Tiedens, Ellsworth, & Mesquita, 2000). Besides, feeling proud influences how one acts in a social group, e.g. behaving in a more dominating style, which will influence the judgement of other group members on one's social status. In a social interaction experiment (Williams & DeSteno, 2009), participants who were manipulated to be in proud status by being praised of their good performance in the previous task, were perceived as more dominating in a subsequent group problem-solving task. Meanwhile they were also more liked by their interaction partners. The results suggest that the pride increased perceived social status and value in a group by promoting dominating behaviors. Moreover, the nonverbal expression of pride is proven to be an automatic indicator of high social status. A series of studies using the implicit association test (Greenwald & Banaji, 1995) and the affect misattribution procedure (Payne, 2005) find out that people react significantly faster when pairing nonverbal expression photos with high status words than any other emotion expressions, such as shame, happiness, disgust, anger and fear; and this strong association cannot be explained by the artifact of aggrandized posture size such as outstretched arms (Shariff & Tracy, 2009). Therefore, individuals displaying pride expressions would be automatically considered to be high in social status. As the resources are always limited in a society, the implicit judgment process is evolutionarily adaptive for the proud individuals to acquire more resources and gain social support.

1.2.2 The definition of pride

The word "proud" comes before the noun form of "pride", originated before 12th century from the old French word "prud" or "prouz", meaning valiant or brave. At that time it was used positively by the Norman knights to describe themselves. Yet, later the word was used by the Anglo-Saxons to describe their invading army in a negative tone, meaning conceited and self-aggrandized (Tracy et al., 2010).

Pride often arises when one appraises a positive, socially valued outcome such as success as to his or her own contribution, such as efforts, personalities, and abilities (Leary, 2007;Tracy et al., 2010; Williams & DeSteno, 2008). Besides, people could also feel proud for owning valuable objects (Leary, 2007), good outcomes from other people they are identified with (e.g., their family members and friends), and even in a more collective level, such as feeling proud for their country (Tracy et al., 2010). Nevertheless, most research studies pride on an individual level as a self-consciousness emotion.

The experience of pride requires a self-evaluation process. This process requires two cognitive capacities (Lewis, Alessandri, & Sullivan, 1992). One is called objective self-awareness, which directs attention introspectively and treats oneself as an evaluable object; the other is called internalized standards of behavior, which could either generated from one's own past experience, or from others' feedback. These two cognitive prerequisites enable us to evaluate our behaviors/qualities according to our inner standards to judge whether they are good enough to make us feel proud of.

Pride has also been described by its unique nonverbal expression. Many studies have reliably found that pride has a universally recognizable bodily and facial expressions, which is distinct from other similar emotions such as happiness and excitement (Tracy & Robins, 2004b, 2007a; Tracy et al., 2010). The typical nonverbal expression of pride includes: a small smile, slightly tilted head, raised arms, and visible expanded posture of upper body. The recognition of the nonverbal expression of pride is as fast and accurate as that of basic emotions

(Tracy & Robins, 2008), indicating that pride may be an innate and evolutionary adaptive emotion.

Despite the universal nonverbal expression, pride serves two highly divergent effects. Historically, pride was viewed as negatively as “the root of all evil” (Baasten, 1986) and the greatest of the Seven Deadly Sins (Alighieri, 2003). In the bible the bad effect of pride was described in the famous Proverbs: “pride goes before destruction, a haughty spirit before a fall” (Proverbs 16:18). On the other hand, Aristotle considered pride as the “the crown of the virtues” (Ross, 1925) and Nietzsche admired the man who “has the pride to live by the values he wills” (Nietzsche, 1968). Therefore, some researchers differentiate pride into two facets: hubristic and authentic (Tracy & Robins, 2004a, 2007b). The first facet is less attached to one’s actual achievement but more out of conceited self-aggrandizement. It is associated with high narcissism, shame-proneness while low implicit self-esteem, and low agreeableness and conscientiousness in the “Big Five” personality traits. The second facet is based on actual achievement and positively correlated with self-esteem, and more adaptive personal traits such as extraversion, agreeableness, and conscientiousness. In terms of mental health and social functioning, people with high trait hubristic pride are more likely to suffer from chronic anxiety, aggression, hostility, rejection sensitivity, Machiavellianism, low dyadic adjustment, low perceived social support and antisocial behaviors such as drug abuse. In contrast, people with high trait authentic pride are less likely to suffer from depression, anxiety, social phobia, and rejection sensitivity; they also have a high level of relationship satisfaction, secured attachment style with their partners, high social support, and conduct more prosocial and achievement-oriented behaviors (Tracy et al., 2010). Therefore, the hubristic pride is a maladaptive, narcissistic view of oneself, and probably a self-defensive mechanism of low self-esteem; while the authentic pride is a prosocial, fact-based view of oneself, it brings achievements and genuine self-worth. In our study, as the pride is elicited by actual achievements we will mainly talk about pride in terms of the authentic pride.

1.2.3 Reward and pride

Pride by nature is usually triggered by acquiring a reward on one's own. The reward could either be success experience, such as accomplishing one's goals and acquiring more possessions or social approval such as compliments from others.

As a self-conscious emotion, the reward has to be related with one's identity and ideal self-representations (Tracy & Robins, 2004a; Tracy et al., 2010). For example, an honest person would not feel proud even if he gets a lot of money by lying to others. And if a woman who doesn't care about her appearance is complimented by her beauty, she would also not take pride in the compliments. Only when the reward aligns with one's values and can add to one's self-worth would it lead to the feeling of pride.

In achievement tasks, the subjective value of a task depends on the task difficulty. Generally, success on a more difficult task is more rewarding than on an easy task. As a result, greater pride responses were observed after the success on difficult tasks than easy tasks (Belsky, Domitrovich, & Crnic, 1997; Lewis et al., 1992).

Pride is not only a positive emotional response to the reward, but also a reinforcer and motivator for self-rewarding behaviors. As what we have discussed in 2.1 the importance of studying pride, pride can promote achievements and achievement-related behaviors. High level of pride makes people persevere on similar tasks, while low level of pride makes people more willing to adaptively change their behavioral patterns and put more efforts to achieve better outcomes (Pekrun et al., 2009; Weidman et al., 2016; Williams & DeSteno, 2008).

1.2.4 Neuroimaging studies of pride

To investigate neural substrates of pride, some studies contrast it as a self-conscious emotion in comparison with some basic emotions such as joy and anger. The first neuroimaging study of pride comes from Takahashi and his colleagues' research comparing the neural correlates of pride and joy (Takahashi et al., 2007). In the study, participants were asked to imagine they were the protagonists in the scenarios depicted in verbal sentences. Pride scenarios included sentences such as "I was awarded a prize for my novel" while joy scenarios included sentences such as "I won a lottery". Besides pride and joy conditions, a neutral condition was also included as a contrast. The neutral scenarios were depicted in sentences like "I had breakfast". Pride compared to the neutral condition induced greater activations in the right pSTS and left temporal pole; however they did not find activations in MPFC as expected. Pride compared to joy condition yielded greater activations in the right pSTS. Besides, the subjective rating of pride was also positively correlated with the activations in the pSTS. Another study similarly used verbal sentences to describe events which could elicit one of the four emotions: pride, guilt, joy and anger. For example, a pride statement was like "It has occurred that I showed assertiveness in front of people (at work, with a service provider)". When a statement was presented, participants were asked to respond whether this event has happened to them or not by pressing yes or no buttons on the left or right. Results showed that pride compared with all the other emotions activated greater activations in a cluster in the vmPFC extending to the OFC. Self-conscious emotions compared with basic emotions in general were related with greater activations in the frontal lobe, including the mPFC, dACC, and the MFG. (Gilead, Katzir, Eyal, & Liberman, 2016).

Studies also compared the brain activations associated with pride and other self-conscious emotions such as shame and guilt. In one fMRI study (Roth, Kaffenberger, Herwig, & Brühl, 2014), participants were asked to recall events during which they either felt ashamed/guilty or pride when they saw certain visual cues. Besides, a neutral condition was added as a control, which required the participants to simply wait for the upcoming neutral pictures. Results showed that pride compared with shame/guilt condition was associated with stronger brain activity in the left superior frontal gyrus, vmPFC, cingulate gyrus, especially the posterior part, inferior temporal gyrus, inferior parietal lobe, caudate body and lateral thalamus. Pride compared with neutral condition yielded greater activations in the mPFC, precentral gyrus, insula, PCC, precuneus, amygdala, and lateral thalamus extending to the caudate body. In another fMRI study which we mentioned before (Gilead et al., 2016), comparing pride with guilt conditions also showed greater activation in ventral part of the mPFC, which was consistent across these two studies.

Besides, there is another study comparing pride as a self-focused emotion with an others-caring emotion compassion (Simon-Thomas et al., 2012). Participants were presented with one of the three kinds of emotion-inducing picture slides: (1) depicting scenes of vulnerable suffering or harm to induce compassion; (2) depicting unspecific in-group achievement and status scenes such as graduation ceremonies to induce pride; and (3) depicting normal life scenes with people and objects to induce neutral feelings. After watching each series of slides, they were also asked to rate their emotion intensity and quality. Results showed that, first, pride condition compared with neutral condition was associated with greater activations in the posterior medial cortex. Moreover, pride compared with compassion involved more brain regions including the posterior medial cortex, parahippocampal gyrus and inferior temporal gyrus.

The neural activity and anatomical basis of pride have also been investigated treating pride as a one of the moral sentiments by Zahn and his colleagues (2014; 2009) . As we have mentioned while reviewing neuroimaging studies of gratitude, the two studies applied the same paradigm: participants were presented with one sentence verbal statements containing their actions towards their best friend either confirm or counter to a social value, or their best friend's action towards them. For example, "Sam (participant) acts generously towards Tom (participant's best friend's name)" was counted as a pride condition. In the fMRI experiment, subjects were simply asked to judge their feelings as pleasant or unpleasant. Results showed that the pride condition was specifically associated with neural activities in the septum. In the subsequent structural MRI study, it was found that grey matter volumes in the cuneus and precuneus were negatively correlated with individual differences in pride. However, in this paradigm, there is no guarantee that the participants really experienced a strong and genuine feeling of pride during the task. Later researchers such as Fox (2015) thinks it is rather a moral judgement task, and Kini (2016) thinks the paradigm is more about identifying emotions in a social narrative, rather than experiencing emotions.

Recently, the intrinsic brain activity relating with the two facets of pride was also investigated (Kong et al., 2017). The Authentic and Hubristic Pride–Proneness Scales (AHPPS) (Tracy & Robins, 2007b) were used to measure the individual differences in the authentic and hubristic facets of pride. The fractional amplitude of low-frequency fluctuations (fALFF) in the brain's resting state was used as a measurement of the intrinsic brain activity to correlate with the behavioral scores of the authentic and hubristic pride. Results showed that, the individual differences in authentic pride were positively correlated with the fALFF in the bilateral superior temporal gyrus. However, individual differences in hubristic pride were positively correlated with the fALFF in the left OFC and PCC.

Taken together, the neuroimaging studies of pride have used verbal narratives (Gilead et al., 2016; Takahashi et al., 2007; Zahn et al., 2014; Zahn et al., 2009), others' success scenario pictures (Simon-Thomas et al., 2012), and recalling participants' own pride memories (Roth et al., 2014) to elicit pride, despite one study (Kong et al., 2017) using questionnaires to measure the individual differences in pride. Converging neuroimaging evidence has shown that pride could probably involve the self-referential processing (mPFC, PCC, and precuneus), reward processing (caudate, vmPFC, septum and OFC), memory retrieval (PCC, temporal pole, parahippocampal gyrus and inferior temporal gyrus), social cognition (right pSTS, superior temporal gyrus), affective processing (amygdala, insula and ventral striatum) and Theory of Mind (mPFC, pSTS and temporal pole).

However, the studies have a few limitations. First, the success scenarios depicted in the verbal narratives and pictures could be far from the participants' own life experience, thus may require participants to take a third person perspective to imagine how they would feel if they were in that situation. That could be the reason why the Theory of Mind regions were involved in these studies. Second, for the recalling success task, it was not possible to explicitly control whether participants were thinking about the pride events or something else. Meanwhile, it was also not clear whether the involvement of the memory regions was due to the nature of the task or the feeling of pride. Third, pride generated real-timely as an emotional response of success or compliments could be fundamentally different than reflecting on it from the narratives (Schilbach & T., 2013). However, none of the previous neuroimaging studies have used achievement tasks to elicit pride directly.

1.3 Reward

1.3.1 The neural correlates of reward representation and reward learning

In recent years, there have been a lot of neuroimaging studies on how humans represent reward and learn to update the value of the reward in the brain. Even though the studies use various different tasks, depending on their complexity, the majority can be classified into 3 categories: (1) passive receiving reward paradigms; (2) instrumental reward paradigms; and (3) reward decision making paradigms (Richards, Plate, & Ernst, 2013). Paradigms in the first category don't require participants to do anything in response to the stimuli to acquire the reward, for example, in the classic conditioning paradigm (J. P. O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003), participants were presented with abstract fractal visual stimuli, and then followed by a delivery of liquid with either a pleasant sweet taste, a neutral taste or a no taste. Whereas in the instrumental reward paradigms, participants have to perform a task correctly in order to get the reward. The challenge in the tasks could be of working memory, such as in the pirate paradigm (Galvan et al., 2006), a perceptual judgement, e.g., the cake gambling task (van Leijenhorst, Crone, & Bunge, 2006), or a motor response, like in the monetary incentive delay task (Knutson, Fong, Adams, Varner, & Hommer, 2001). The actions performed in the tasks are associated with a single expected value in the instrumental reward paradigms, which is a big difference from the reward decision making paradigms. The expected value or the expectation of the reward is an estimation formed based on participants' experience during the task, and it is updated once being compared with the actual gain or loss. The discrepancy between the actual reward and the expected reward is called the reward prediction error, which is considered the major indicator of reinforcement learning (Montague, Dayan, & Sejnowski, 1996; Rescorla & Wagner, 1972; Sutton, 1988). In the reward decision making tasks, participants are usually presented with several different options and each is linked with a different reward probability or magnitude. Therefore, participants have to form different expected values and compare between the choices to select the one with the highest expected value to maximize gain and minimize risk of loss. For example, in the wheel of fortune task (Ernst et al., 2004), participants were presented with a wheel with two colored slices, each representing the probability of winning for a choice, and two choice tags displaying different amount of money corresponding with the color slices indicating its occurrence probability. The high winning probability is usually paired with small amount of money and vice versa. There is no absolute correct choice to gain the maximum reward as in the instrumental reward tasks, the choice reflects the decision maker' preference on risk-taking in this task, and in other reward decision making tasks some other preferences such as delayed gratification, novelty could also be tested (Richards et al., 2013).

