SUPPLEMENTARY MATERIALS for article entitled "Comparison of social cognition using an adapted Chinese version of the Reading the Mind in the Eyes Test in drug-naive and regularly medicated individuals with chronic schizophrenia and healthy controls in rural China" by Deng et al.

Changes in the administration of the RMET-CV-R

The original 36-item RMET and the 70-item Chinese version (RMET-CV) are computerized tests which are self-completed by respondents. However, in this study 59.4% of the sample (n=384) were over 60 years of age or had less than 5 years of schooling, and 88% had never used a computer. To conduct the test with these respondents it was necessary to make the following changes in the administration of the test (without changing the presented pictures or optional descriptors provided in the original versions of the test).

1) All RMET-CV-R tests are conducted in the presence of an interviewer. Unlike the administration of the RMET, for the RMET-CV-R the interviewer reads aloud the requirements of the test and shows the respondent how to make responses during the training session prior to starting the main test.

2) For all participants with less than five years of schooling, the two training items and the 70 items in the RMET-CV-R main test were 'interviewer-completed': the interviewer reads aloud the four options for each picture at a standard speed (1 second for two-Chinese-character-words, 2 seconds for four-Chinese-character-words) and clicked the mouse to select the descriptor chosen by the respondent. The participant was then asked if the picture was of a male or a female and the interviewer clicked the corresponding response on the screen.

3) For participants with five or more years of schooling, they are first trained to use the computer mouse by having them independently manipulate the mouse to move the position of the cursor over 5 numbered buttons on the computer screen and click each button in sequence. If the respondent was unable to complete this task after several trials, the test was interviewer-completed.

4) Presentation of the two pictures used for training in the RMET-CV-R is managed by the interviewer who explains the expectations at each step (i.e., to select an appropriate emotional descriptor and then indicate the gender of the presented eyes) and asks the respondent to make a selection when presented with the pictures. If the respondent has difficulty in completing this task, the interviewer makes up to 6 attempts to help the respondent understand the requirements of the task prior to starting the main trial.

5) Respondents with more than five years of education who are able to learn how to use a mouse are expected to complete the computerized main test on their own (i.e., a 'self-completed' test). However, if a respondent clearly has difficulty reading the words on the slide or manipulating the mouse either during the training slides or during the main test, the interviewer reverts to the 'interviewer-completed' approach.

6) One of the main difficulties these respondents had completing the test was that they often did not understand some of the emotional descriptors used. In the RMET and some of the other translated versions of the instrument (such as the Spanish and Italian version provided at https://www.autismresearchcentre.com/tests/eyes-test-adult/), respondents are provided a definition handout for terms used in the slides that they can refer to while doing the test; but this was not done in the RMET-CV or other translated versions (such as the German and French

versions, provided at https://www.autismresearchcentre.com/tests/eyes-test-adult/). Similarly, in our revised Chinese version (RMET-CV-R) we have not provided a definition handbook because many of the respondents were illiterate or semi-literate, so having the interviewer read the definitions would have inordinately extended the duration of the test. However, to reduce the number of respondents who were excluded because of limited vocabulary, respondents are allowed to skip up to 4 of the 70 pictures (skipped items are considered 'incorrect' when computing total number of correct responses). When the respondent fails to complete more than 4 items, the test is terminated and considered incomplete.

7) To balance any potential order effects, 327 participants enrolled before September 2017 were first presented Asian then Caucasian pictures, while 320 participants enrolled after September 2017 were first presented Caucasian then Asian pictures.

8) Participants who have started the main test but refuse to continue before finishing are encouraged a maximum to two times with the phrase 'please try to complete the test'; if they persist in their request to stop the test, the test is terminated and the results considered incomplete.

The computerized RMET-CV-R was generated using Java. After completion or termination of each test a separate text file recording the selected emotional descriptor, the selected gender, and the time to make both of these decisions for each of the 70 pictures and a summary table separating results for the 36 Caucasian and 34 Asian pictures is downloaded. These data were then automatically extracted into a csv database using Python.

Assessment of RMET-CV-R completion status

Given the possibility that some older or uneducated respondents could not successfully complete RMET-CV-R, we classified the completion status of each administration of the test based on the respondent's willingness to complete the test, ability to complete the test, and understanding of the required task, as follows:

a) Respondent explicitly refuses to complete the practice examples, to start the main test *or* to respond to 1 or more pictures during a main test in which a total of 5 pictures have been skipped (either due to refusal or failure to select a response).

b) Respondent skips a total of 5 pictures (occurring at any point in the series of 70 pictures) none of which are skipped due to active refusal *and* has indicated in some way that he/she understands what is expected (either by choosing an emotional descriptor or by stating that they are unable to choose a descriptor for any slide). In almost all cases this occurs because the respondent has limited vocabulary and, thus, does not understand the provided emotional descriptors.

c) Respondent provides no active response during the practice examples or during the first five pictures in the main test *and* has not indicated in any way that he/she at least partially understands what is expected.

d) Respondent completes 66-70 pictures.

