The neural mechanisms of Self-deception

Self-deception

- Messick and Bazerman, 1996:
  - Self-deception is defined as being unaware of the processes that lead us to form our opinions and judgments.

- Tenbrunsel and Messick, 2004:
  - Self-deception allows one to behave self-interestedly while, at the same time, falsely believing that one's moral principles were upheld.
  - … the fact that it is unclear whether such deception is the result of a conscious act or an unconscious process.

- Peterson, et al., 2003:
  - These findings support a model of self-deception as ignoring evidence of error and reinforce the argument that self-deception may be maladaptive.

- Trivers et al., 2011
  - Self-deception means people decive themself, which makes it easier to deceive others.
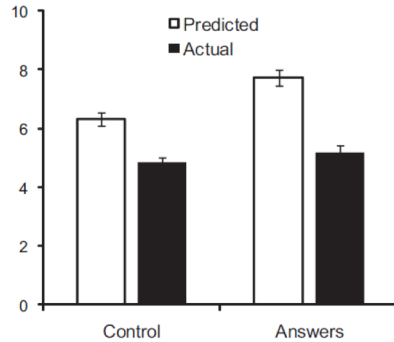
How to engage in self-deception?

Fig. 1. Predicted vs. actual performance on test 2 (experiment 2). The group with the previous advantage of the answer key again anticipated superior performance. However, we observed no difference in actual performance between the two groups.

Table 1. Self-deception proves costly: Experiment 3

	Control	Answers
Test 1 score	4.58	7.61***
Test 2 prediction	4.98	7.24***
Test 2 score	4.45	4.47
Earnings	$17.75	$14.47***

Participants given the answer key for the first test performed better on that test than those in the control condition, and erroneously predicted higher scores on the second test, resulting in lower earnings due to their similar performance. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.
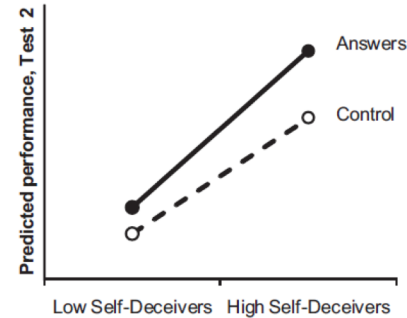

Fig. 2. Regression model interaction with dispositional self-deception (experiment 2). Predictions of performance on the second test are moderated by dispositional self-deception, with high self-deceivers (+1 SD above the mean) showing greater inflation in the answers condition than low self-deceivers (−1 SD below the mean).

People who exploited opportunities to cheat on tests were likely to engage in self-deception .

– Chance, Norton, Gino, & Ariely, 2011

Neural correlates of self-deception

➤ Design:
- response (fake bad, fake good, count)
- condition: (impression management or self-deception)

➤Conclusions:

✓Self-deception, impression-management and faking bad: the mPFC and left vlPFC

✓Impression- management: the left dlPFC

✓Faking bad: the right vlPFC, left TPJ and right cerebellum

Self-deception vs. counting baseline. p<0.05 FWE. Extent threshold = 10 voxels

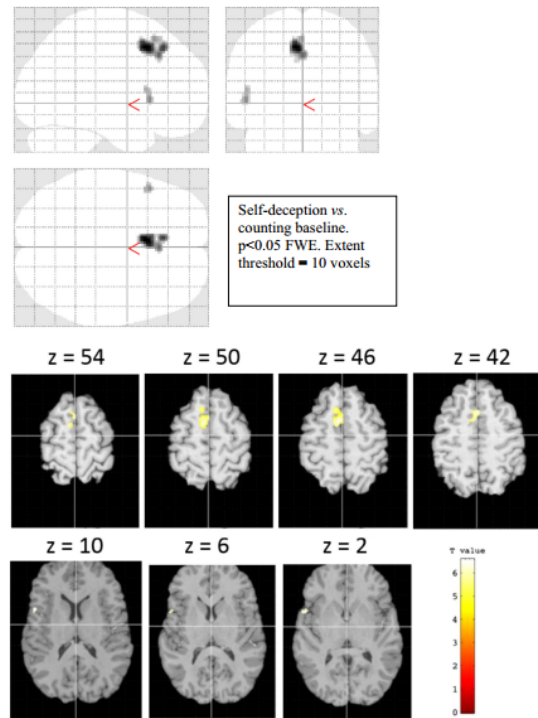z = 54   z = 50   z = 46   z = 42

z = 10   z = 6   z = 2

T value

Fig. 2. Self-deception vs. counting baseline. Flexible factorial design. p < 0.05 family-wise error (FWE). Extent threshold = 10 voxels. See Table 2 for anatomical descriptions and co-ordinates.

Neuropsychologia 67 (2015) 159–174

−Farrow, Burgess, Wilkinson, & Hunter, 2015

4

Questioans/Hypotheses

Q1: Will monetary reward enhance self-deception?

Q2: What's the neural mechanisms of deceiving oneself and others?