The brain regions involving in reward processing and reward learning mainly include: the vmPFC, OFC, ventral striatum, dorsal striatum, and amygdala. The vmPFC represents value signals, encodes stimulus values and reward outcomes (Glimcher & Fehr, 2013; Knutson & Cooper, 2005; J. P. O'doherty, Cockburn, & Pauli, 2017). For example, an fMRI experiment asked participants to rate how pleasant the presented images were: images of face, house and painting in the scanner, and found that the value signals, which were the subjective pleasantness ratings, positively correlated with activities in the vmPFC (Lebreton, Jorge, Michel, Thirion, & Pessiglione, 2009). Furthermore, the vmPFC is essential for learning and updating reward values. In a reversed-reward contingency task, patients with lesion in vmPFC made significantly more errors than the groups of damage in dlPFC and normal controls when the stimulus-reward association contingencies had changed: they tended to still choose the old stimuli which were rewarded before, indicating a deficit in updating the stimulus-reward associations (Fellows & Farah, 2003). Last but not least, the vmPFC acts as a common currency among different kinds of reward. In a willingness-to-pay task, the vmPFC is correlated with three categories of reward: food, non-food consumables, and money (Chib, Rangel, Shimojo, & O'Doherty, 2009). One step further, another study demonstrated that the activation level in vmPFC was proportionally scaled according to the subjective value of the stimuli regardless of the categories they belonged to (Levy & Glimcher, 2012). Adjacent to vmPFC, the OFC also plays an important role representing value signals. It encodes stimulus value across diverse sensory modalities, such as gustatory, visual, auditory, olfactory and somatosensory (O'Doherty, 2004). Meanwhile, it is also activated in more abstract reward representations, such as money and art, together with vmPFC (Knutson & Cooper, 2005). Besides, the OFC is involved in reward prediction. In a classic conditioning paradigm, selectively devaluing one of the rewards caused a decrease in the OFC, amygdala and ventral striatum signals associated with the predictive cue paired with the devalued reward (Gottfried, O'doherty, & Dolan, 2003). In addition, the OFC also computes the reward prediction error: it is significantly correlated with a positive prediction error when an unexpected reward is delivered and a negative prediction error when an expected reward is omitted (J. P. O'Doherty et al., 2003). The ventral striatum consists of the olfactory tubercle and the nucleus accumbens. It receives mid-brain dopamine projections, and is associated with reward prediction and reward prediction errors during reinforcement learning and gambling tasks (J. O'doherty et al., 2004; J. P. O'Doherty et al., 2003; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). The dorsal striatum consists of the putamen and the caudate nucleus. Studies show that, even though the dorsal striatum also receives prominent dopamine projections as the ventral striatum, it plays a relatively different role in reward learning (Balleine, Delgado, & Hikosaka, 2007; J. O'doherty et al., 2004). It is especially involved in encoding action-outcome association in goal-directed decision making, and action-selection based on expected values of those actions. Therefore, it may mainly contribute to the stimulus-response-reward learning. The amygdala has been considered to relate with aversive stimuli (Zald & Pardo, 1997), however, later in reward processing studies (Anderson et al., 2003; Small et al., 2003), it was found to represent the stimulus reward value and specifically, it responds more to stimulus intensity rather than stimulus valence. Besides, the amygdala is also involved in reward expectation and prediction. It encodes the conditioned stimuli which predict the subsequent delivery of appetitive or aversive outcomes (J. P. O'doherty et al., 2017; O'Doherty, 2004).

1.3.2 The emotions underlying the reward processing

Although we have known much about the neural responses to reward and the learning process as discussed above, the emotions related with reward during the process have been largely ignored. People may generally feel happy when they receive a reward, and upset when they don't. More specifically, when the task requires effortful thoughts and actions to earn the reward, people may feel proud of themselves once they win. When someone else contributes to the good results by collaborating or offering voluntary help, people may feel grateful as an acknowledgement for others' kindness.

In many instrumental reward paradigms and reward decision making paradigms, the participants perform the task alone and have to make efforts to estimate the upcoming reward magnitude and probability, and come up with the best policy to maximize the total reward. Participants have to win the challenges in the tasks in various forms, for example the pirate paradigm (Galvan et al., 2006) relating working memory, the cake gambling task (van Leijenhorst et al., 2006) on perceptual judgement, the monetary incentive delay task (Knutson et al., 2001) on motor response, and the wheel of fortune task (Ernst et al., 2004) on balance of risk and return. Winning the reward in these tasks activated brain regions such as the nucleus accumbens, OFC, mPFC, ACC, ventral lateral PFC, putamen, dorsal caudate. However, participants could also feel proud as they gained the reward through their own efforts and abilities to accomplish the tasks by themselves and maximize the reward. Therefore, part of the brain activations during the reward time could be explained by pride. This idea could also be supported by the facts that the neuroimaging studies of pride not using the reward paradigms also reported some reward regions such as mPFC and OFC, as we have discussed before. Considering the close relationship between reward and pride, it is surprising so far no neuroimaging studies have examined how reward contributes to the feeling of pride directly.

In recent years, reward decision making in a social interactive setting has becoming more and more popular. In these new paradigms more social emotions are involved. These paradigms usually involve another player (or confederate) to perform the task with the participants, and many of the paradigms are based on game theory, for example, the ultimatum game, trust game and the prisoner's dilemma game. In the ultimate game, two players split certain amount of money. One player acts as the proposer, and offers a plan of division. The other player acts as the responder, and can either accept or reject the offer. If the proposal is accepted, the money will be divided accordingly. If the proposal is rejected, neither of the players will get any money. The game is usually played only once with the same player to avoid reciprocity. If human beings are purely rational, then no matter how unfair the offer is, they should accept the offer so they could at least receive some money. However, it is consistently found that most subjects would reject the low offers (20% of the total money) about half of the times (Bolton & Zwick, 1995; Henrich et al., 2001). This paradigm has showed the influence of emotions in human decision making behavior, especially anger or disgust for unfairness. Evidence from neuroimaging studies (Corradi-Dell'Acqua, Civai, Rumiati, & Fink, 2012; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003) revealed that the anterior insula, known for processing negative emotions, including anger and disgust (Damasio et al., 2000), had a higher response to rejected unfair offers, suggesting the significant role of inequality aversion interfering decision making. The similar effect is also seen in the trust game, during which one player (investor) first invest some of his or her own money, and then the money is tripled and transferred to the other player (trustee). Then the trustee decides how much to give back to the investor. The investor could apply costly punishment when receiving unfair returns from the trustee. The activation in the investor's caudate nucleus was associated with the magnitude of punishment, indicting a hedonic effect of the punishment (De Quervain, Fischbacher, Treyer, & Schellhammer, 2004). In the prisoner's dilemma game, two players have to choose to cooperate or defect without any communication. If one player chooses to defect while the other one chooses to cooperate, then the one who defects will receive a high payoff but the other one will get a punishment. If both of them choose to defect, then both will get a punishment, whereas if both of them choose to cooperate, both of them will win a reward. The problem is that, the players do not know the other player's decision beforehand, thereby creating the dilemma. Neuroimaging studies showed that, the participants showed greater brain activation in the reward regions such as the mPFC, OFC and ventral striatum when the other player chose to cooperate, and decreased activation in the same regions when the other player defected (Rilling et al., 2002; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004). Furthermore, the unreciprocated cooperation elicited increased activity in bilateral anterior insula, which usually involves in negative emotions, and its functional connectivity with lateral OFC in response to unreciprocated cooperation can predict subsequent defection (Rilling et al., 2008). However, even though participants reported the feeling of happiness, camaraderie, trust, and relief during mutual cooperation condition, anger, irritation and disappointment during other-defected condition, happiness, guilt and relief during self-defected condition, and general low and undifferentiated emotions during both-defected condition (Rilling et al., 2008), how these emotions would interact with the experiment conditions and interfere with the results were not examined.

Taken together, previous social decision making studies revealed that humans are not perfectly rational, and their decisions are often influenced by emotions, especially the negative emotions generated by violation against social norms such as fairness and reciprocity. The brain activities in regions such as anterior insula could be contributed by the involvement of those emotions. However, to what extent those emotions influence decision making behavior, how they update their value and how they relate with brain regions modulating the effect has not been studied. Besides, people not only prefer fairness, but also conduct altruistic behaviors irrespective of their own interest. In daily life, we often help strangers find their ways somewhere, give up our seats to the elderly and children, and at the same time also receive sincere help and care from people around us. The feeling of thankfulness or gratitude while being helped could have a positive long term effect in promoting cooperation and prosocial behavior than the mere monetary interest.

The emotion of gratitude seems to have a very rewarding nature as we have discussed before, however, how this emotion is learned in social interactions and whether it is updated in the brain in a similar way as the monetary reward is still not known yet.

There has been an fMRI study specifically investigating the emotion of happiness in reward decision making paradigm (Rutledge, Skandali, Dayan, & Dolan, 2014). In the study, the participants played a probabilistic reward task during which they could choose between a certain reward and a gamble. Every two or three trials, they were asked to rate "How happy are you at this moment?" by moving a cursor on a scale bar. Thus, the gains and losses in the game induced rapid fluctuations in affective states and the experience sampling method allowed a quantitative measurement of the participants' subjective well-being. In the end, a computational model inspired by dopamine function was applied to model the subjective rating of happiness. In the model, happiness was explained by the expected value of the chosen gamble, the prediction error of the experienced and predicted reward, the certain reward, and a constant factor. This study has demonstrated an example of using computational model and fMRI to decode the neural mechanism underlying the changes in emotion.

1.4 Unaddressed problems and limitations of previous research

Although previous studies have contributed a lot discovering the psychological and social benefits of gratitude and pride, understanding the brain reward circuit, and developing paradigms to measure gratitude and pride behaviorally and with neuroimaging, there are still some unaddressed problems and limitations arising from previous research.

First, although our emotional reaction to a situation need adjustment over to time for us to be more socially adaptive, to date, no studies has identified the neural substrates behind such bookkeeping of gratitude and pride over a longer period of time and updating the strength of the motion based on new experience. Previous neuroimaging studies of gratitude and pride were either unable to measure the changes in gratitude and pride or overlooked the fluctuations in emotions during the experiments. Therefore, it is also unknown whether the brain updates these positive emotions using the same functional regions as reward updating.

Second, currently no computational models of gratitude and pride are established to quantify the influence of multiple environmental and mental factors. For example, apart from the well-known contribution of reward per se to gratitude and pride, whether the expectation of others, the anticipation for future reward, the mental comparison (prediction errors) of what we actually get and expect, and the task difficulty influence our feelings of gratitude or pride is still unknown.

Third, the verbal narratives paradigms that most neuroimaging studies of gratitude and pride applied have three key inherent limitations: First of all, perspective-taking and Theory of Mind are indispensable to be able to engage in the depicted scenarios. This implicates Theory of Mind brain regions in both gratitude and pride conditions, making it hard to identify the exact contribution of Theory of Mind in generating these two emotions. Secondly, some of the scenarios may be far from the participants' own experience, thus making it hard to generate genuine and strong emotional feelings, resulting in a small effect size on the brain. Lastly, the neural processing of emotions generated in real-time social interactions could be fundamentally different than “observing” them from narratives (Schilbach & T., 2013).

Fourth, a lack of ecological validity and good experiment control also exists in previous neuroimaging studies. For the interactive paradigms in gratitude studies, participants were forced to repeatedly receive money or share pain in each trial. It was very different from natural social interactions. For the recalling success memory task in pride studies, it was impossible to explicitly control the content participants were contemplating about. Mind wandering or thinking about something else may cause the activation of brain areas such the default mode network, which may have little to do with the emotion of pride. Besides, the memory task also makes it difficult to determine whether the involvement of the memory regions was due to the nature of the task or the feeling of pride.

Fifth, whether gratitude and pride have shared or dissociated neural substrates needs further exploration. So far only one early study (Zahn et al., 2009) on social values labeled two of the four conditions as gratitude and pride in one verbal narrative paradigm. However, later researchers (Kini et al., 2016) think the paradigm is more about identifying emotions in a social narrative, rather than experiencing emotions. What’s more, in this study no significant group effects were found specific to gratitude when contrasting gratitude with pride, indignation/anger conditions. This could be due to little emotional involvement in the paradigm and it could also be seen from the overall small effect size in the fMRI results.

1.5 Aims of study

To address the problems stated above, in the current study, we aim to investigate the neural substrates and computational processes underlying gratitude and pride through a new social interactive paradigm which is more ecologically validated, explore the learning mechanisms of gratitude and pride, and ascertain the contributions such as reward and Theory of Mind to gratitude and pride.

1.6 Research questions

We mainly would like to answer the following research questions:

1) What are the neural substrates of gratitude?
2) What are the neural substrates of pride?
3) What are the unique neural substrates to gratitude and pride in comparison with each other?
4) How does the brain update gratitude and pride?
5) Which psychological factors (environment, learning…) can best explain gratitude and pride?

1.7 Hypotheses

Here are the hypotheses regarding to the research questions:

First, based on previous psychological and neuroimaging studies, gratitude may involve brain regions that also process Theory of Mind, such as the dmPFC, TPJ, precuneus, temporal pole, and IFG; reward processing, such as the vmPFC, OFC, VTA, ventral striatum, caudate and putamen; moral cognition, such as the mPFC and ACC, and basic emotions such as amygdala and insula.

Second, pride may involve brain regions which are also engaged in self-referential processing, such as the mPFC, PCC, and precuneus; reward processing, such as the ventral striatum, caudate, putamen, vmPFC, septum and OFC; memory retrieval, such as the PCC, hippocaumpus, parahippocampal gyrus and inferior temporal gyrus; social cognition such as the right pSTS and STS; affective processing such as the amygdala and insula. Contrary to previous studies, we assume that without the story-based paradigm, pride would not involve Theory of Mind processing, which typically includes the mPFC, precuneus, pSTS, TPJ and temporal pole. As the self-referential processing also includes the mPFC, PCC, and precuneus, we speculate that the pride would specifically not involve regions such as TPJ and temporal pole.

Third, gratitude compared with pride may involve more brain regions regarding social interaction and Theory of Mind, which mainly include pSTS, TPJ, mPFC, precuneus, PCC, temporal pole, and IFG; pride compared with gratitude may involve more brain regions for self-referential processing, such as the mPFC, PCC, and precuneus. Although the central midline structures such as mPFC, PCC and precuneus are both involved in Theory of Mind and self-referential processing, according to the simulation theory for the Theory of Mind (Saxe, 2006), people use their own mind as a model to understand and predict the other mind, thus the self-referential processing may be an underlying basic processing for Theory of Mind. Therefore, we speculate that the central midline structures such as mPFC, PCC, and precuneus may have increased activity for more complicated Theory of Mind related processing, as in our study the gratitude condition.

Fourth, as gratitude and pride are both positive emotions, we could infer from how the brain updates the value of happiness (Rutledge et al., 2014). Our brain may update positive emotions similarly as reward, also through prediction error signals. Generally, when we get more than what we expected, a positive prediction error arises, and this may active many dopamine-projected brain regions, such as the ventral striatum, mPFC, OFC, ACC, etc. and we feel elevated emotions. While specifically for gratitude, to form expectation requires an understanding of the benefactor's mind to estimate whether this person would still help them in the long run. For pride, the expectation is always based on a person's previous success experience. As the mPFC is both involved in understanding others' mind and self-referential processing, we hypothesize that the mPFC could be the neural substrate to update the values for both gratitude and pride.

Last, we speculate that the factors influencing gratitude may include: the magnitude of reward, how much the beneficiary needs help, personality trait, the expectation of reward, the prediction error of reward. Besides, we also hypothesize that the gratitude level also depends on our expectation of the benefactor: when we get more help than expected, we would be more likely to feel grateful. By contrast, when we have high expectation of others while get less than expected, we would probably feel less grateful. Therefore, even when we get the same amount of help which is valuable for us, we may feel totally different depending on our expectations. Hence the expectation of others' helping behavior and the prediction error of the help are also taken into consideration as two independent factors for gratitude. However, for pride, the most important factors may include: the magnitude of reward (success), how challenging the task is, personal trait, the expectation of reward and the prediction error of reward.

1.8 Methodology

In this study, we developed a novel fMRI paradigm called the “Millionaire Game”, which allows the participants to have social interaction with a real person and generate genuine feeling of gratitude and pride. To elicit the fluctuations in emotions, the task difficulty, whether the benefactor is helpful, and win or loss were manipulated during the task. To measure the changes in gratitude and pride reliably, we used the well-established experience sampling method to record the moment-to-moment emotional states on a numerical scale. Besides, we applied the same experimental procedure for both gratitude and pride conditions to control the amount of stimulus and level of social interaction were equivalent in both conditions.

We applied an event-related fMRI design rather than a block design to avoid the emotional fatigue from exposing to one condition of emotion for a long time. The brain activities were associated with events and parametric regressors in a design matrix in a general linear model to identify the neural substrates for gratitude, pride and the learning parameters.

Finally, we used computational modeling to see which environmental and psychological factors contributed most to the feelings of gratitude and pride.

1.9 Overview of study

In the following chapters, the methods of the current study will be descripted in details in chapter 2, including the fMRI task, experiment design, fMRI data acquisition, and statistical analyses. The results will be displayed in chapter 3, including the neural substrates of gratitude and pride, neural contrast of gratitude and pride, neural correlates of updating values in gratitude and pride and the computational models of gratitude and pride. In chapter 4, we will discuss the results in context of previous findings, and give potential explanations and implications for further research. In chapter 5, we will have a broader discussion of the contributions and limitations of the current study to this research field, and propose several further research directions. In the last chapter, we will conclude the current study relating to our research questions and hypotheses as well as its implications for further studies.