Tests classified as 'a', 'b' or 'c' are considered 'incomplete' and those classified as 'd' are considered 'successfully completed'.

Supplementary Table 1A. Multivariate logistic regression results comparing rates of successful completion of RMET-CV-R between untreated cases (UT) and treated controls (TC) in the full sample and in different subsamples, including duration of illness as a covariate

		Total sample[a] n=387		Respondents assigned to self-completed RMET-CV-R (self-completion subsample) n=136		Respondents assigned to have RMET-CV-R administered by interviewer (interviewer-completed subsample) n=249[a]		Respondents with ≥ 5 years of schooling and <60 (educated-younger subsample) n=163		Respondents with <5 years of schooling or ≥ 60 (undereducated-older subsample) n=224	
		OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Group	TC v. UT*	**2.61**	**1.58-4.29**	1.55	0.66-3.67	**3.19**	**1.71-5.92**	1.72	0.80-3.72	**3.25**	**1.65-6.41**
Age in years		0.99	0.96-1.02	0.99	0.92-1.06	0.99	0.96-1.02	0.98	0.92-1.04	0.99	0.96-1.02
Gender (1=female*, 2=male)		**1.76**	**1.06-2.94**	**2.34**	**1.01-5.44**	1.65	0.85-3.23	1.60	0.75-3.39	2.03	0.99-4.15
Years of schooling	0 years	Reference group		----	-----	Reference group		----	----	Reference group	
	1-4 years	**2.79**	**1.51-5.16**	-----	-----	**2.93**	**1.53-5.59**	----	----	**3.01**	**1.55-5.85**
	5-7 years	**3.62**	**1.79-7.33**	Reference group		**3.58**	**1.16-11.04**	Reference group		**2.90**	**0.69-12.21**
	8-10 years	**5.77**	**2.58-12.92**	1.71	0.63-4.59	**6.56**	**1.83-23.45**	1.63	0.68-3.90	**10.57**	**1.35-82.87**
	11+ years	**6.91**	**1.66-28.82**	3.19	0.57-17.77	----	----	2.03	0.46-9.02	----	----
Residence (1=urban*, 2=rural)		0.38	0.14-1.01	0.72	0.16-3.16	**0.26**	**0.08-0.89**	0.50	0.10-2.60	0.30	0.09-1.06
Ethnicity (1=Han*, 2=minority group)		1.12	0.68-1.83	1.09	0.41-2.92	1.13	0.63-2.05	1.30	0.53-3.19	1.05	0.57-1.95
Duration of illness	<10 years v. 30+years*	**3.43**	**1.28-9.22**	2.61	0.38-18.01	**4.03**	**1.25-13.00**	1.51	0.30-7.57	**6.19**	**1.86-20.62**
	10-19 years v. 30+ years*	2.20	0.99-4.87	2.10	0.40-11.09	2.17	0.84-5.62	1.57	0.39-6.28	**2.68**	**1.01-7.12**
	20-29 years v. 30+ years*	1.69	0.85-3.37	2.00	0.50-8.05	1.65	0.73-3.77	1.88	0.55-6.45	1.54	0.63-3.74
Adjusted percent (95% CI) of successful completion of RMET-CV-R computed from regression											
UT		50.8% (45.1%-56.5%)		74.1% (64.5%-83.6%)		38.7% (31.8%-45.6%)		72.6% (63.5%-81.8%)		35.8% (28.7%-42.9%)	
TC		69.0% (62.0%-76.1%)		81.1% (71.4%-90.7%)		62.5% (52.7%-72.2%)		81.6% (73.0%-90.2%)		59.9% (49.1%-70.7%)	

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes

statistically significant results are **bolded**

* denotes the reference group

[a] Two cases (one in untreated group and one in treated control group) were excluded from the analysis as outliers because there were only two cases in the education group '11 years or above'

Supplementary Table 1B. Multivariate logistic regression results comparing rates of successful completion of RMET-CV-R between untreated cases (UT) and treated controls (TC) in the full sample and in different subsamples, including duration of illness and PANSS total score as covariates