—H1: The degree of self-deception is greater in the money condition (pre-behavior study)
- More motivated
    - Psychological benefit/positive illusions
    - Monetary rewards

—H2: The medial prefrontal cortex (mPFC) will associate with self-deception (fMRI study)

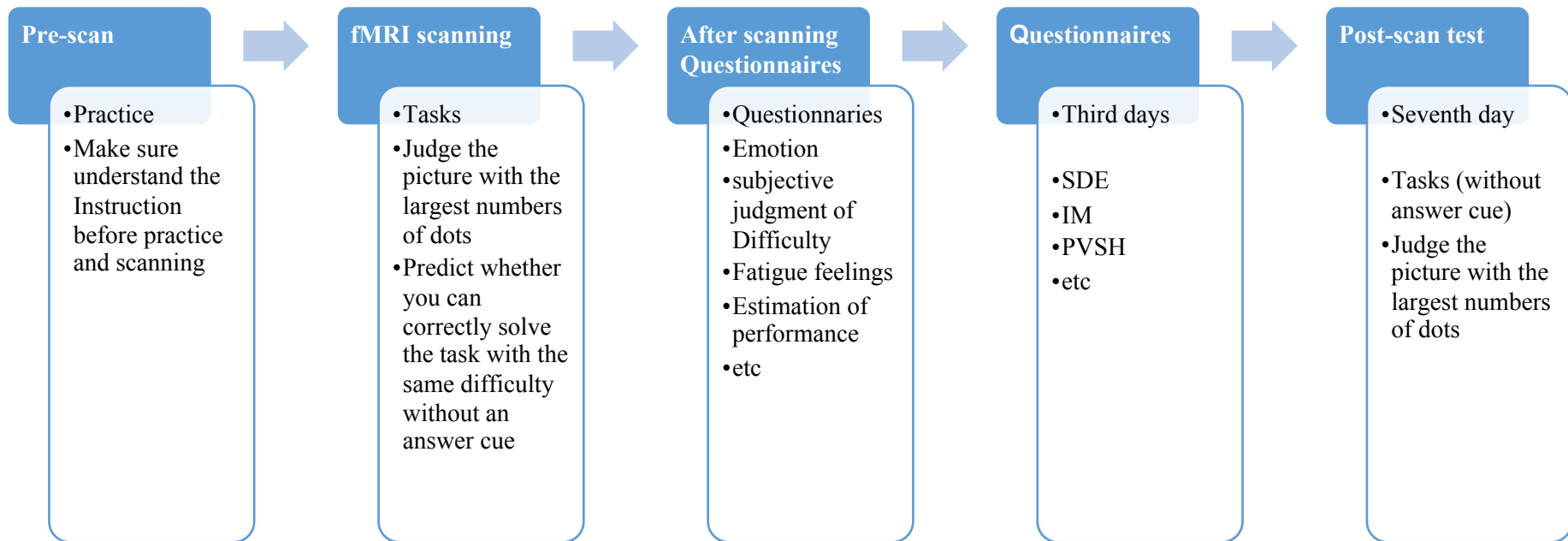—H3: The N2 and P3 will associate with self-decption (ERP study)

fMRI Study

Procedure

Pre-scan

- Practice
- Make sure understand the Instruction before practice and scanning

fMRI scanning

- Tasks
- Judge the picture with the largest numbers of dots
- Predict whether you can correctly solve the task with the same difficulty without an answer cue

After scanning Questionnaires

- Questionnaries
- Emotion
- subjective judgment of Difficulty
- Fatigue feelings
- Estimation of performance
- etc

Questionnaires

- Third days

- SDE
- IM
- PVSH
- etc

Post-scan test

- Seventh day

- Tasks (without answer cue)
- Judge the picture with the largest numbers of dots

Design

- Conditions:
  - Answer condition (115 trails):
    - ➢ Control (15 trails) / Middle (50 trails) / Hard difficuty (50 trails)
    - Choose the picture with the largest numbers of dots
      - Right: +1
      - Answer presents with pictures
    - "Can you correctly solve the task with the same difficulty without an answer cue?"
      - Can/ Can not
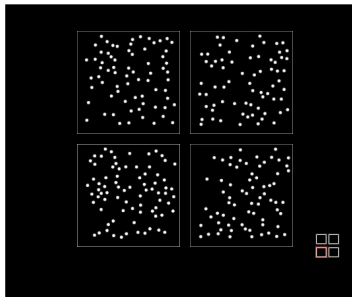   - *If the difference between your prediction and your actual score in part two (NS) is smaller than 10%, then you can get extra money ￥30 .*
  - No Answer condition (115 trails):
    - Choose the picture with the largest numbers of dots
      - Right: +1
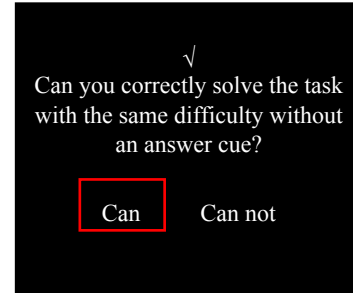
Experimental program

Answer Session (AS)

Prediction Phase (PP)
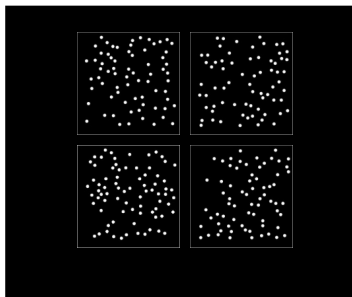
√
Can you correctly solve the task
with the same difficulty without
an answer cue?