2. Methods

2.1 Participants

28 right handed participants, 19-28 years of age (M = 24.39, SD = 2.44, 18 female), participated in the study. All of them were from western cultural background to limit potential confounding of gratitude expression through variations in different cultural backgrounds (Watkins, 2014). Six additional participants were not included in the analysis due to technical problems (severe signal loss in the prefrontal region) or because they did not complete the experiment. All participants had normal or corrected-to-normal vision and no history of either neurological/psychiatric illness or any other contraindications to the MRI environment. Participants gave written informed consent and were later paid a compensation of 25 to 30 euro depending on their performance level. The study was approved by the ethics committee of the Ludwig-Maximilian-University Munich.

2.2 fMRI task

We developed a new task "The Millionaire Game" inspired by famous TV show "Who Wants to Be a Millionaire" using Cogent Graphic (http://www.vislab.ucl.ac.uk/cogent_graphics.php). The task consisted of either 30 trials (pre-session used for training outside the scanner) or 120 trials (fMRI experiment) of knowledge-based questions with four answer choices.

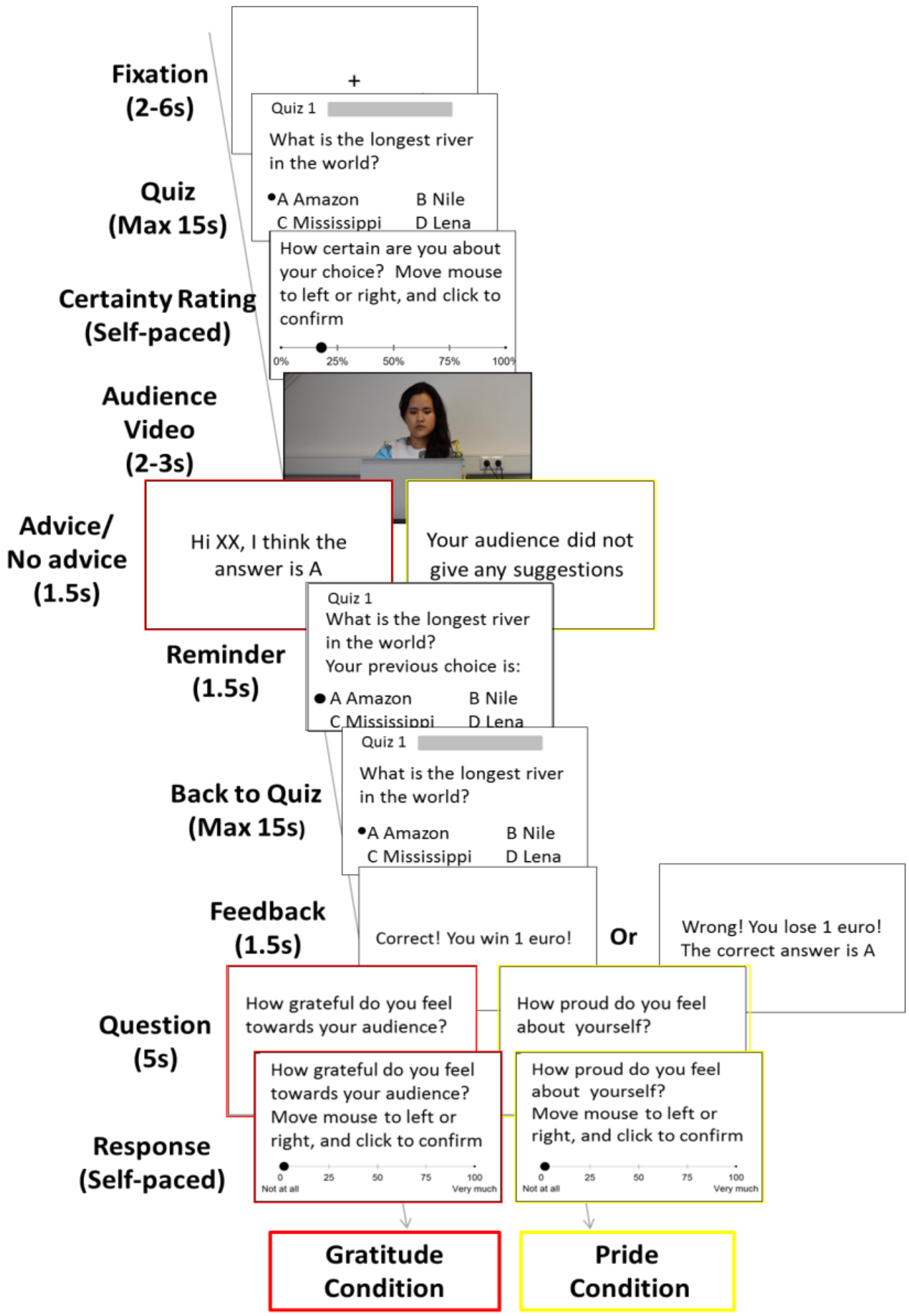

Figure 2. The millionaire game task

Each trial began with a screen presenting a knowledge question from our database, four possible answers, and a time bar counting down within 15s. To control the difficulty of the task, the question pool consisted of 60 easy and 60 difficult questions. Participants had to answer each question in 15 seconds expired by clicking on one of the four possible answers. Next, they had to rate how certain they were about the answer by moving a black dot on a 0-to-100% slider. This certainty rating gave us some information on how urgent the participants actually needed help on this trial. We assigned participants to either one of two conditions: 1) gratitude condition or 2) pride condition based on that information: trials in which participants' certainty ratings were lower or equal to 50%, were set to the gratitude condition, otherwise the pride condition. This ensured that subjects were really in a situation where they needed help and would not be annoyed by any advice in a situation where they already knew the answer.

In the gratitude condition, subjects were then presented with a video clip showing a xx sec. recording of another person who was introduced to the subject as audience lifeline in the game. In the video, the person looked at the quiz on a laptop in front of her and then pressed a button as if to give a suggestion. The suggestions could be either right (80%) or wrong (20%) to simulate a normal human's knowledge and also to prevent participants from getting satiated and reaching a plateau of gratitude quickly. Then the suggestion was presented to the participants. After showing the advice from the audience, a brief reminder of the participants' previous own choice was presented for 1.5 seconds. Participants were then directed back to the question screen with four choices and the time bar, where they could click on their final choice. After the click, a feedback of whether the answer was correct or not and the monetary reward was presented. If the answer was wrong, the correct answer would be revealed to the participants as well on the screen. Finally, we asked subjects how grateful they felt towards the audience. This question was shown for

5s to give participants time to contemplate on their feelings. Then a slider would appear on the screen under the question so that they could rate their feeling from 0 (not at all) to 100 (very much).

Conversely, the pride condition was entered when the participants' certainty rating was higher than 50%. Following the certainty rating a screen showed up displaying the audience person looking at the quiz on a laptop in front of her but did not give any suggestion. Afterwards a notice of no suggestions was presented to participants. As in the gratitude condition, a brief reminder of participants' previous choice was then shortly presented. Participants were then directed back to the question screen, where they could choose their final answer. Then they could see the feedback whether the answer was correct. After the feedback was presented, similarly, we presented the participants with a question screen (for 5 s) asking how proud they felt about themselves. Then a slider would appear under the question so that the participants could rate their feeling from 0(Not at all) to 100(very much).

2.3 Experimental design

One important goal of our task was to elicit true feelings of gratitude. Thus we had to ensure that the participants believed that the audience was a real person and helping them when they were in need.

Environment Setting. To make the social interaction setting of the task more credible, each participant met with the "audience" in person before the experiment. Unbeknownst to the participants, the "audience" was a confederate, a member of our lab, and all videos shown during the experiments were pre-recorded.

The participants were also informed that the "audience" would get paid based on her fixed-time participation, therefore had neither the shared or conflict interest to help them. In this way, we set the expectation of acquiring the help from "audience" to zero at the beginning of the experiment.

Video Stream. Participants were informed that the "audience" was sitting outside the scanner room and would be connected to them through a real-time video stream in each trial. While in fact, short video clips of the "audience" sitting in front of the computer and pressing a button were video-taped in advance. This way we could control that body movements and facial expressions of the "audience" was identical across participants. During the videotaping, we ensured that the audience person had neutral facial expression and acted as if she was giving suggestions to the participants. 123 video clips for the gratitude condition 184 video clips for pride condition were recorded in total. (As we had 120 trials in total including both conditions) The video clips were played non-repeatedly and sequentially to the participants.

Pre-session Meeting. Around one week before the fMRI experiment, each participant was invited to the lab to meet the "audience" in "The Millionaire Game" in person and do a short exercise of the task. During the pre-session meeting, participants performed a 30 trial version of the fMRI task (these 30 questions would not appear in the subsequent fMRI experiment). Suggestion from the audience was given after the screen displaying "Connecting…" rather than the video stream. Participants were told that they would see the "audience" on the screen in the fMRI environment settings. After completion of the task, we checked that participants believed in this social interaction setting. Only those who performed reasonably well and believed the social interaction were invited back for the fMRI experiment.

fMRI Experiment. On the scanning day, participants performed the full version of The Millionaire Game (120 trials). Before entering the MRI scanner, participants again met the member of our lab who acted as the “audience”. The task was divided into 4 EPI runs of 30 questions and between 25 to 30 minutes each. After the fMRI scan, a brief questionnaire was given to participants to assess their feelings during the game.

2.4 fMRI data acquisition

Neuroimaging data were collected using a 3T whole-body Siemens MAGNETOM Skyra scanner with 64 channel head coil located at the Klinikum der Universität München. T2* echo-planar images (with a multiband factor of 4) were collected with repetition time (TR) = 1760 ms, a matrix size of 106 x 106, an echo time (TE) = 38.6 ms, 45 deg flip angle, and field of view (FOV) = 212 x 212 mm. 60 axial slices were included in every volume. The voxel resolution is 2.0 x 2.0 x 2.0 mm^3. The phase encoding direction was from posterior to anterior commissure. Five dummy scans at the beginning of each run were discarded to allow for stabilization of the MR signal. The structural brain images were acquired applying an MP-RAGE T1-weighted sequence (TR/TE = 2.4 / 2.17 ms) with 320 x 256 matrix size, resolution 0.75 x 0.75 x 0.75 mm^3, 12 deg flip angle, and a FOV of 240 x 240 mm.

2.5 Statistical analysis

2.5.1 Computational models for gratitude and pride

First, we collected available factors which could contribute to feelings of gratitude and pride (see Table 1 below) in the game.

We then used, separately for gratitude and pride, stepwise regression to determine the best model to predict the trial-by-trial gratitude and pride ratings by those potential factors.

As potential factors we considered the following:

$$Rating = w_0 + w_1 * Reward + w_2 * EV_r + w_3 * RPE_r + +w_4 * EV_b + w_5 * RPE_b + w_6 * Certainty$$

In both conditions, participants rated how certain they felt on the question (Factor Certainty), which reflected the subjective difficulty for the current trial, and got rewarded by their correct answers (Factor Reward). Furthermore, we hypothesized that gratitude and pride were not purely triggered by the gain of reward, but involved a higher level of comparison in human mind between the expected reward/help and the reality (RPE_r/ RPE_b). If people expected less and gained more, they were more likely to experience the feeling of gratitude or pride, than those who expected more but gained less. And more specifically, such expectation could either be of the reward (EV_r), or of the person who might help them (EV_b). The latter could be more important to the feeling of gratitude, in which people could feel really grateful when they did not expect the help from the audience, but actually got the right advice contributing to winning in the game. We also assumed that the two kinds of expectations would create a good feeling like hope, which would eventually contribute to the feeling of gratitude or pride. The constant factor w_0 accounted for individual differences in the baseline level of gratitude or pride.

Table 1. List of Model Factors

Factors	Names	Explanations
w_0	Constant factor	Accounts for the baseline level of gratitude or pride; conceptually similar as mood; this could be influenced by genes, education, personality, etc.
Reward	Trial reward	The win or lose for each trial, values could be 1, or -1.
EV_r	Expectation of reward	Whether participants could get reward or not for the next trial based on learning from previous experience. $EV_r(i+1) = EV_r(i) + \alpha * (Reward(i) - EV_r(i))$
RPE_r	Reward prediction error	The differences between the actual reward and expected reward. $RPE_r(i) = Reward(i) - EV_r(i)$
EV_b	Expectation of the benefactor	Whether the audience would give right, wrong or no advice (Outcome = 1,-1 or 0) for the next trial based on learning from previous experience. $EV_b(i+1) = EV_b(i) + \beta * (Outcome(i) - EV_b(i))$
RPE_b	Benefactor prediction error	The differences between the real feedback from the audience(Outcome) and EV_b $RPE_b(i) = Outcome(i) - EV_b(i)$
Certainty	Certainty rating	The number participants scaled on the bar from 0 to 100%; converted to a [0, 1] ranged score.

Note:

i marked the trial index, for example i = 3 referred to the third trial.

α was the learning rate of the reward, 0< α<1, to indicate how fast people could learn about the upcoming reward.

β was the learning rate of the benefactor, 0< β <1, to indicate how fast people could learn about the audience's behavior, how likely the audience gave the correct or wrong advice.

Note that evaluation of the model is complicated by the fact that EV and RPE itself contain free model parameters and that those variables are functions that require data from trial 1..i for evaluation.

We therefore used a cost function based on the Bayesian Information criterion to determine the goodness of fit for each regression model. The BIC also accounts for model complexity by adding a penalty for each parameter.

Parameters Fitting. The fmincon function in Matlab Optimization Toolbox (https://de.mathworks.com/products/optimization.html) was used to find the minimum of the squared differences of the model predicted values of gratitude/pride and the true gratitude/pride ratings. Once the least square error was found, the parameter values for the current model were determined.

We used the residual sum of squares (RSS) to calculate the BIC for each model, the function was as below:

$$\mathrm{BIC} = n * \ln\left(\frac{RSS}{n}\right) + k * \ln(n)$$

In which n was the number of trials, RSS was the differences of actual rating of gratitude/pride and the model predicted values, and k was the number of parameters used in the model.

Then the BIC values for all the models were compared to identify the models with the smallest BIC for gratitude and pride.

As validation, we generated random gratitude rating for each trial each subject and used the same model building procedure as we had used before, and found that the model with smallest BIC is the one factor model with the constant factor w0. The same is with the randomly generated pride ratings: the model with smallest BIC is the one factor model with w0 only.

2.5.2 fMRI data analysis

The imaging data were analyzed with SPM12 (www.fil.ion.ucl.ac.uk/spm/software/spm12/).

Preprocessing. Unwrapped field maps which measured the magnetic field inhomogeneity were converted into voxel displacement maps and later used for unwarping. The EPI images were realigned to correct for head motion and unwarped to correct for geometric distortions. Then the T1 images were coregistered to the mean unwrapped images and structurally segmented using CAT12 toolbox (http://www.neuro.uni-jena.de/cat/). DARTEL based normalization (Ashburner, 2007) was applied to the segmented grey matter and white matter images to the template space and Montreal Neurological Institute (MNI) space and then applied the generated flow fields in the previous step to normalize the functional images in MNI space with smoothing (Gaussian blurring kernel of 5 mm Full width at half maximum; FWHM).

Besides, the segmented grey matter images were smoothed in 6mm FWHM and used to create grey matter mask images at the threshold of grey matter probability > 0.1.

The event related design was based on the general linear model (GLM) approach with random effects implemented in SPM12 on two levels. In the first level analysis, each subject's GLM was estimated using a canonical hemodynamic response function with no derivatives, a high-pass filter cutoff of 150 seconds, within individual subject's grey mater mask, masking threshold 0.4 for model estimation.

General Linear Model. Event regressors in the GLM model included the presentation of the questions, the action of clicking on the choice, the presentation of certainty rating, the click on certainty rating, the presentation of the video clips, the feedback from the audience of advice or no suggestions, the reminder of participants' previous choice, the presentation of the question again, the action of clicking on the final choice, the presentation of feedback whether the answer was correct or not (separate regressors for gratitude and pride trials), the time of the question inquiring the feelings of gratitude and pride, the consecutive presentation of the slider under the question, and the subjects' clicks on the slider to rate for gratitude and pride.

For the first GLM we were interested in the neural correlates of gratitude rating, pride rating, and the specific neural markers for gratitude and pride. Therefore in this model we included three parametric regressors: certainty rating, gratitude rating, and pride rating. The time when the certainty rating was shown was parametrically modulated by certainty rating. According to our post fMRI experiment questionnaire, participants felt the highest level of gratitude or proud when they received feedback of whether their answers were correct or not. Therefore the event of the quiz outcome in gratitude condition was parametrically modulated by gratitude rating, and the event of quiz outcome in pride condition was parametrically modulated by pride rating.

In a second GLM model we would like to see whether the brain had common updating neural substrates for gratitude and pride. We calculated the trial-to-trial differences in gratitude, and combined with trial-to-trial differences in pride as one parametric regressor named Combo, as a measure of updating values in these two positive emotions in general. The event during the time of quiz outcome of all the trials was modulated by the Combo regressor. The event regressor for rating on certainty was still parametrically modulated by values of certainty rating.

In the third GLM model we would like to explore which brain regions were sensitive to changes in the feelings of gratitude and pride separately. First, we calculated the trial-to-trial differences in gratitude rating and marked this variable as GGi. Trial-to-trial difference in pride rating was saved as PPi. The event of quiz outcome in gratitude condition was parametrically modulated by GGi, and the event of quiz outcome in pride condition was parametrically modulated by PPi. Same as usual, the certainty rating event regressor was parametrically modulated by the values of certainty rating.

The six head movement regressors were also included in all the models as regressors of no interest. Linear weighted contrasts were computed to identify effects of interest in the first level analysis, providing contrast images in each subject to evaluate the random effect further on a group level.

Second Level Analysis. One sample t-tests were performed for each regressor of interest on a group level based on individual subject contrasts of parameter estimates derived from the first level GLM analysis. Besides, to find the specific neural correlates for gratitude and pride, we also made a second-level contrast of gratitude vs. pride conditions of the events when the participants saw the quiz outcome as highest level of gratitude and pride feelings were reported.

For all analysis, we used Family-Wise Error (FWE) correction for multiple comparisons at the cluster level. We considered results as significant at $P < 0.05$ using a cluster defining threshold of $P = 0.001$.

fMRI results were viewed and displayed in bspmview toolbox for SPM12 (version 20161108) (http://www.bobspunt.com/bspmview/).

3 Results

3.1 Behavioral statistics

Participants rated their feeling of gratitude and pride on a scale of 0 (Not at all)-100 (very much), which were converted to values between 0 and 1. Data showed that our experiment was effective in inducing gratitude and pride: participants had an average gratitude rating of 0.67 (sd = 0.15), and an average pride rating of 0.55 (sd = 0.17). Participants got correct answers and thus being rewarded in 77.47% of the gratitude trials and 86.06% of the pride trials.

Although participants were correct in most of the times, whether they gain or lose reward had a significant influence on their gratitude ratings and pride ratings. For the gratitude condition, the mean gratitude rating was 0.75 in the gain trials and 0.24 in the loss trials. This difference was significant: $t(27) = 26.50$, $p < 0.001$. Moreover, for each gain trial, the gratitude rating increased by 0.11 on average, while for the loss trials the gratitude rating decreased by 0.27 on average. These changes between the gain and loss trials were also significant: $t(27) = 19.24$, $p < 0.001$. There was a similar situation for the pride condition: the gain trials had an average pride rating of 0.59, while the loss trials had only 0.15, which yielded a significant difference ($t(27) = 23.31$, $p < 0.001$). On average, for each gain trial the pride rating increased by 0.07, while for each loss trial the pride rating decreased by 0.22— such a difference was also significant ($t(27) = 20.66$, $p < 0.001$).

In gratitude condition, the advice offered by the “audience” also had a significant influence on gratitude ratings ($t(27) = 27.29$, $p < 0.001$): for trials given correct advice, the average gratitude rating was 0.75, while for those given wrong advice, the average gratitude rating was 0.17, even lower than the average loss trials. Each piece of correct advice increased the gratitude rating by 0.10 on average, while each piece of wrong advice decreased the gratitude rating by 0.33 on average, which had a significantly different effect ($t(27) = 20.66$, $p < 0.001$).

However, the relationship between the certainty rating of participants' own choice and the gratitude rating or the pride rating was not significant.

3.2 Neuroimaging results

3.2.1 Specific neural representation for gratitude and pride

To get strong and selective neural response signals for gratitude and pride, we made a contrast between gratitude and pride conditions as a control to each other, during the time when participants saw the reward. We choose this event because participants reported feeling most grateful or proud during this time.

We expected to find brain regions processing Theory of Mind, such as the temporal parietal junction (TPJ), pSTS, mPFC, precuneus, PCC, and inferior frontal gyrus (IFG) being more active in gratitude condition, because understanding others' action and intention is necessary to feel grateful. Meanwhile, we expected brain regions involving self-referential processing, such as the mPFC, Precuneus, PCC and also memory, such as the hippocampus, are more responsive in the pride condition, as one needs to compare the current event with previous experience and encode rewarded events. In addition, since both gratitude and pride are positive emotions and could relate with the reward system, the brain regions involved in basic emotions and reward processing could be more active in either condition or absent as a result of contrasting. In addition, as both conditions may involve the mPFC, precuneus and PCC, it is uncertain whether they will be more engaged in either condition or absent from the results because of contrasting.

We found that, compared with pride, gratitude was related to increased activation in brain regions including (see Figure 3 & Table 2): the bilateral TPJ (rPTJ peak at MNI 50 -52 28, t = 8.15, 719 voxels; lTPJ peak at MNI -58 -56 18, t = 6.12, 853 voxel), IFG, right dorsomedial prefrontal cortex (dmPFC), and precuneus.

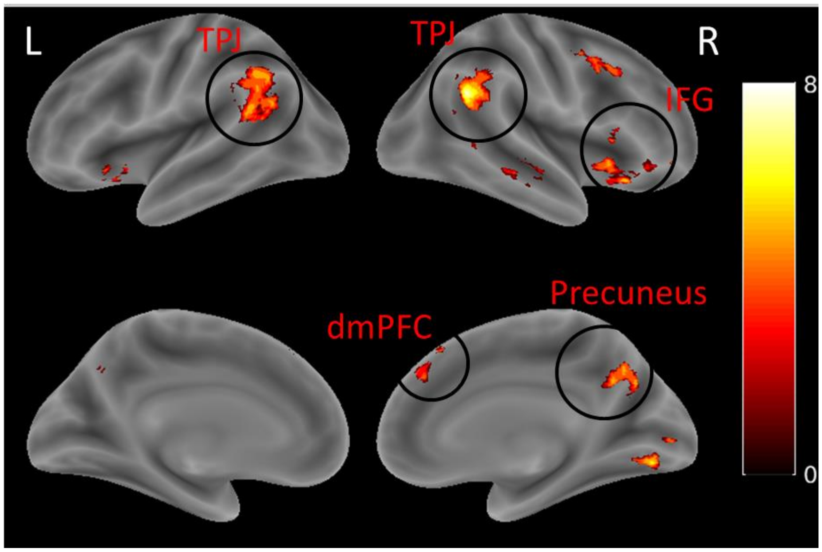

Figure 3. Contrast of gratitude > pride conditions

Table 2. Contrast of gratitude and pride

Contrast Name				**MNI Coordinates**		
	Region Label	**Extent**	**t-value**	**x**	**y**	**z**
Gratitude > Pride	R TPJ	719	8,148	50	-52	28
	R Middle Temporal	719	4,136	60	-48	8

	Gyrus					
	R Cerebelum (VI)	160	6,246	16	-76	-8
	L TPJ	853	6,121	-58	-56	18
	L Inferior Parietal Lobule	853	5,783	-52	-54	46
	R IFG (p. Orbitalis)	747	5,950	30	20	-20
	Location not in atlas	747	5,339	48	36	-18
	R Insula Lobe	747	4,321	44	16	2
	R dmPFC	188	5,699	6	38	38
	R Precuneus	343	5,367	12	-62	30
	L Precuneus	343	3,550	-6	-66	38
	R Middle Temporal Gyrus	202	5,032	50	-20	-10
	L IFG (p. Orbitalis)	117	4,847	-28	12	-18
	R Middle Frontal Gyrus	192	4,744	36	4	46
Pride > Gratitude	R Caudate Nucleus	201	6,688	20	-2	24
	Location not in atlas	201	4,918	18	24	14
	R Hippocampus	131	5,870	28	-40	6

Note: abbreviations: R: right; L: left; TPJ: temporal parietal junction; dmPFC: dorsomedial prefrontal cortex; IFG: inferior frontal gyrus. Whole brain Family-Wise Error (FWE) corrected at the cluster level.

In contrast, the brain regions such as the right caudate nucleus (peak MNI 20 -2 24, t = 6.69, 201 voxel) and the right hippocampus (see Figure 4 & Table 2) were more engaged in the pride condition.

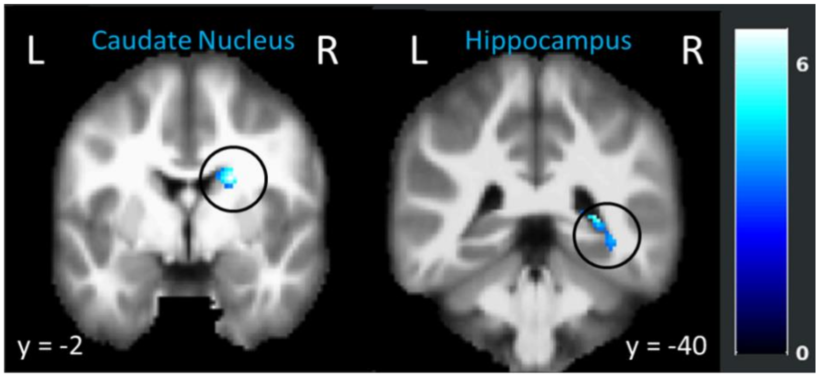

Figure 4. Contrast of pride > gratitude conditions

3.2.2 Neural substrates of subjective feelings of gratitude and pride

Next, we would like to characterize the neural substrates accounting for the moment-to-moment fluctuated subjective feelings of gratitude and pride. To this end, we regressed BOLD activity during the time when participants had the strongest emotional response (when they saw whether their answers were correct or not) on subsequent ratings of gratitude or pride.

We expect that gratitude feeling may relate with brain regions involving Theory of Mind as stated above, reward (e.g. striatum, OFC), and emotion (e.g. insula). On the other hand, pride may relate with regions involving self-referential processing as stated above, reward (e.g. striatum, OFC) and emotion (e.g. insula).

Parametric Modulation by Gratitude. As a result, brain regions positively modulated by the later gratitude rating mainly involved some parts in the temporal cortex, striatum and motor cortex (see Figure 5 and Table 3). As expected, the more grateful participants felt, the more brain activity was found in the temporal pole (peak MNI 56 8 -2, t = 4.03, 281 voxel) and bilateral superior temporal gyrus. Two subcortical regions in striatum: the right ventral striatum (peak MNI 10 6 -12, t = 5.85, 118 voxel) and the left putamen (peak MNI -28 -6 2, t = 5.28, 96 voxel) were also positively modulated by gratitude rating. Unexpectedly for the motor cortex, higher BOLD signals in the left precentral gyrus (peak MNI -36 -20 64, t = 5.53, 1651 voxel) extending to postcentral gyrus were associated with feeling more grateful.

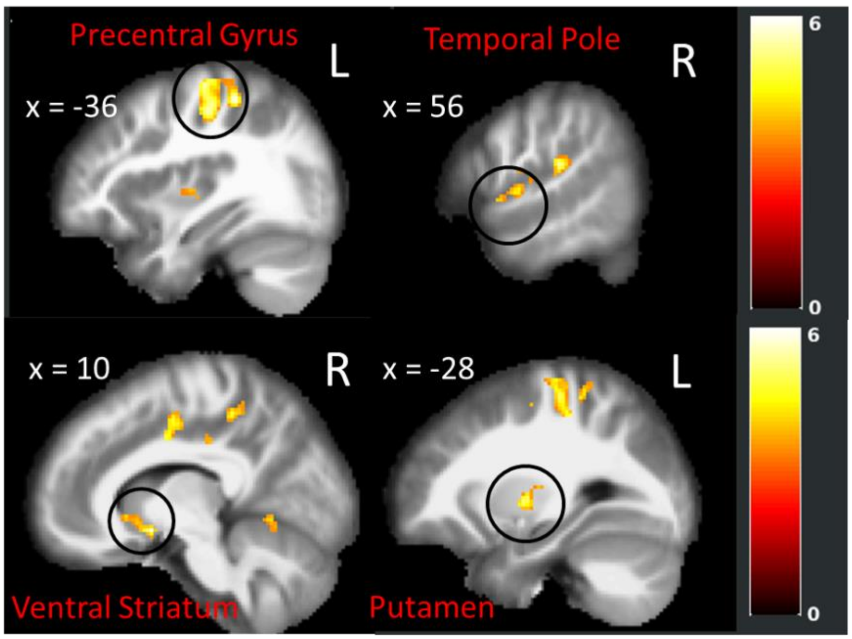

Figure 5. Brain regions positively modulated by gratitude ratings

Table 3 Neural correlates of gratitude ratings

Contrast Name				MNI Coordinates		
	Region Label	**Extent**	**t-value**	**x**	**y**	**z**
	L Postcentral Gyrus	1651	6,124	-34	-38	54

Positive	Location not in atlas	1651	5,607	-6	-30	36
	L Precentral Gyrus	1651	5,532	-36	-20	64
	R Superior Temporal Gyrus	218	6,030	54	-28	18
	R Ventral Striatum	118	5,847	10	6	-12
	Location not in atlas	118	4,242	20	24	2
	L Putamen	96	5,277	-28	-6	2
	L Superior Temporal Gyrus	118	5,231	-56	0	4
	Location not in atlas	281	5,125	38	-8	0
	R Rolandic Operculum	281	4,589	48	-16	16
	R Temporal Pole	281	4,031	56	8	-2
	R Cerebelum (VI)	231	5,110	22	-56	-22
	Cerebellar Vermis (4/5)	231	4,655	4	-58	-4
	R Postcentral Gyrus	138	5,065	24	-38	68
	R MCC	92	4,823	10	-2	42
	R Middle Temporal Gyrus	93	4,657	48	-68	4
	R Insula Lobe	1416	-7,194	36	20	2
Negative	R IFG (p.	1416	-5,889	50	22	26

Triangularis)					
Location not in atlas	1416	-5,410	34	6	30
L Middle Occipital Gyrus	333	-6,493	-30	-74	32
Location not in atlas	333	-6,010	-28	-52	38
R Precuneus	532	-6,274	4	-58	48
R Superior Parietal Lobule	532	-3,889	22	-70	58
L IFG (p. Opercularis)	641	-6,071	-48	12	28
L Middle Frontal Gyrus	641	-5,172	-40	2	56
L IFG (p. Orbitalis)	389	-5,856	-38	14	-8
R Middle Occipital Gyrus	108	-5,692	32	-80	38
L Superior Medial Gyrus	404	-5,519	0	22	54

Note: abbreviations: R: right; L: left; IFG: inferior frontal gyrus. Whole brain Family-Wise Error (FWE) corrected at the cluster level.

Besides, we also found higher gratitude rating was associated with decreased neural activities in right insula (peak MNI 36 20 2, t = -7.19, 1416 voxel) (see Figure 5).