Can   Can not
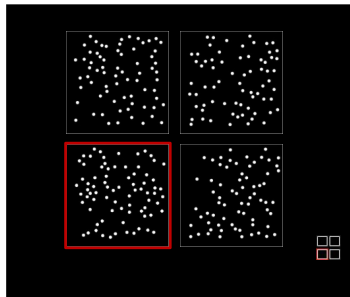
115 trails

4 s   2-6 s   3 s   2-6 s   15s
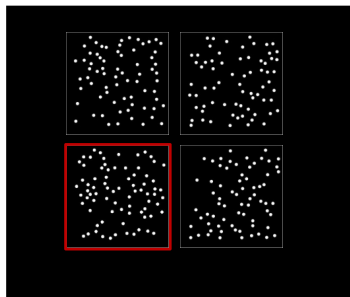
No-answer Session(NAS)

115 trails

(15control, 50middle,50hard)

4 s   2-6 s

8s

(15+8) x 115 =
44.09')

RAW & RAR

After scanning Questionnaires:

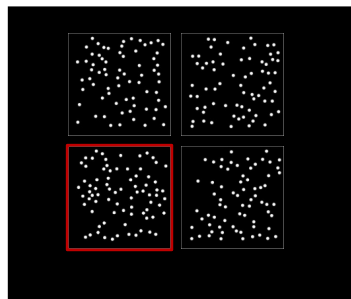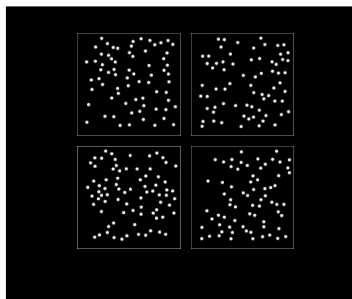(Please answer the following questions according to your feelings in the experiemnt.)

✓ (Subjective Correctness) Please estimate your performance in each session with percentage.

✓ (Subjective Emotion feelings) Rate your emotion feelings in each session: 1(very negative)-7 (very positive)

✓ (Subjective Difficuty Level) Rate the overall difficuty of the task you feel in each session: 1(very easy)-7(very difficuty)

✓ (Subjective Fatigue feelings) Rate the degree of fatigue you feel in each session: 1(not at all)-7(very much)

✓ (Subjective proficiency) Do you feel you are good at identifying the picture with more dots? 1 not at all - 7 vey much

• Did you learned the patterns in the first session and used the patterns in the second session? 1 not at all - 7 vey much

• Did Answers provided in the first session influenced you perform?

  • 1 significantly decrease - 7 significantly increase

Questionnaires (Scales)

- Balanced inventory of desirable responding (MIDR) (40 items: 1,not ture--7,very ture;)

    - *Self-deception enhancement (SDE ↑)*

    - *Impression management (IM ↑)*

- *Narcissistic personality inventory (*NPI) (40 items: 0-20, ↑),

- Generalized Self-efficacy scale (GSE) (Schwarzer & Jerusalem, 1995; 10 items: 1, Not at all true-4,Exactly true↑)

- Personal values scale-honesty (PVSH) (Scott, 1965; 20 items: 0-20, Honesty ↑)

- Defensive pessimisim questionnaire (Norem, 2002; 12 items: 1,Not at all true of me---7Very true of me ↑)

- Self-handicapping scale (Martin, 1998; 1,Not at all true of me---5Very true of me ↑)

    - *Behavioral self-handicapping (9 items)*

    - *Claimed self-handicapping (16 items)*

- Well Being （ 1,Not at all ---5Very happiness↑)

- PHQ4 （ 0, not at all, 1, several days, 2, more than half the days, 3, nearly every day ↑)
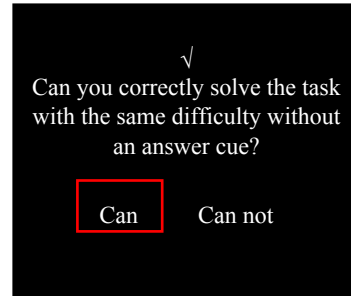
Post-test program

Answer Session (AS)

Prediction Phase (PP)

√
Can you correctly solve the task
with the same difficulty without
an answer cue?

Can   Can not

15s

30 trails
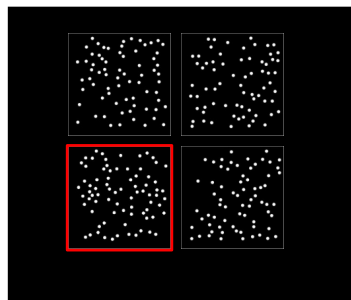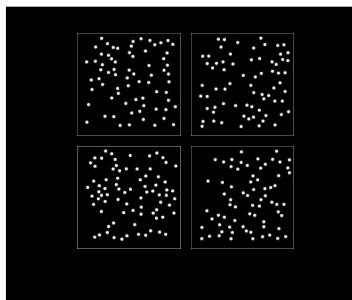
4 s   2-6 s   3 s   2-6 s

NoAnswer Session(NAS)

8s

30 trails

(10control, 10middle,10hard)

4 s   2-6 s

(15+8) x 30 =
11.5')

14

Behavioral data analyses

➢ **Participants.**

• Thity-six undergraduate students (female = 15; Mean age = 20.19, *SD* = 1.43) participated in this study for monetary rewards. One participants deleted because of system settings error and one participants deleted becasuse of head move.

➢ **Scales**.

– *Impression management (IM),* Cronbach's alpha = .572.

– *Self-deception enhancement (SDE)*, Cronbach's alpha = .625.

– *Narcissistic personality inventory (NPI) ,* Cronbach's alpha = .858.