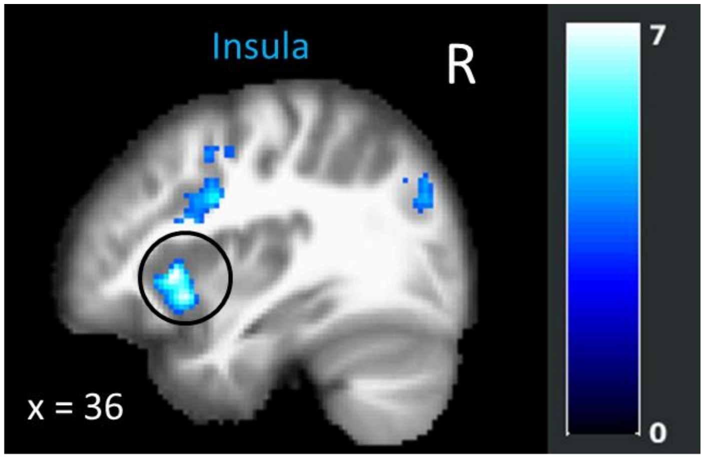

Figure 6. Brain regions negatively modulated by gratitude ratings

Parametric Modulation by Pride. The more proud people felt of themselves, the higher BOLD signals were found in the two clusters in the striatum—bilateral putamen (see Table 4, Figure 7). The cluster in left putamen was not overlapped with the one parametrically modulated by gratitude rating, but located more anterior to it. We did not find any regions involving self-referential processing such as mPFC, Precuneus, and PCC nor basic emotion such as insula.

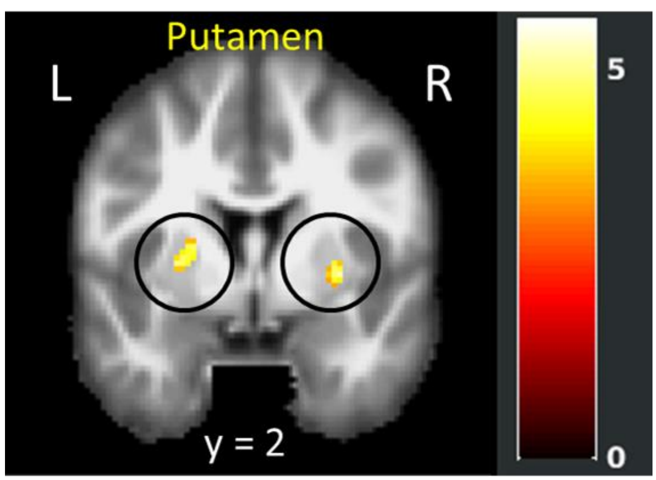

Figure 7. Brain regions positively modulated by pride ratings

Table 4 Neural correlates of pride ratings

Contrast Name				MNI Coordinates		
	Region Label	**Extent**	**t-value**	**x**	**y**	**z**
Positive	R Putamen	76	5,627	24	-4	6
	L Putamen	74	4,811	-28	2	4
Negative	None					

Note: abbreviations: R: right; L: left; whole brain Family-Wise Error (FWE) corrected at the cluster level.

3.2.3 Common value updating mechanisms of gratitude and pride in mPFC

Since we had observed that the brain is tracking the moment-to-moment subjective feelings of gratitude and pride, we wanted to explore further how the brain updates the values of these feelings. Based on our results that both gratitude and pride modulate the activities in striatum, which represented social as well as monetary reward (Izuma, Saito, & Sadato, 2008; Lin, Adolphs, & Rangel, 2011), we hypothesized that the brain updates the feelings of gratitude and pride in a similar way as updating reward, through a common neural substrate. Therefore, we calculated the differences in the gratitude ratings and pride ratings trial by trial and took it as the error signal, and then regressed the whole brain BOLD activity also during the time when participants saw the reward on it. We expected to find brain regions representing prediction error signals such as the ventral striatum, ACC, mPFC or OFC to be the neural substrates for updating gratitude and pride.

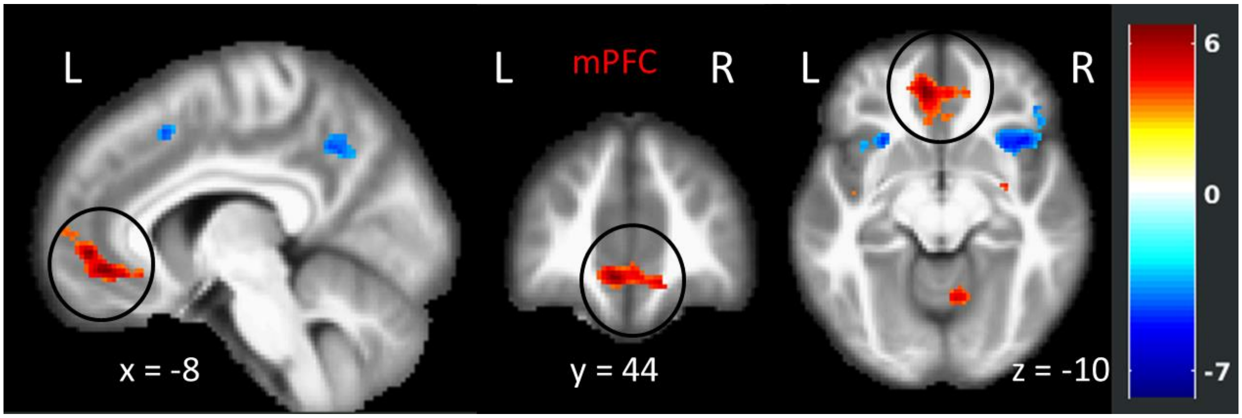

Figure 8. Parametric modulation of the general updating values in gratitude and pride

We found one large cluster in the mPFC (Figure 8) with two local peaks (MNI -8 44 -10, t = 6.70 and MNI -4 56 6, t = 5.01, 486 voxels) which positively predicted later changes in gratitude and pride ratings. The higher the BOLD signal was in the mPFC when participants saw the reward, the greater increase was found in the ratings of gratitude or pride later in the trials. The other clusters can be seen in Table 5.

Since gratitude may use more other-referential information and pride may use more self-referential information to update the feelings, we explored further how specifically the value updating in gratitude and pride were represented in mPFC. For this purpose, we plotted the neural correlates of GGi and PPi within the mPFC cluster we had acquired in the previous step for better visualization (Figure 9).

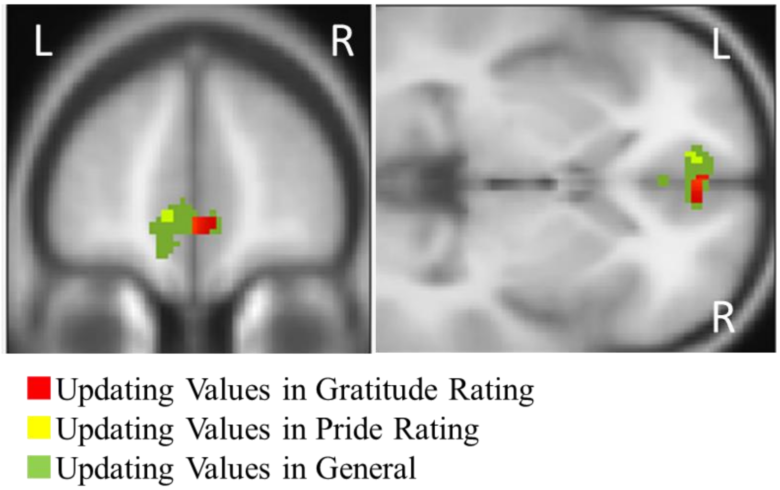

Figure 9. Separate plots for parametric modulations of updating values in gratitude and pride ratings compared with the prior-defined mPFC

We observed a clear dissociation for gratitude and pride updating in the mPFC. The neural correlates of updating values in gratitude ratings and pride ratings in the mPFC were adjacent to each other, but not overlapped. Value updating in gratitude was more closely associated to the medial part in the right hemisphere, while value updating in pride was more closely related to the lateral part in the left hemisphere.

Table 5 Neural correlates of updating values in gratitude and pride

Contrast Name				**MNI Coordinates**		
	Region Label	**Extent**	**t-value**	**x**	**y**	**z**
Positive	L mPFC	486	6,697	-8	44	-10
	L mPFC	486	5,007	-4	56	6
	L Middle Frontal Gyrus	178	5,950	-28	24	48

	Cerebellar Vermis (4/5)	187	5,699	6	-62	-12
	R Cerebelum (VI)	187	4,343	26	-50	-22
	R SupraMarginal Gyrus	587	5,658	60	-30	28
	R Superior Temporal Gyrus	587	4,956	54	-4	2
	R Superior Temporal Gyrus	587	3,889	58	-24	8
	R Putamen	203	5,077	28	-8	6
	R Middle Cingulate Cortex	96	5,063	8	-4	42
	L Insula Lobe	139	5,006	-40	-4	-2
	L Rolandic Operculum	97	4,873	-48	-20	20
	L Precentral Gyrus	246	4,859	-34	-28	62
	R Postcentral Gyrus	112	4,518	30	-38	60
Negative	R Insula Lobe	592	-7,993	34	18	2
	R IFG (p. Orbitalis)	592	-4,532	44	36	-12
	R IFG (p. Orbitalis)	592	-4,021	30	10	-16
	L Posterior-Medial Frontal	595	-7,917	2	16	50

Location not in atlas	295	-6,321	-28	-54	38
L IFG (p. Orbitalis)	342	-6,301	-30	22	-4
R Precuneus	259	-6,267	4	-60	48
R IFG (p. Triangularis)	543	-6,053	38	12	30
Location not in atlas	543	-4,485	32	2	48
R Middle Occipital Gyrus	122	-5,688	32	-72	34
Location not in atlas	142	-5,355	-28	-68	34
L IFG (p. Opercularis)	389	-5,314	-48	18	38
Location not in atlas	389	-4,354	-34	10	26

Note: abbreviations: R: right; L: left; mPFC: medial prefrontal cortex; IFG: inferior frontal gyrus. Whole brain Family-Wise Error (FWE) corrected at the cluster level.

3.2.4 Computational models of gratitude and pride

Having found the neural markers of updating our feelings on gratitude and pride, we further analyzed what environmental and psychological factors influence the fluctuations in gratitude and pride. Specifically in our study, we postulate seven factors that may influence the brain to update gratitude or pride in Table 1. Using all possible linear combinations of these factors, we generated 127 models for gratitude and 127 models for pride. We ranked the models by their BIC values and took the ones with the smallest BIC as the best models for gratitude and pride because they explained more variance with fewer model factors.

The Model of Gratitude. The model that captured participants' moment-to-moment subjective feelings of gratitude with the smallest BIC was as follows:

$$Gratitude(i) = w_0 + w_1 * Reward(i) + w_2 * EV_b(i) + w_3 * RPE_b(i) + w_4 * Certainty(i)$$

In this model, five factors contributed to the feeling of gratitude: (1) a constant factor (w_0) which represented each subject's baseline level of gratitude feeling; (2) the reward gained in each trial; (3) the expectation of the benefactor EV_b (whether the audience would help them or not); (4) the prediction error of the benefactor RPE_b (the difference between expectation of the benefactor and the actual helping behavior of the benefactor); and (5) certainty (how certain participants felt about their own choice before receiving the help from the audience), which was an indicator of how much help they need. Combining the influence of these factors, we found the model fitting the dynamics in gratitude ratings well (seen Figure 10).

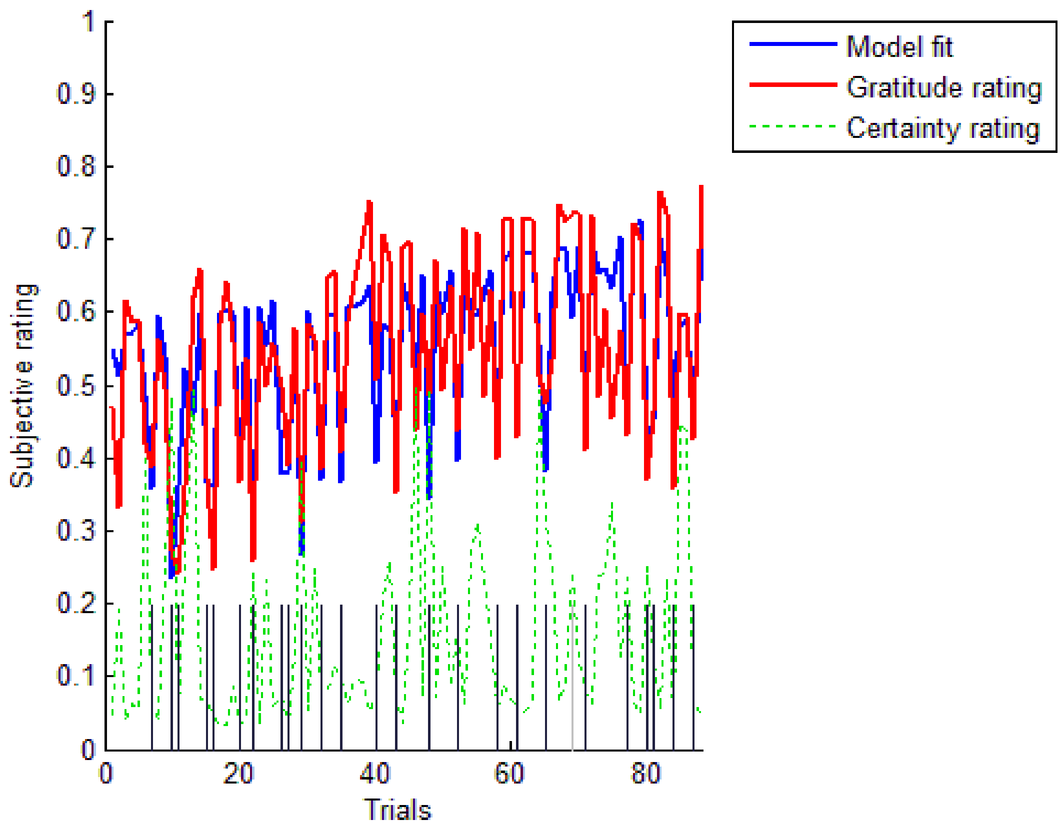

Figure 10. The gratitude model fit in an example participant

However, it's worth noticing that the expectation for reward itself (EV_r), as well as the prediction error of the reward (RPE_r; the differences of participants' actual gain or loss compared to the expectation of the upcoming reward) were not in the optimal model to explain or predict the feeling of gratitude.

The Model of Pride. The computational model of pride was even simpler than that of gratitude. It could be formalized as bellow:

$$Pride(i) = w_0 + w_1 * Reward(i) + w_2 * Certainty(i)$$

Here, three factors made major contribution to the feeling of pride: (1) the constant factor w_0, which varied between individuals as the baseline level of pride; (2) the reward gained in each trial; and (3) certainty (how certain they were about each question before answering it), which reflects the difficulty or the value of the current trial. The more uncertain people felt, the more difficult the task was, thus more rewarding for people to solve it. Neither the expected future reward nor the reward prediction error significantly influenced pride. From Figure 11 we can see the influence of reward and certainty on the fluctuations of pride, and this model fitted the pride ratings reasonably well.

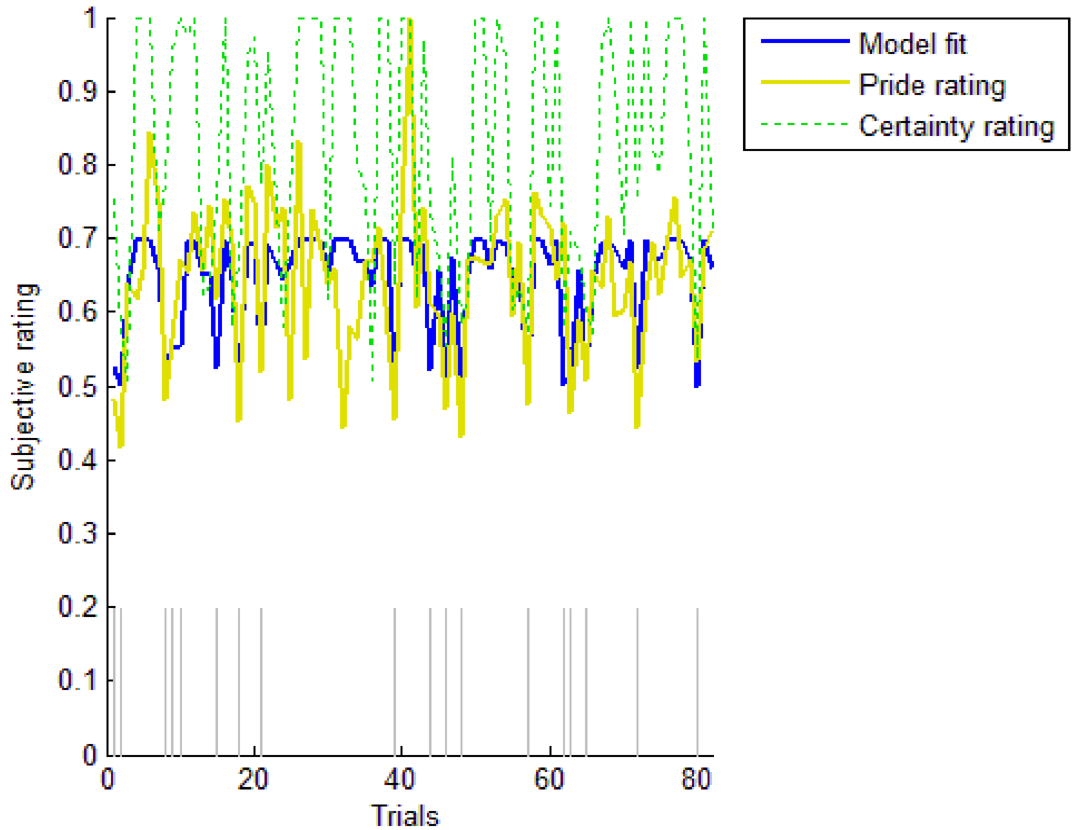

Figure 11. The pride model fit in an example participant

Through the models above, we found that the dynamics in gratitude and pride were influenced by the common factors of reward and certainty, and they were also affected by unique factors such as individual differences in the baseline level of gratitude and pride, which could arise from personality traits, genes and social cultural environment. Importantly, as we can see in the gratitude model, people were learning the actions of the benefactor through prediction error, but not the reward itself.

4. Discussion

At the frontier of social and affective neuroscience, complex social emotions such as gratitude and pride have recently being investigated with functional neuroimaging. However, given the relative few studies, little is known the neural substrates underlying the dynamics of gratitude and pride over time, nor their differences in the neural substrates and underlying computational mechanisms. Discovering the differences of gratitude and pride in the neural substrates and computational mechanisms will deepen our understanding of these emotions, and facilitate developing new interventions to promote these positive emotions and motivate prosocial behaviors in everyday life.

In the current study, we applied a novel social interactive fMRI task and computational modeling to track the fluctuations in gratitude and pride in the brain and behavior. The results showed that, gratitude compared with pride conditions involved more brain regions processing Theory of Mind, such as the TPJ, dmPFC, IFG and precuneus; while pride compared with gratitude conditions involved more brain regions processing reward, such as the caudate nucleus, and memory such as the hippocampus. Specifically for the neural substrates accounting for the subjective feelings of gratitude, higher level of gratitude was associated with higher activation in the motor cortex such as the precentral gyrus, the reward system such as the ventral striatum and putamen, and the Theory of Mind network such as the temporal pole; while lower level of gratitude was associated with lower activation in the insula. In contrast, for pride, only positive correlation of subjective rating and the activation in bilateral putamen was found. Furthermore, the brain updated the values of gratitude and pride primarily with the mPFC, an area which is known for learning reward in various forms and a hub for processing information about others and the self in general. The computational models that best explained the dynamics in gratitude and pride both included parameters of a constant factor accounting for the mood, reward, and certainty. Gratitude in particular involved a learning process of the benefactor's behavior, not the reward itself.

In the following sessions we will mainly discuss these findings comparing to previous studies.

4.1 Gratitude vs. pride

Positive emotions have long been considered less differentiated than negative emotions (Smith, Tong, & Ellsworth, 2014). For example, the broad-and-build theory believed that all the positive emotions have the similar elevated positive subjective experience and the motivational action tendency to broaden cognitive functions and build resources. However, even though gratitude and pride are both positive emotions related with reward or success, they seem to have different attributions and motivational urges. People attributing success to external sources (e.g. others' help, good opportunity) will experience a feeling of gratitude, while people attributing success to internal sources (e.g. one's ability, personality) will experience a feeling of pride. People who feel grateful are more likely to engage in prosocial behaviors, while people who feel proud are more likely to engage in rewarding self and showing off (Smith et al., 2014). Therefore, we hypothesized that gratitude and pride may have different neural substrates.

Comparing gratitude with pride conditions in the brain reveal that, gratitude involves more brain regions which are thought to be typical Theory of Mind functional regions, while pride involved brain regions more related with reward and memory. Theory of Mind is the ability to understand others' actions, beliefs and intentions, and functional neuroimaging studies have found the neural substrates of the Theory of Mind including TPJ, temporal pole, IFG, dmPFC and precuneus/PCC (Saxe & Kanwisher, 2003; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014; Van Overwalle & Baetens, 2009). More specifically, the bilateral TPJ were selectively involved in understanding others' thoughts rather than general social information such as bodily sensation and appearance (Saxe & Powell, 2006). In our study, the bilateral TPJ, IFG, dmPFC and precuneus were more engaged in the gratitude condition. It indicates that, to feel grateful, one has to have a good knowledge of others' mental states to understand the good intentions behind such actions. Meanwhile, understanding benefactor's different beliefs, for example, in our paradigm, understanding that the "audience" may give a piece of wrong advice which she thought to be right, will help people keep a grateful mind. In contrast, pride was associated with stronger brain activities in the caudate nucleus, which is part of the reward system. The caudate nucleus is associated with short-term reward, and is one of the main brain regions for reward-based behavioral learning (Haruno et al., 2004). Increased BOLD signals have also been found in the caudate nucleus when people reported the euphoric "high" feeling during acute cocaine intoxication (Risinger et al., 2005). Therefore, the neural activities found in the caudate nucleus indicate that pride is an intrinsically rewarding and pleasant feeling. In addition, the hippocampus was also more engaged in the pride condition. The memory function of hippocampus is well-documented (Bliss & Collingridge, 1993; Burgess, Maguire, & O'Keefe, 2002). In the pride condition, without others' help, participants must remember what they did to achieve the task; while in the gratitude condition, they only need to accept the help from the helper.

Our study is the first evidence of specific neural activities associated with gratitude and pride conditions. An early study (Zahn et al., 2009) on social values had both gratitude and pride conditions in the same paradigm: subjects were presented with one sentence verbal statements descripting their actions towards their best friend either confirming or countering to a social value, or their best friend's action towards them. For example, "Sam acts generously towards Tom", and "Tom acts stingily towards Sam". In the fMRI experiment, subjects were simply asked to judge their feelings as pleasant or unpleasant. After the experiment, they were asked to classify these verbal statements into self-other agency and positive-negative categories. Later researchers (Kini et al., 2016) think the paradigm is more about identifying emotions in a social narrative, rather than experiencing emotions. What's more, in this study no significant group effects were found specific to gratitude when contrasting gratitude with pride or indignation/anger conditions. This could be due to the less emotional involvement in the paradigm and it could also be seen from the overall small effect size in all the fMRI results.

4.2 Gratitude and Theory of Mind

Our results were in line with previous gratitude research which found that Theory of Mind processing was involved given the evidence in the dmPFC (Fox et al., 2015) and the left SPL (Kini et al., 2016). The left SPL is close to our finding in left TPJ. Moreover, we also found strong right TPJ involvement. The review on neural basis of Theory of Mind have highlighted that although both mPFC and TPJ are reliably activated during Theory of Mind tasks, the mPFC plays a general role in integration and prospection, while the TPJ selectively responds to mental state information across many recent studies (Mahy, Moses, & Pfeifer, 2014). Therefore, our study provides strong evidence of the Theory of Mind processing preferentially engaging in the gratitude condition.

4.3 Gratitude, prosocial intentions, reward, and Theory of Mind

Self-reported feelings of gratitude were positively correlated with neural activities in the motor cortex, reward system and Theory of Mind network, but negatively correlated with neural activities in brain regions representing negative emotions. In the motor cortex, the left precentral gyrus is also reported in previous gratitude studies, and it is related to the desire to help (Kini et al., 2016). When people feel grateful, they have a greater tendency to act proscially as a return for kindness. In consistent with a lot of behavioral studies (Bartlett & DeSteno, 2006; Tsang, 2006), gratitude induced more prosocial behaviors than neutral and even other positive emotions. Therefore this could also explain that, the more grateful people felt, the greater activation was found in the motor cortex. For the brain reward system, we had two regions positively modulated by gratitude: the right ventral striatum and the left putamen. The ventral striatum receives mid-brain dopamine projections, and is associated with reward prediction and reward prediction errors during reinforcement learning and gambling tasks (J. O'doherty et al., 2004; J. P. O'Doherty et al., 2003; Pessiglione et al., 2006). It indicates that gratitude may involve a learning process, rather than a simple hedonic response to reward. The left putamen parametrically modulated by gratitude is located more posterior to the left putamen cluster parametrically modulated by pride rating. The posterior putamen is associated with habit learning (Tricomi, Balleine, & O'Doherty, 2009) and memory-guided movement (Menon, Anagnoson, Glover, & Pfefferbaum, 2000). In our study, when participants chose to take the benefactor's advice and got rewarded, the activation in the posterior putamen helped the participants to memorize the event and form the habit of taking the benefactor's advice again in the future. The temporal pole is part of the Theory of Mind network, and it plays an important role in understanding others' action intention (Den Ouden, Frith, Frith, & Blakemore, 2005). This is also consistent with our previous finding that the other brain regions in the Theory of Mind network such as bilateral TPJ and dmPFC are more involved in the gratitude condition, and also in previous studies of gratitude (Fox et al., 2015). In addition, we also find a negative modulation of gratitude rating on the right insula. This is also consistent with previous studies (Fox et al., 2015; Yu et al., 2017). Feeling more grateful was correlated with decreased insula activities—this reflects a positive effect of gratitude on reducing negative emotions.

4.4 Pride and reward

Self-reported feeling of pride was associated with activities in the bilateral putamen, which are part of the brain reward system. Previous pride studies also reported the involvement of reward regions such as in the OFC (Gilead et al., 2016), but not in the putamen. Clinical studies have shown that a lack of putamen activity is linked to a lack of positive emotions: one study showed that compared with healthy controls, patients with major depression had significantly lower activation in the putamen during reward anticipation (Pizzagalli et al., 2009). Another recent study on childhood chronic fatigue syndrome has shown that lower putamen activity was correlated with lower reward sensitivity and more fatigue (Mizuno et al., 2016). On the other hand, a structural MRI study on general self-efficacy, which is the overall confidence of one's general ability to achieve on tasks, conceptually close to pride, discovered that the low mean diffusivity (which means higher neuronal density and better functioning) in putamen was associated with high general self-efficacy values (Nakagawa et al., 2017). Furthermore, as part of the dorsal striatum, the putamen is densely connected to various motor regions and generally involved in motor learning (Doyon, Penhune, & Ungerleider, 2003). In particular, it encodes stimulus-action-reward association (Haruno & Kawato, 2006). It is likely that, in pride situations, the activity in the putamen encodes our rewarded actions to enable and motivate us to repeat the achievements again in the future. This effect could be seen in previous behavioral studies in which participants experiencing pride feelings on a success task can work longer on a subsequent similar task (Williams & DeSteno, 2008).

4.5 Pride and self-referential processing

Previous studies on pride have highlighted the central midline structures like the mPFC, PCC, and precuneus (Gilead et al., 2016; Roth et al., 2014; Simon-Thomas et al., 2012). However, we did not find strong posterior midline involvement in the pride condition. One reason could be that previous studies used narratives which required participants to put themselves in the scenarios far from their own experience, or recall their own success moments. This may activate the mPFC due to a perspective-taking processing, and the precuneus and PCC as of more basic functions in visual-spatial imagery and episodic memory retrieval (Cavanna & Trimble, 2006; Northoff et al., 2006). While in our study, pride is a direct emotional response to the current successful event, without the potential confounding of perspective-taking and past memory retrieval.

4.6 The common neural currency in the mPFC

The mPFC is well known as a part of the "reward circuitry" (Russo & Nestler, 2013), especially its role in representing reward values and reward prediction errors (Knutson & Cooper, 2005). It has been shown that the mPFC represent different kinds of values, such as food, money, and other nonfood consumables (Chib et al., 2009).

Similar to tracking the differences in reward, we have observed that the mPFC tracks the changes in gratitude and pride, which are positive reinforcers for social attachment and self-enhancement. Unlike the dynamics in happiness, which primarily involves the ventral striatum (Rutledge et al., 2014), the dynamics in gratitude and pride involve evaluating the subjective values regarding others and oneself. In social learning, the prediction errors of others and self are both tracked at the ventral part of mPFC (Joiner, Piva, Turrin, & Chang, 2017). Meanwhile, as the mPFC is a hub involved in both Theory of Mind (Van Overwalle & Baetens, 2009) and self-referential processing (Gusnard, Akbudak, Shulman, & Raichle, 2001; Northoff et al., 2006), it could provide enough information about others and self to support the value judgment and learning process. The finding suggests a domain general role of mPFC—it is not only a common currency for values, but also supports the changes for higher order social emotions.

4.7 The predictive model of gratitude and the reactive model of pride

Emotions have long been considered as responses to certain events. They either serve as a functional signal to threats or rewards in the environment, or the products of appraisals, such as attributions and evaluations of the environment relevant to one's beliefs and goals. However, recently the constructed emotion theory (Barrett, 2017) proposed that emotions are not reactive, but predictive. The brain works as a Bayesian filter: people form expectations about how things should be based on their prior experience, and then use expectations to filter and process the unexpected sensory stimuli to generate feelings. This theory matches with our model of gratitude very well, but not with pride. People form expectations from interacting with the benefactor about how likely they would help them and generate an internal model to predict the others' behaviors. When the outcome reveals, people generate the strongest feelings, but the feeling is not only a reaction to the outcome, but biased by expectations.

From our computational model of gratitude, we can see that the updating process is influenced by the learning process of the actions of the benefactor, rather than the reward itself. In particular, we discovered the important roles of both the expectation and prediction error of the benefactor's action in contributing to the dynamics of gratitude. When we are unexpected helped by someone, we generate a positive prediction error and feel grateful, which reinforces us to assign social value to this person and motives us to build up a relationship. If this person helps us again and again, we gradually decrease the prediction error and increase the expectation, which is estimation of how likely this person is going to help us in the future. We feel grateful when we believe someone is going to come and help us, even before the help actually takes place. Think about a family member or close friend: even when he or she does nothing for us at the moment, we still feel grateful to this person. Yet the role of expectation in gratitude is like a double-edged sword as high expectation will dampen the effect of a positive event or even cause a negative prediction error when an expected help is absent or below the expectation. That explains why we take the help from our family and close friends for granted and sometimes even become ungrateful when they offer less than we expect.

These findings in expectation and prediction error add a new dimension to traditional psychological theories on gratitude. According to McCullough (2001), people would feel grateful if they receive help that is valuable, of high efforts/cost, intentional, and gratuitous. Wood and his colleagues (2008) developed this idea and proposed a social-cognitive theory of gratitude. In this theory, benefit appraisal is the cause of state gratitude (in contrast to trait gratitude), and the appraisal includes the evaluation of benefit in terms of its value, cost, and genuine helpfulness. However, we notice that over a long period of time, people do not keep feeling the same level of gratitude towards the same valuable benefit given by the benefactor with good intentions. We constantly compare how people treat us with an expectation "how they should be", which makes us grateful for a small favor from a stranger, while less so intensive to the same favor offered by an old friend. While with time goes on, we have smaller prediction error about a person's behavior and the expectation is becoming stable. Then we can base on the expectations to decide who we can count in the future. From a functional perspective, this computational process on a micro level may benefit us in social life: the updating in the expectations and prediction errors of the benefactors modify the social values we assign to the people we interact with, so that we can identify and invest more in building up beneficial relationships with those significant others.

In addition, we find the contribution of the certainty factor from the computational model is consistent with previous studies. The subjective value of each question is reflected by participants' rating how certain they are about the choice. It acts as an independent model factor besides the real monetary reward to influence both gratitude and pride. Even though the object reward for each correct answer remains constant at one euro, in a gratitude situation, the less certain people are about their choice, the more they need the help, making the help more valuable. Therefore people feel more grateful after receiving reward. The evaluation of subjective value is an important component in benefit appraisal, and significantly influences the gratitude level (Wood, Maltby, Stewart, et al., 2008). In pride situations, more certainty people feel about the tasks(less difficult) means less challenge and lower value, thus making the success less rewarding when people solve it without help. This finding was consistent with previous studies which showed that greater pride responses were observed when people get success on difficult tasks than on easy tasks (Belsky et al., 1997; Lewis et al., 1992).