– *Generalized self-efficacy scale (GSE)*, Cronbach's alpha = .816.

– *Questionnaire of Personal Values Scale-Honesty (PVSH),* Cronbach's alpha = .842.

– *Defensive Pessimism Questionnaire (DPQ)*, Cronbach's alpha = .868.

– *Claimed subscale of the SHS (BSH),* Cronbach's alpha = .660.

– *Behavioral subscale of the SHS (CSH)*, Cronbach's alpha = .541.

– *Well Being,* Cronbach's alpha = .652.

– *PHQ4*, Cronbach's alpha = .759.

Self-deception Parameters Estimation

(1) AS: $logit(\pi_i) = -Difficulty_j + Deception_{ij} + Ability_i$

(2) PP: $logit(\pi_i) = -Difficulty_j + Self\text{-}deception_{ij} + Ability_i + error_{ij}$

(3) NAS: $logit(\pi_i) = -Difficulty_j + Ability_i$

} fMRI

(4) NAS: $logit(\pi_i) = -Difficulty_j + Ability_i$

(5) PP: $logit(\pi_i) = -Difficulty_j + Ability_i + error_{ij}$
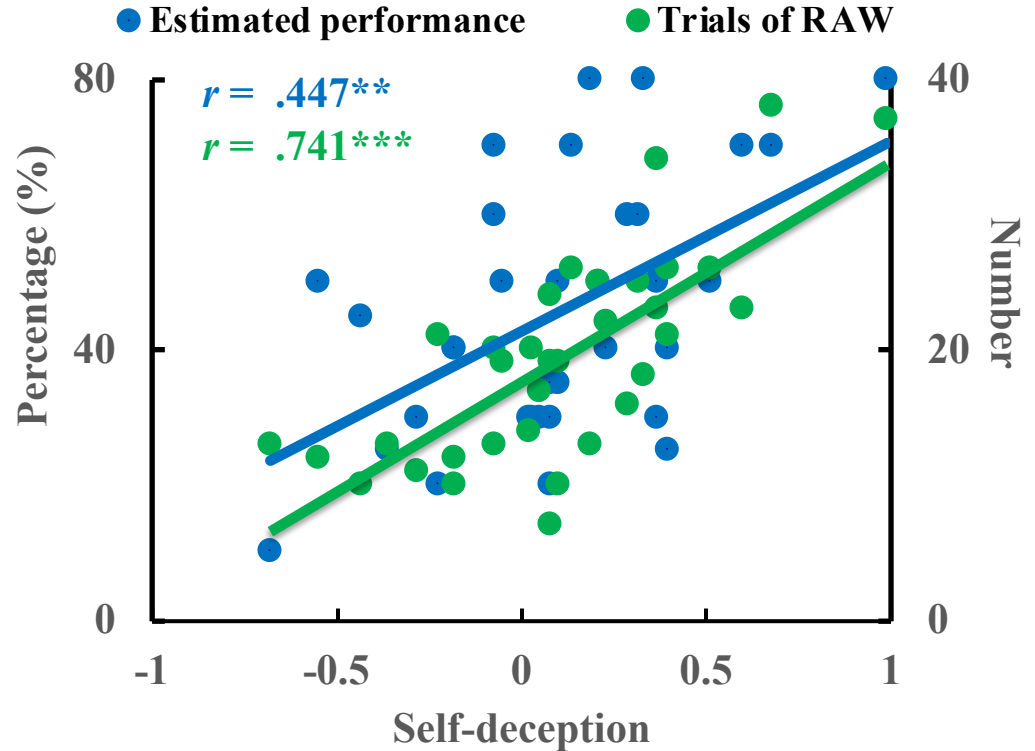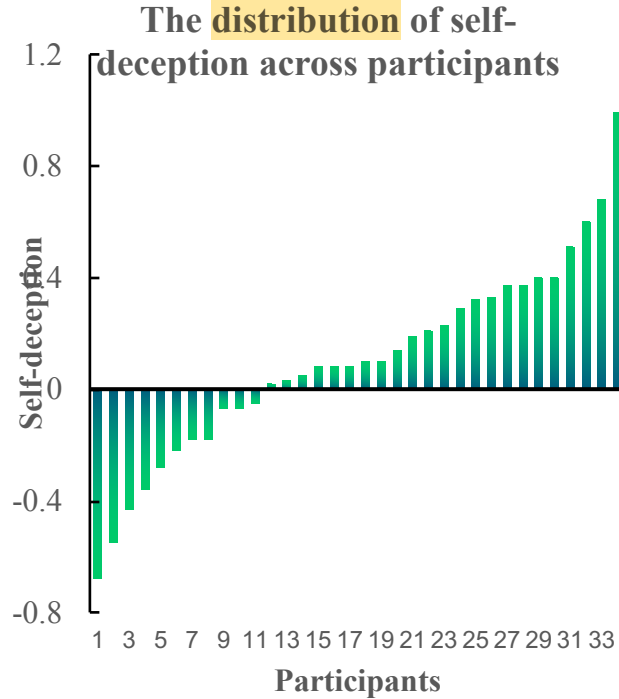
(6) NAS: $logit(\pi_i) = -Difficulty_j + Ability_i$

} Post-test

Note: i, subject; j, difficulty level. Eq. 1-3 estimated by using fMRI data, and Eq. 4-6 estimated by using post-scan behavioral test (measured on the seventh day after fMRI scanning).