4.8 Limitations and further directions

Our study has a few limitations. First, we do not have a large enough data set to test and replicate our behavioral models. As the limited scanning time in fMRI, and the effect of emotional fatigue after repeated trials, we have 120 trials in total, which are not strong enough to establish very solid computational models. Further studies are needed to replicate these findings. Moreover, there are more factors in real life that could influence the feelings of gratitude and pride. Therefore, the extent to which we can explain from our models is limited. Second, in our study, the participants only met the benefactor twice; this kind of short-term relationship may yield different model weights of gratitude from that of the long-term relationship. Third, there are cultural differences in gratitude and pride. Even though we control this by only taking all the participants from the western culture, there could be still a little cultural difference between different countries. For example, some German participants reported feeling rather reluctant to admit that they were "proud" of themselves as they thought it would be similar as being arrogant, while the American participants had rather positive attitude reporting to be proud. Fourth, the definitions of both gratitude and pride are broad: people may feel grateful towards the beautiful nature or they could feel proud for others or their country, therefore we cannot generalize our models and neural correlates to these definitions from these different perspectives. Fifth, in our fMRI experiment, we asked the participants to directly report on their feelings of gratitude and pride. The method of self-report may lead to participants reporting more socially desirable feelings such as gratitude. However, if we take an indirect measurement (donating money or other prosocial behaviors), it is hard to explain the motivations behind the behaviors because the causes could be something else than the emotions.

Base on the current study, there are a few interesting follow-up studies we can explore further. First, we can see if the functional connectivity between brain regions in the reward system and the Theory of Mind network could predict gratitude. There has been evidence that children with autism spectrum disorders (ASD) have both deficits in Theory of Mind and reward learning (Scott-Van Zeeland, Dapretto, Ghahremani, Poldrack, & Bookheimer, 2010), and the frontal-striatal neural activities were specially abnormal in social reward learning. The ASD patients suffer from avoiding social interaction and expressing interpersonal emotions. The new disease model of ASD believes that the social problems in ASD patients mainly stem from the deficit in social reward learning, rather than the the traditional belief in Theory of Mind. We hypothesize that, gratitude, serving as the "social glue" in interpersonal relationship, could be modulated by the connectivity by the strength of connectivity between the reward system and the Theory of Mind network. Second, we can collect large behavioral data online or in real life to validate our models of gratitude and pride. For example, if we have 300 subjects, we can use half of the data set to build the models and then test them on the second half. Besides, we can also test if there is a difference in the gratitude model when subjects meet the benefactor once and when they meet and play with the benefactor for multiple times. Furthermore, we could do intervention studies. Previous gratitude interventions normally asked subjects to recall things they feel grateful for. However, according to our model, appreciating rewarding experience only influences gratitude to a certain degree: the expectation of others and the subsequent prediction error contrasting to reality also play important roles.

Therefore, adjusting the expectation of others could be another promising way to increase gratitude.

5. Conclusions

In summary, our research identified the specific neural substrates of gratitude and pride. Brain regions in the Theory of Mind network were more involved in processing gratitude, while brain regions related with reward and memory were more involved in processing pride. Furthermore, the neural substrates of updating the positive emotions of gratitude and pride were also located in the mPFC, which is a common neural currency for different kinds of reward as well as a social brain hub. Moreover, computational models showed that the reward and certainty both contributed to feelings of gratitude and pride, but the predictive coding of the benefactor's actions was specific to gratitude.

6. General discussion

What is the essence of emotions? Do different emotions have unique biological basis? Why do emotions drive us to approach or avoid things? These questions have motivated scientists to explore the nature of emotions from different perspectives generation by generation.

From the early evolutionary view such as seen in Darwin's *On the Origin of Species*, emotions are considered to serve as signals to environmental changes. Emotions help individuals to react quickly and be more adaptive in the society. As different emotions are triggered by different stimuli and lead to different adaptive functions, such as fear signifying danger, and sadness signifying loss, more contemporary functional theorists follow this idea and look for unique biomarkers of different emotions. For example, Paul Ekman has identified unique facial expressions of six basic emotions (anger, disgust, fear, happiness, sadness, and surprise) across different ethnics and cultures. However, all positive emotions are less differentiated compared to negative emotions, although they have different conceptual constructs and serve different functions. For example, gratitude helps us to identify important benefactors and build up meaningful relationship while pride helps us to identify how well we can do on certain things so that we can better choose the tasks we are good at. Meanwhile, expressing gratitude can improve interpersonal relationships and expressing pride can improve our social status perceived by others. Nevertheless, such conceptual differences have little corresponding biological evidence.

Our study is the first to identify the largely differentiated functional neural substrates of gratitude and pride. Gratitude involves large brain regions processing reward and the Theory of Mind, while pride involves mainly the reward and memory system. This is an important piece of evidence to support the functional perspective, especially on the positive emotions spectrum, which is generally blurred. In addition, we have also found different behavioral models for the subjective feelings of gratitude and pride.

Furthermore, the functional perspective of emotions has a loophole: how does the brain know what and to what extend we should react to the environmental stimuli? The same stimuli may cause different people react differently. For example, seeing a dog makes some people happy while others afraid. Each individual's past experience influences how one sees the stimuli as pleasant or threating. Therefore, psychologists developed the appraisal theories to explain how people interpret the stimuli.

According to the appraisal theories, emotions are not only reactions to environmental stimuli, but the products of a subjective and meaningful "analysis" or "evaluation" in terms of one's own goals, needs, desires and abilities. According to the appraisal theories, there are three benefit-related emotions: happiness, gratitude and pride (Smith et al., 2014). Happiness is the general elevated emotion for achieving one's goals or satisfying one's needs and desires. The attribution of reward to external source such as others' help is linked to gratitude while to internal source is linked to pride. From this distinction, we expected to find a self-other neural dissociation when comparing pride with gratitude conditions. However, there is no neural evidence so far to support the obvious hypothesis. We have found that the "Others-related" brain regions largely involved in gratitude conditions, while the typical "Self-related" regions were absent from the pride conditions. It could be that event-evoked pride is automatically generated rather than going through a slow self-evaluative process. This is consistent with previous studies of pride which considered pride as a basic emotion given that it has unique and crosscultural facial expressions and body gestures (Tracy & Robins, 2004b; Tracy et al., 2010).

Moreover, with our computational models, we can see the clear difference between gratitude, pride, and happiness. In Rutledge (2014)'s study, Happiness = w_0 (constant factor) + R (certain reward) + EV_r (the expectation of reward) + RPE_r (the prediction error of reward). In our study, Gratitude = w_0 + R + EV_b (the expectation of the benefactor) + RPE_b (the prediction error of the benefactor) + Certainty. There we can see that although happiness and gratitude both involve the prediction of the future, happiness tracks the reward itself while gratitude tracks the benefactor's action. Being sensitive to others' help is important for social survival. Building a proper expectation of others will help people to identify and build up long-term relationship, while through the prediction error we can adjust the expectations to be more adaptive. Whereas, the model of pride is more concise, pride = w_0 + R + Certainty. We can see that comparing to gratitude and happiness, pride does not have to involve a learning process of the reward or other people. What matters is the actual success experience. This algorithm helps individuals to have a more accurate estimation based on facts rather than expectation. As too much pride is also considered hubristic and socially negative, this algorithm keeps people to be socially adaptive. There could also be a reason that the learning process mainly happens in childhood, as our participants are all adults, they have already a rather stable self-evaluation of their own abilities and knowledge. By comparing the three positive emotions on a computational level, we can see how they are updated in different ways. This new contribution will take us to a deeper understanding of the nuances of positive emotions.

Our study also has an implication that some emotions are predictive, not reactive. This is consistent with the newest emotional theory: the constructed emotion theory

(Barrett, 2017). The theory thinks human brain works like a Bayesian filter, we form expectations and then everything happens to us causes a reaction to our expectation, rather than the reality. This is different from the appraisal theory, which thinks emotions are generated by interpretations of the reality. We can see from our fMRI studies and behavioral models that the emotion happen quickly, and people feel the strongest emotions in the moment of the reward reveal, not during contemplating emotions afterwards. Our studies partly support the constructed emotion theory, in terms of the predictive process of gratitude. However, we cannot deny the appraisal theories and the functional theories based on limited evidence.

Our study also has further clinical implications. We could develop a new therapy based on our findings. We think that the Cognitive behavioral therapy is not efficient enough. As the appraisal theories are the cornerstones of the main stream Cognitive Behavioral Therapy, and the treatment effect mainly comes from changing the irrational appraisal/interpretation process of various events, this process is rather post-hoc and takes a slow process to train people to identify the many different kinds of cognitive distortions (some of which I think are inevitable as humans). However, if we can insert the right mental models of what to expect beforehand, we can have healthy reactions, even towards negative life events. For example, a patient who gets anxious all the time may have a lot of cognitive distortions (catastrophe, all-or-nothing, etc.). Even when he or she realize it and re-interpret it, anxiety may happen automatically without thinking. If we insert the right mental models for the patient, e.g. that everyone makes mistakes, and we learn through mistakes etc., we can help the patient to practice, and prepare for the upcoming challenging situations.

References

Algoe, S. B., Fredrickson, B. L., & Gable, S. L. (2013). The social functions of the emotion of gratitude via expression. *Emotion, 13*(4), 605.

Algoe, S. B., Gable, S. L., & Maisel, N. C. (2010). It's the little things: Everyday gratitude as a booster shot for romantic relationships. *Personal relationships, 17*(2), 217-233.

Algoe, S. B., & Haidt, J. (2009). Witnessing excellence in action: The 'other-praising'emotions of elevation, gratitude, and admiration. *The journal of positive psychology, 4*(2), 105-127.

Algoe, S. B., Haidt, J., & Gable, S. L. (2008). Beyond reciprocity: gratitude and relationships in everyday life. *Emotion, 8*(3), 425.

Alighieri, D. (2003). The divine comedy (J. Ciardi, Trans.). *New York: New American Library*.

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D., Glover, G., . . . Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature neuroscience, 6*(2), 196.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage, 38*(1), 95-113.

Baasten, M. (1986). Pride according to Pope Gregory the Great: A study of the Moralia: Lewinston, NY: Edwin Mellen Press.

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *Journal of Neuroscience, 27*(31), 8161-8165.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review, 84*(2), 191.

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Soc Cogn Affect Neurosci, 12*(1), 1-23.

Bartlett, M. Y., Condon, P., Cruz, J., Baumann, J., & Desteno, D. (2012). Gratitude: Prompting behaviours that build relationships. *Cognition & emotion, 26*(1), 2-13.

Bartlett, M. Y., & DeSteno, D. (2006). Gratitude and prosocial behavior: Helping when it costs you. *Psychological science, 17*(4), 319-325.

Belsky, J., Domitrovich, C., & Crnic, K. (1997). Temperament and Parenting Antecedents of Individual Differences in Three-Year-Old Boys' Pride and Shame Reactions. *Child development, 68*(3), 456-466.

Bliss, T. V., & Collingridge, G. L. (1993). A synaptic model of memory: long-term potentiation in the hippocampus. *Nature, 361*(6407), 31.

Bolton, G. E., & Zwick, R. (1995). Anonymity versus punishment in ultimatum bargaining. *Games and Economic behavior, 10*(1), 95-121.

Brown, J. D., & Marshall, M. A. (2001). Self-esteem and emotion: Some thoughts about feelings. *Personality and Social Psychology Bulletin, 27*(5), 575-584.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron, 35*(4), 625-641.

Carey, J. R., Clicque, S. H., Leighton, B. A., & Milton, F. (1976). A test of positive reinforcement of customers. *The Journal of Marketing*, 98-100.

Carlson, M., Charlin, V., & Miller, N. (1988). Positive mood and helping behavior: a test of six hypotheses: American Psychological Association.

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain, 129*(3), 564-583.

Chen, L. H., & Kee, Y. H. (2008). Gratitude and adolescent athletes' well-being. *Social Indicators Research, 89*(2), 361-373.

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience, 29*(39), 12315-12320.

Clark, H. B., Northrop, J. T., & Barkshire, C. T. (1988). The effects of contingent thank-you notes on case managers' visiting residential clients. *Education and Treatment of Children*, 45-51.

Corradi-Dell'Acqua, C., Civai, C., Rumiati, R. I., & Fink, G. R. (2012). Disentangling self-and fairness-related neural mechanisms involved in the ultimatum game: an fMRI study. *Soc Cogn Affect Neurosci, 8*(4), 424-431.

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature neuroscience, 3*(10), 1049.

De Quervain, D. J., Fischbacher, U., Treyer, V., & Schellhammer, M. (2004). The neural basis of altruistic punishment. *Science, 305*(5688), 1254.

Den Ouden, H. E., Frith, U., Frith, C., & Blakemore, S.-J. (2005). Thinking about intentions. *Neuroimage, 28*(4), 787-796.

Doyon, J., Penhune, V., & Ungerleider, L. G. (2003). Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia, 41*(3), 252-262.

Dunn, J. R., & Schweitzer, M. E. (2005). Feeling and believing: the influence of emotion on trust. *Journal of personality and social psychology, 88*(5), 736.

Emmons, R. A., & McCullough, M. E. (2003). Counting blessings versus burdens: an experimental investigation of gratitude and subjective well-being in daily life. *Journal of personality and social psychology, 84*(2), 377.

Emmons, R. A., & Shelton, C. M. (2002). Gratitude and the science of positive psychology. *Handbook of positive psychology, 18*, 459-471.

Ernst, M., Nelson, E. E., McClure, E. B., Monk, C. S., Munson, S., Eshel, N., . . . Towbin, K. (2004). Choice selection and reward anticipation: an fMRI study. *Neuropsychologia, 42*(12), 1585-1597.

Fellows, L. K., & Farah, M. J. (2003). Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain, 126*(8), 1830-1837.

Fox, G. R., Kaplan, J., Damasio, H., & Damasio, A. (2015). Neural correlates of gratitude. *Front Psychol, 6*, 1491. doi: 10.3389/fpsyg.2015.01491

Freitas, L. B. d. L., O'Brien, M., Nelson, J. A., & Marcovitch, S. (2012). The understanding of gratitude and Theory of Mind in 5-year-olds. *Psicologia: Reflexão e Crítica, 25*(2), 330-338.

Frith, C. D., & Frith, U. (1999). Interacting minds--a biological basis. *Science, 286*(5445), 1692-1695.

Froh, J. J., Kashdan, T. B., Ozimkowski, K. M., & Miller, N. (2009). Who benefits the most from a gratitude intervention in children and adolescents? Examining positive affect as a moderator. *The journal of positive psychology, 4*(5), 408-422.

Froh, J. J., Miller, D. N., & Snyder, S. F. (2007). *Gratitude in children and adolescents: Development, assessment, and school-based intervention.* Paper presented at the School Psychology Forum.

Froh, J. J., Yurkewicz, C., & Kashdan, T. B. (2009). Gratitude and subjective well-being in early adolescence: Examining gender differences. *Journal of adolescence, 32*(3), 633-650.

Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *Journal of Neuroscience, 26*(25), 6885-6892.

Geraghty, A. W., Wood, A. M., & Hyland, M. E. (2010a). Attrition from self-directed interventions: Investigating the relationship between psychological predictors, intervention content and dropout from a body dissatisfaction intervention. *Social science & medicine, 71*(1), 30-37.

Geraghty, A. W., Wood, A. M., & Hyland, M. E. (2010b). Dissociating the facets of hope: Agency and pathways predict dropout from unguided self-help therapy in opposite directions. *Journal of Research in Personality, 44*(1), 155-158.

Gilead, M., Katzir, M., Eyal, T., & Liberman, N. (2016). Neural correlates of processing "self-conscious" vs. "basic" emotions. *Neuropsychologia, 81*, 207-218. doi: 10.1016/j.neuropsychologia.2015.12.009

Glimcher, P. W., & Fehr, E. (2013). *Neuroeconomics: Decision making and the brain*: Academic Press.