Behavioral results

(N=34)

Behavioral results

The **distribution** of self-deception across participants

Self-deception / Participants

Estimated performance · Trials of RAW

$r = .447^{**}$
$r = .741^{***}$

Percentage (%) / Number / Self-deception
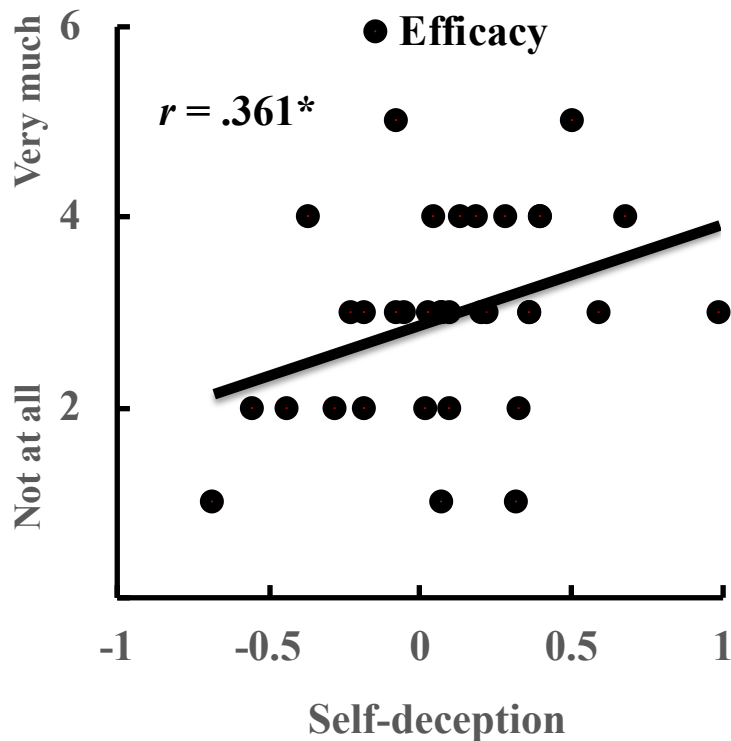
SD & Subjective proficiency (efficacy)

➢ **The feelings of efficacy**

— Do you feel you are good at identifying the picture with more dots? 1 not at all - 7 vey much

—你觉得自己擅长识别点数最多的图片吗？（1,完全不擅长；7,非常擅长）

Table 1
Predictors of subjective proficiency

Variables	Beta	t	Sig
Constant		.471	.641
Ability	-.053	-.298	.768
Deception	-.034	-.201	.842
Self-deception	.374*	2.106	.044

*$p < .05$ **$p < .01$

$r = .361*$

Psychological benefits of self-deception

Emotion

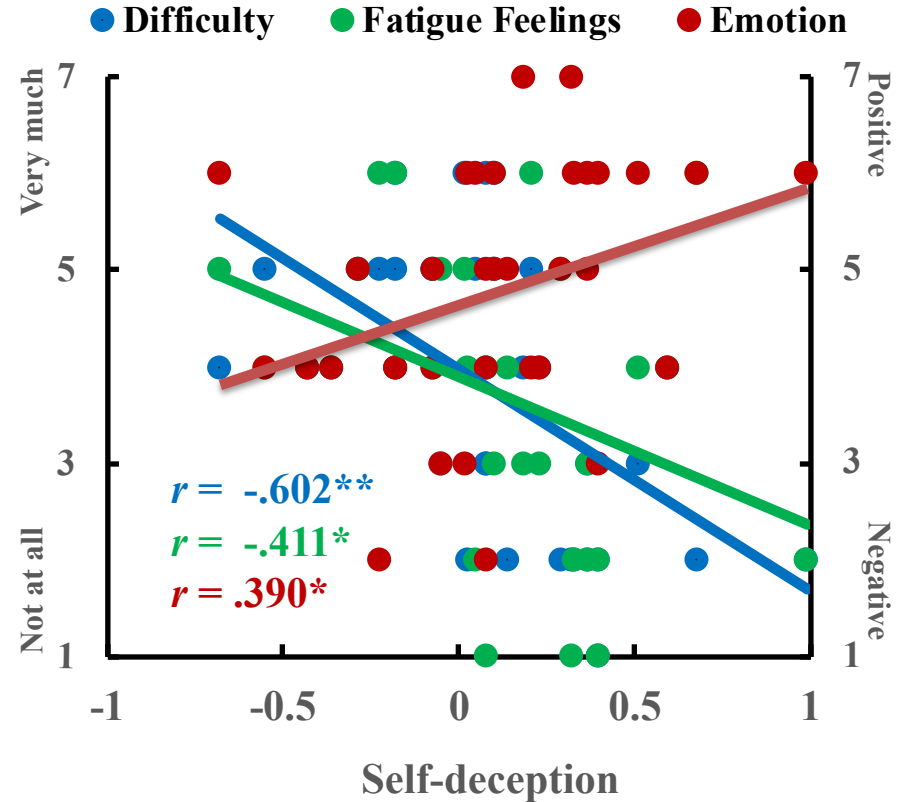─Rate your emotion feelings in each session: 1(very negative)-7 (very positive)

Difficuty Feelings

─Rate the overall difficuty of the task you feel in each session: 1(very easy)-7(very difficuty)

Fatigue feelings

─Rate the degree of fatigue you feel in each session: 1(not at all)-7(very much)

fMRI data analyses

Method

- SPM12 was adopted for fMRI data analysis. For each participants, EPI images were first realigned and corrected for slice timing. Data sets that exhibited movement of >3 mm or 3° of rotation were not included. The anatomical image was coregistered to the mean EPI image, and segmented, generating parameters for normalization to MNI space. EPI data were then projected onto MNI space with a $2 \times 2 \times 2$ mm^3 resolution and smoothed using an 8-mm FWHM (full width at half maximum) isotropic Gaussian kernel. High-pass temporal filtering with a cut-off of 128 s was performed to remove low-frequency drifts.

- **Participants**

- Thity-six undergraduate students (female = 15; Mean age = 20.19, SD = 1.43) participated in this study for monetary rewards. One participants deleted because of system settings error, two participants deleted becasuse of head move, and five participants were removed from the analysis due to insufficient trials (<5) for at least one of regressors of interest. Therefore, data from 28 participants (11 female) were included.

Neuroimaging results

Results of regression analysis in the contrast of RAW versus RAR in the AS with estimated self-deception parameter as a covariate. (N = 32)
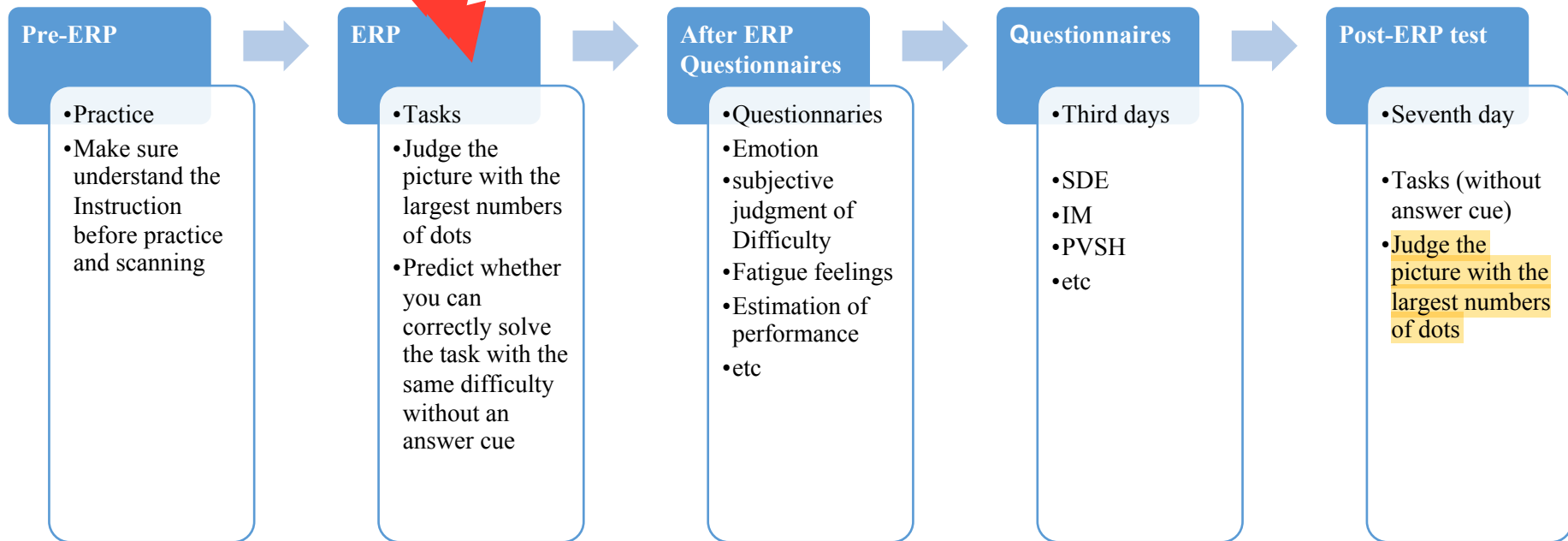