Gordon, C. L., Arnette, R. A., & Smith, R. E. (2011). Have you thanked your spouse today?: Felt and expressed gratitude among married couples. *Personality and Individual Differences, 50*(3), 339-343.

Gottfried, J. A., O'doherty, J., & Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science, 301*(5636), 1104-1107.

Grant, A. M., & Gino, F. (2010). A little thanks goes a long way: Explaining why gratitude expressions motivate prosocial behavior. *Journal of personality and social psychology, 98*(6), 946.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review, 102*(1), 4.

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences, 98*(7), 4259-4264.

Happé, F. G., Winner, E., & Brownell, H. (1998). The getting of wisdom: Theory of Mind in old age. *Developmental psychology, 34*(2), 358-362.

Haruno, M., & Kawato, M. (2006). Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *Journal of neurophysiology, 95*(2), 948-959.

Haruno, M., Kuroda, T., Doya, K., Toyama, K., Kimura, M., Samejima, K., . . . Kawato, M. (2004). A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *Journal of Neuroscience, 24*(7), 1660-1665.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review, 91*(2), 73-78.

Isen, A. M., & Simmonds, S. F. (1978). The effect of feeling good on a helping task that is incompatible with good mood. *Social Psychology*, 346-349.

Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron, 58*(2), 284-294.

Joiner, J., Piva, M., Turrin, C., & Chang, S. W. (2017). Social learning through prediction error in the brain. *npj Science of Learning, 2*(1), 8.

Karns, C. M., Moore III, W. E., & Mayr, U. (2017). The Cultivation of Pure Altruism via Gratitude: A functional MRI Study of Change with Gratitude Practice. *Frontiers in human neuroscience, 11*, 599.

Kendler, K. S., Liu, X.-Q., Gardner, C. O., McCullough, M. E., Larson, D., & Prescott, C. A. (2003). Dimensions of religiosity and their relationship to lifetime psychiatric and substance use disorders. *American Journal of Psychiatry, 160*(3), 496-503.

Kini, P., Wong, J., McInnis, S., Gabana, N., & Brown, J. W. (2016). The effects of gratitude expression on neural activity. *Neuroimage, 128*, 1-10. doi: 10.1016/j.neuroimage.2015.12.040

Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Current opinion in neurology, 18*(4), 411-417.

Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport, 12*(17), 3683-3687.

Kong, F., He, Q., Liu, X., Chen, X., Wang, X., & Zhao, J. (2017). Amplitude of Low-Frequency Fluctuations During Resting State Differentially Predicts Authentic and Hubristic Pride. *Journal of Personality*.

Krause, N. (2006). Gratitude toward God, stress, and health in late life. *Research on Aging, 28*(2), 163-183.

Krause, N. (2009). Religious involvement, gratitude, and change in depressive symptoms over time. *The International journal for the psychology of religion, 19*(3), 155-172.

Lambert, N. M., Clark, M. S., Durtschi, J., Fincham, F. D., & Graham, S. M. (2010). Benefits of expressing gratitude: Expressing gratitude to a partner changes one's view of the relationship. *Psychological science, 21*(4), 574-580.

Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annu. Rev. Psychol., 58*, 317-344.

Lebreton, M., Jorge, S., Michel, V., Thirion, B., & Pessiglione, M. (2009). An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron, 64*(3), 431-439.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current opinion in neurobiology, 22*(6), 1027-1038.

Lewis, M., Alessandri, S. M., & Sullivan, M. W. (1992). Differences in shame and pride as a function of children's gender and task difficulty. *Child development, 63*(3), 630-638.

Lin, A., Adolphs, R., & Rangel, A. (2011). Social and monetary reward learning engage overlapping neural substrates. *Soc Cogn Affect Neurosci, 7*(3), 274-281.

Mahy, C. E., Moses, L. J., & Pfeifer, J. H. (2014). How and where: Theory-of-mind in the brain. *Developmental cognitive neuroscience, 9*, 68-81.

McAdams, D. P., & Bauer, J. J. (2004). Gratitude in modern life: Its manifestations and development. *The psychology of gratitude*, 81-99.

McCullough, M. E., Emmons, R. A., & Tsang, J.-A. (2002). The grateful disposition: a conceptual and empirical topography. *Journal of personality and social psychology, 82*(1), 112.

McCullough, M. E., Kilpatrick, S. D., Emmons, R. A., & Larson, D. B. (2001). Is gratitude a moral affect? *Psychological bulletin, 127*(2), 249.

McGovern, L., Ditzian, J. L., & Taylor, S. P. (1975). The effect of one positive reinforcement on helping with cost. *Bulletin of the Psychonomic Society, 5*(5), 421-423.

Menon, V., Anagnoson, R. T., Glover, G. H., & Pfefferbaum, A. (2000). Basal ganglia involvement in memory-guided movement sequencing. *Neuroreport, 11*(16), 3641-3645.

Mizuno, K., Kawatani, J., Tajima, K., Sasaki, A. T., Yoneda, T., Komi, M., . . . Watanabe, Y. (2016). Low putamen activity associated with poor reward sensitivity in childhood chronic fatigue syndrome. *NeuroImage: Clinical, 12*, 600-606.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience, 16*(5), 1936-1947.

Nakagawa, S., Takeuchi, H., Taki, Y., Nouchi, R., Kotozaki, Y., Shinada, T., . . . Yokoyama, R. (2017). Lenticular nucleus correlates of general self-efficacy in young adults. *Brain Structure and Function*, 1-10.

Nietzsche, F. W. (1968). *Basic writings of Nietzsche* (Vol. 1): Modern Library (Hardcover).

Northoff, G., Heinzel, A., De Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain—a meta-analysis of imaging studies on the self. *Neuroimage, 31*(1), 440-457.

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*(5669), 452-454.

O'doherty, J. P., Cockburn, J., & Pauli, W. M. (2017). Learning, reward, and decision making. *Annual review of psychology, 68*, 73-100.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron, 38*(2), 329-337.

O’Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current opinion in neurobiology, 14*(6), 769-776.

Payne, B. K. (2005). Conceptualizing control in social cognition: how executive functioning modulates the expression of automatic stereotyping. *Journal of personality and social psychology, 89*(4), 488.

Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of educational Psychology, 101*(1), 115.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature, 442*(7106), 1042.

Pizzagalli, D. A., Holmes, A. J., Dillon, D. G., Goetz, E. L., Birk, J. L., Bogdan, R., . . . Fava, M. (2009). Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder. *American Journal of Psychiatry, 166*(6), 702-710.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a Theory of Mind? *Behavioral and Brain Sciences, 1*(4), 515-526.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, 9*(7), 545-556.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory, 2*, 64-99.

Richards, J. M., Plate, R. C., & Ernst, M. (2013). A systematic review of fMRI reward paradigms used in studies of adolescents vs. adults: the impact of task design and implications for understanding neurodevelopment. *Neuroscience & Biobehavioral Reviews, 37*(5), 976-991.

Rilling, J. K., Goldsmith, D. R., Glenn, A. L., Jairam, M. R., Elfenbein, H. A., Dagenais, J. E., . . . Pagnoni, G. (2008). The neural correlates of the affective response to unreciprocated cooperation. *Neuropsychologia, 46*(5), 1256-1266.

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron, 35*(2), 395-405.

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). Opposing BOLD responses to reciprocated and unreciprocated altruism in putative reward pathways. *Neuroreport, 15*(16), 2539-2243.

Rind, B., & Bordia, P. (1995). Effect of server's "thank you" and personalization on restaurant tipping. *Journal of Applied Social Psychology, 25*(9), 745-751.

Risinger, R. C., Salmeron, B. J., Ross, T. J., Amen, S. L., Sanfilipo, M., Hoffmann, R. G., . . . Stein, E. A. (2005). Neural correlates of high and craving during cocaine self-administration using BOLD fMRI. *Neuroimage, 26*(4), 1097-1108.

Rosenberg, M. (1965). *Society and the adolescent self-image*: Princeton university press.

Ross, D. (1925). trans. Aristotle: The Nicomachean Ethics: New York: Oxford University Press. First published in.

Roth, L., Kaffenberger, T., Herwig, U., & Brühl, A. B. (2014). Brain activation associated with pride and shame. *Neuropsychobiology, 69*(2), 95-106.

Russo, S. J., & Nestler, E. J. (2013). The brain reward circuitry in mood disorders. *Nature Reviews Neuroscience, 14*(9), 609.

Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2014). A computational and neural model of momentary subjective well-being. *Proceedings of the National Academy of Sciences, 111*(33), 12252-12257.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300*(5626), 1755-1758.

Saxe, R. (2006). Uniquely human social cognition. *Current opinion in neurobiology, 16*(2), 235-239.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: the role of the temporo-parietal junction in “Theory of Mind”. *Neuroimage, 19*(4), 1835-1842.

Saxe, R., & Powell, L. J. (2006). It's the thought that counts: specific brain regions for one component of Theory of Mind. *Psychological science, 17*(8), 692-699.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht,, & T., V., K. . (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences, 36*, 393-414.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating Theory of Mind: a meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews, 42*, 9-34.

Scott-Van Zeeland, A. A., Dapretto, M., Ghahremani, D. G., Poldrack, R. A., & Bookheimer, S. Y. (2010). Reward Processing in Autism. *Autism Res, 3*(2), 53-67. doi: 10.1002/aur.122

Seligman, M. E., Steen, T. A., Park, N., & Peterson, C. (2005). Positive psychology progress: empirical validation of interventions. *American psychologist, 60*(5), 410.

Shariff, A. F., & Tracy, J. L. (2009). Knowing who’s boss: Implicit perceptions of status from the nonverbal expression of pride. *Emotion, 9*(5), 631.

Simon-Thomas, E. R., Godzik, J., Castle, E., Antonenko, O., Ponz, A., Kogan, A., & Keltner, D. J. (2012). An fMRI study of caring vs self-focus during induced compassion and pride. *Soc Cogn Affect Neurosci, 7*(6), 635-648. doi: 10.1093/scan/nsr045

Small, D. M., Gregory, M. D., Mak, Y. E., Gitelman, D., Mesulam, M. M., & Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron, 39*(4), 701-711.

Smith, C. A., Tong, E. M., & Ellsworth, P. C. (2014). The differentiation of positive emotional experience as viewed through the lens of appraisal theory. *Handbook of positive emotions*, 11-27.

Stanculescu, E. (2012). The self-conscious emotion of pride as mediator between selfesteem and positive affect. *Procedia-Social and Behavioral Sciences, 33*, 263-267.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine learning, 3*(1), 9-44.

Takahashi, H., Matsuura, M., Koeda, M., Yahata, N., Suhara, T., Kato, M., & Okubo, Y. (2007). Brain activations during judgments of positive self-conscious emotion and positive basic emotion: pride and joy. *Cerebral cortex, 18*(4), 898-903.

Tesser, A., Gatewood, R., & Driver, M. (1968). Some determinants of gratitude. *Journal of personality and social psychology, 9*(3), 233.

Tiedens, L. Z., Ellsworth, P. C., & Mesquita, B. (2000). Sentimental stereotypes: Emotional expectations for high-and low-status group members. *Personality and Social Psychology Bulletin, 26*(5), 560-575.

Toussaint, L., & Friedman, P. (2009). Forgiveness, gratitude, and well-being: The mediating role of affect and beliefs. *Journal of Happiness Studies, 10*(6), 635.

Tracy, J. L., Cheng, J. T., Robins, R. W., & Trzesniewski, K. H. (2009). Authentic and hubristic pride: The affective core of self-esteem and narcissism. *Self and identity, 8*(2-3), 196-213.

Tracy, J. L., & Robins, R. W. (2004a). " Putting the Self Into Self-Conscious Emotions: A Theoretical Model". *Psychological Inquiry, 15*(2), 103-125.

Tracy, J. L., & Robins, R. W. (2004b). Show your pride: Evidence for a discrete emotion expression. *Psychological science, 15*(3), 194-197.

Tracy, J. L., & Robins, R. W. (2007a). Emerging insights into the nature and function of pride. *Current directions in psychological science, 16*(3), 147-150.

Tracy, J. L., & Robins, R. W. (2007b). The psychological structure of pride: a tale of two facets. *Journal of personality and social psychology, 92*(3), 506.

Tracy, J. L., & Robins, R. W. (2008). The automaticity of emotion recognition. *Emotion, 8*(1), 81.

Tracy, J. L., Shariff, A. F., & Cheng, J. T. (2010). A naturalist's view of pride. *Emotion Review, 2*(2), 163-177.

Tricomi, E., Balleine, B. W., & O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *European Journal of Neuroscience, 29*(11), 2225-2232.

Tsang, J.-A. (2006). Gratitude and prosocial behaviour: An experimental test of gratitude. *Cognition & emotion, 20*(1), 138-148.

van Leijenhorst, L., Crone, E. A., & Bunge, S. A. (2006). Neural correlates of developmental differences in risk estimation and feedback processing. *Neuropsychologia, 44*(11), 2158-2170.

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage, 48*(3), 564-584.

Watkins, P. C. (2014). *Gratitude and the good life*: Springer.

Watkins, P. C., Uhder, J., & Pichinevskiy, S. (2015). Grateful recounting enhances subjective well-being: The importance of grateful processing. *The journal of positive psychology, 10*(2), 91-98.

Watkins, P. C., Woodward, K., Stone, T., & Kolts, R. L. (2003). Gratitude and happiness: Development of a measure of gratitude, and relationships with subjective well-being. *Social Behavior and Personality: an international journal, 31*(5), 431-451.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology, 54*(6), 1063.

Weidman, A. C., Tracy, J. L., & Elliot, A. J. (2016). The benefits of following your pride: Authentic pride promotes achievement. *Journal of Personality, 84*(5), 607-622.

Weiner, B., Russell, D., & Lerman, D. (1979). The cognition–emotion process in achievement-related contexts. *Journal of personality and social psychology, 37*(7), 1211.

Williams, L. A., & DeSteno, D. (2008). Pride and perseverance: the motivational role of pride. *Journal of personality and social psychology, 94*(6), 1007.

Williams, L. A., & DeSteno, D. (2009). Pride: Adaptive social emotion or seventh sin? *Psychological science, 20*(3), 284-288.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103-128.

Wood, A. M., Brown, G. D., & Maltby, J. (2011). Thanks, but I'm used to better: a relative rank model of gratitude. *Emotion, 11*(1), 175.

Wood, A. M., Froh, J. J., & Geraghty, A. W. (2010). Gratitude and well-being: A review and theoretical integration. *Clinical psychology review, 30*(7), 890-905.

Wood, A. M., Joseph, S., Lloyd, J., & Atkins, S. (2009). Gratitude influences sleep through the mechanism of pre-sleep cognitions. *Journal of psychosomatic research, 66*(1), 43-48.

Wood, A. M., Maltby, J., Gillett, R., Linley, P. A., & Joseph, S. (2008). The role of gratitude in the development of social support, stress, and depression: Two longitudinal studies. *Journal of Research in Personality, 42*(4), 854-871.

Wood, A. M., Maltby, J., Stewart, N., Linley, P. A., & Joseph, S. (2008). A social-cognitive model of trait and state levels of gratitude. *Emotion, 8*(2), 281.

Yu, H., Cai, Q., Shen, B., Gao, X., & Zhou, X. (2017). Neural substrates and social consequences of interpersonal gratitude: Intention matters. *Emotion, 17*(4), 589-601. doi: 10.1037/emo0000258

Zahn, R., Garrido, G., Moll, J., & Grafman, J. (2014). Individual differences in posterior cortical volume correlate with proneness to pride and gratitude. *Soc Cogn Affect Neurosci, 9*(11), 1676-1683. doi: 10.1093/scan/nst158

Zahn, R., Moll, J., Paiva, M., Garrido, G., Krueger, F., Huey, E. D., & Grafman, J. (2009). The neural basis of human social values: evidence from functional MRI. *Cerebral cortex, 19*(2), 276-283.

Zald, D. H., & Pardo, J. V. (1997). Emotion, olfaction, and the human amygdala: amygdala activation during aversive olfactory stimulation. *Proceedings of the National Academy of Sciences, 94*(8), 4119-4124.

List of figures and tables