X = -5
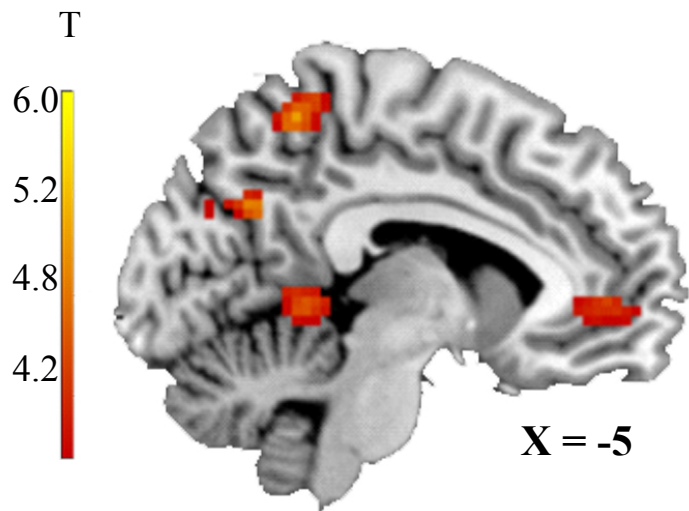
Significant activation was found in the ==ventromedial prefrontal cortex== (vmPFC; voxel-level threshold $p < 0.001$ uncorrected, cluster-level $p < 0.05$, FWE corrected; $N = 32$).
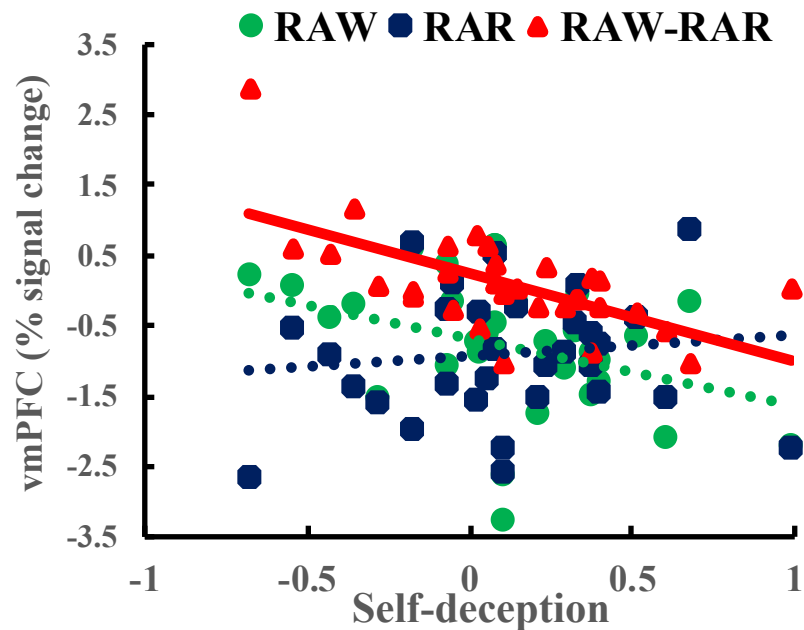
Parameter estimates were extracted from the ==whole activated cluster== in the vmPFC in conditions of RAW vs. RAR, RAW and RAR.

ERP Study

Procedure

Pre-ERP

- Practice
- Make sure understand the Instruction before practice and scanning

ERP

- Tasks
- Judge the picture with the largest numbers of dots
- Predict whether you can correctly solve the task with the same difficulty without an answer cue

After ERP Questionnaires

- Questionnaries
- Emotion
- subjective judgment of Difficulty
- Fatigue feelings
- Estimation of performance
- etc

Questionnaires

- Third days

- SDE
- IM
- PVSH
- etc

Post-ERP test

- Seventh day

- Tasks (without answer cue)
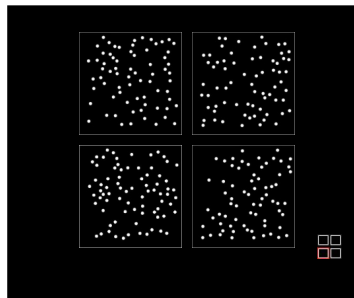- Judge the picture with the largest numbers of dots

Design

- Conditions:
  - Answer condition (230 trails):
    - ➤ Control (30 trails) / Middle (100 trails) / Hard difficuty (100 trails)
    - Choose the picture with the largest numbers of dots
      - Right: +1
      - Answer presents with pictures
    - "Can you correctly solve the task with the same difficulty without an answer cue?"
      - Can/ Can not
      - *If the difference between your prediction and your actual score in part two (NAS) is smaller than 10%, then you can get extra money ￥20 .*
  - No Answer condition (230 trails):
    - Choose the picture with the largest numbers of dots
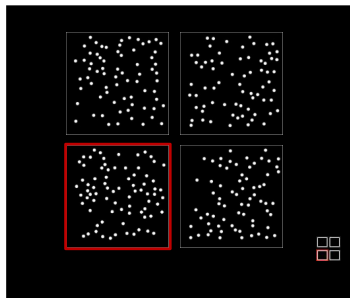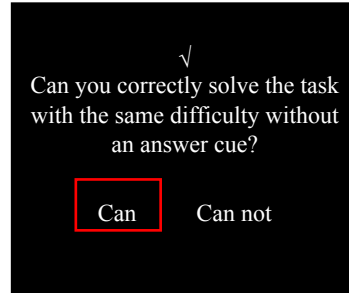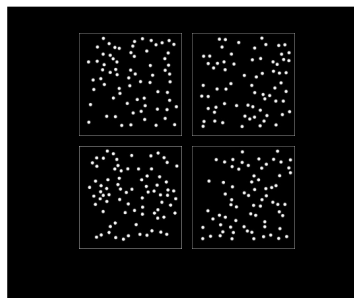   - Right: +1

Experimental program

After ERP Questionnaires:

(Please answer the following questions according to your feelings in the experiemnt.)

✓ (Subjective Correctness) Please estimate your performance in each session with percentage.

✓ (Subjective Emotion feelings) Rate your emotion feelings in each session: 1(very negative)-7 (very positive)

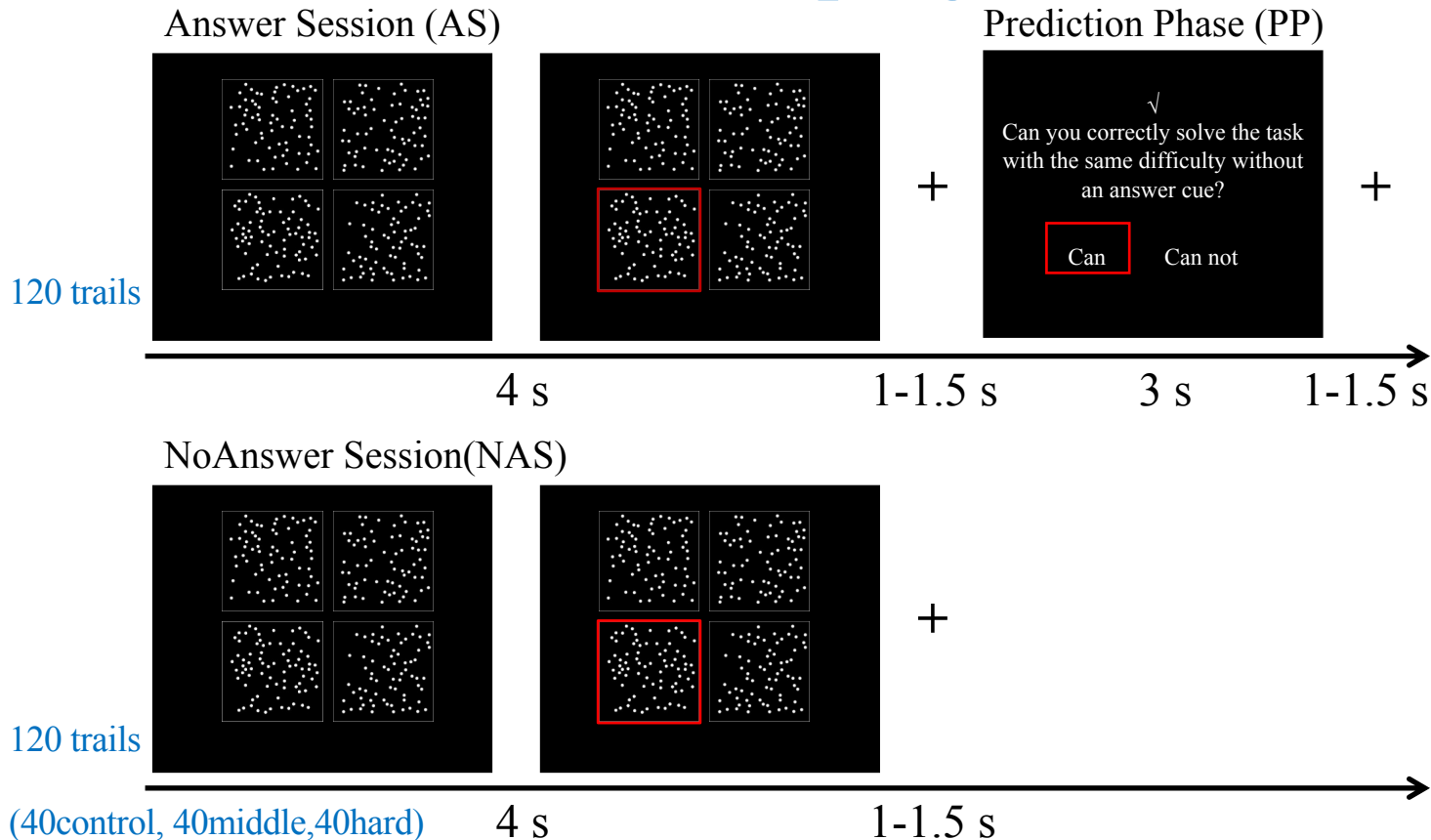✓ (Subjective Difficuty Level) Rate the overall difficuty of the task you feel in each session: 1(very easy)-7(very difficuty)

✓ (Subjective Fatigue feelings) Rate the degree of fatigue you feel in each session: 1(not at all)-7(very much)

✓ (Subjective proficiency) Do you feel you are good at identifying the picture with more dots? 1 not at all - 7 vey much

• Did you learned the patterns in the first session and used the patterns in the second session? 1 not at all - 7 vey much

• Did Answers provided in the first session influenced you perform?

      • 1 significantly decrease - 7 significantly increase

Questionnaires (Scales)

- Balanced inventory of desirable responding (MIDR) (40 items: 1,not ture--7,very ture;)

    - *Self-deception enhancement (SDE ↑*)

    - *Impression management (IM ↑)*

- *Narcissistic personality inventory (*NPI) (40 items: 0-20, ↑)*,*

- Generalized Self-efficacy scale (GSE) (Schwarzer & Jerusalem, 1995; 10 items: 1, Not at all true-4,Exactly true↑)

- Personal values scale-honesty (PVSH) (Scott, 1965; 20 items: 0-20, Honesty ↑)

- Defensive pessimisim questionnaire (Norem, 2002; 12 items: 1,Not at all true of me---7Very true of me ↑)

- Self-handicapping scale (Martin, 1998; 1,Not at all true of me---5Very true of me ↑)

    - *Behavioral self-handicapping (9 items)*

    - *Claimed self-handicapping (16 items)*

- Well Being （ 1,Not at all ---5Very happiness↑ ）

- PHQ4 （ 0, not at all, 1, several days, 2, more than half the days, 3, nearly every day ↑ ）

Post-ERP program

Answer Session (AS)

Prediction Phase (PP)

120 trails

√
Can you correctly solve the task
with the same difficulty without
an answer cue?

Can   Can not

+   +

4 s   1-1.5 s   3 s   1-1.5 s

NoAnswer Session(NAS)

120 trails

+

(40control, 40middle,40hard)   4 s   1-1.5 s

Method

- **Participants**

- Thity undergraduate students (female = XXX; Mean age = XXX, SD = XXX) participated in this study for monetary rewards.