		Full sample n=387		Respondents assigned to self-completed RMET-CV-R (self-completion subsample) n=136		Respondents assigned to have RMET-CV-R administered by interviewer (interviewer-completed subsample) n=249[a]		Respondents with ≥ 5 years of schooling and <60 (educated-younger subsample) n=163		Respondents with <5 years of schooling or ≥ 60 (undereducated-older subsample) n=224	
		OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Group	TC v. UT*	1.47	0.85-2.56	0.52	0.17-1.58	**2.03**	**1.06-3.90**	0.50	0.18-1.41	**2.24**	**1.11-4.51**
Age in years		0.99	0.96-1.02	0.99	0.91-1.07	0.99	0.96-1.02	1.01	0.94-1.08	0.99	0.96-1.03
Gender (1=female*, 2=male)		1.57	0.93-2.65	**2.57**	**1.00-6.59**	1.42	0.72-2.78	1.81	0.77-4.25	1.73	0.84-3.53
Years of schooling	0 years	Reference group		----	-----	Reference group		----	-----	Reference group	
	1-4 years	**3.13**	**1.62-6.07**	----	-----	**3.21**	**1.64-6.27**	----	-----	**3.23**	**1.64-6.34**
	5-7 years	**3.27**	**1.53-6.98**	Reference group		3.03	0.95-9.67	Reference group		2.69	0.57-12.62
	8-10 years	**5.74**	**2.45-13.44**	1.76	0.57-5.50	**7.90**	**2.31-27.01**	2.21	0.79-6.25	**9.43**	**1.35-65.81**
	11+ years	**7.72**	**1.88-31.72**	3.91	0.60-25.47	No observation		2.57	0.56-11.74	No observation	
Residence (1=urban*, 2=rural)		0.37	0.13-1.05	0.53	0.13-2.15	**0.26**	**0.07-0.95**	0.28	0.07-1.19	0.32	0.09-1.17
Ethnicity (1=Han*, 2=minority group)		1.18	0.69-2.02	0.66	0.21-2.09	1.31	0.70-2.44	0.70	0.24-2.01	1.23	0.65-2.33
Duration of illness	<10 years v. 30+years*	**3.76**	**1.38-10.22**	3.15	0.35-28.34	**4.19**	**1.28-13.69**	3.27	0.51-20.96	**6.06**	**1.75-21.00**
	10-19 years v. 30+ years*	2.21	0.99-4.94	2.36	0.31-17.77	2.12	0.85-5.31	2.38	0.41-13.97	2.47	0.96-6.33
	20-29 years v. 30+ years*	2.05	0.99-4.25	3.31	0.58-19.01	1.87	0.81-4.28	4.15	0.80-21.46	1.63	0.67-3.93
PANSS total score		**0.96**	**0.95-0.98**	**0.94**	**0.91-0.96**	**0.97**	**0.96-0.99**	**0.93**	**0.90-0.96**	**0.98**	**0.96-0.99**
Adjusted percent (95% CI) of successful completion of RMET-CV-R computed from regression											
UT		55.4% (50.0%-60.8%)		80.1% (72.6%-87.7%)		42.5% (35.6%-49.4%)		79.8% (73.0%-86.6%)		38.7% (31.5%-45.9%)	
TC		62.1% (54.6%-69.6%)		71.6% (60.7%-82.5%)		56.0% (46.0%-66.0%)		70.7% (60.6%-80.8%)		54.3% (43.3%-65.3%)	

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes; PANSS, Positive and Negative Syndrome Scale

statistically significant results are **bolded**

* denotes the reference group

[a] Two cases (one in untreated cases group and one in treated control group) were excluded from the analysis as outliers because there were only two cases in the education group '11 years or above'

Supplementary Table 2A. Multivariate linear regression comparing RMET-CV-R scores (percent correct responses about emotions in 70 pictures) between untreated cases (UT) and treated controls (TC) in the full sample and in different subsamples, including duration of illness as a covariate

	All respondents who successfully completed RMET-CV-R n=224		Respondents who successfully self-completed RMET-CV-R (self-completion subsample) n=105		Respondents who successfully completed RMET-CV-R administered RMET-CV-R by interviewers (interviewer-completion subsample) n=118[a]		Respondents who successfully completed RMET-CV-R who were with ≥ 5 years of schooling and <60 (educated-younger subsample) n=125		Respondents who successfully completed RMET-CV-R with <5 years of schooling or ≥ 60 (undereducated-older subsample) n=99	
	Coef	P value	Coef	P value	Coef	P value	Coef	P value	Coef	P value
Group TC v. UT*	**0.042**	**0.016**	**0.066**	**0.015**	0.017	0.440	**0.058**	**0.025**	0.021	0.413
Age in years	**-0.002**	**0.034**	-0.0004	0.824	-0.002	0.054	-0.001	0.627	**-0.003**	**0.035**
Gender (1=female*, 2=male)	0.036	0.053	0.016	0.558	0.044	0.084	0.031	0.231	0.048	0.106
Years of schooling 0 years	Reference group		----	----	Reference group		----	----	Reference group	
1-4 years	**0.079**	**<0.001**	----	----	**0.072**	**<0.001**	----	----	**0.069**	**0.001**
5-7 years	**0.132**	**<0.001**	Reference group		**0.082**	**0.020**	Reference group		0.066	0.225
8-10 years	**0.150**	**<0.001**	0.036	0.246	0.058	0.081	0.013	0.639	**0.108**	**0.020**
11+ years	**0.169**	**<0.001**	0.038	0.405	No observation		0.034	0.449	No observation	
Residence (1=urban*, 2=rural)	**-0.077**	**0.004**	**-0.089**	**0.017**	-0.062	0.067	**-0.081**	**0.029**	**-0.090**	**0.017**
Ethnicity (1=Han*, 2=minority group)	-0.028	0.097	-0.010	0.726	-0.037	0.077	-0.005	0.856	**-0.054**	**0.021**
Duration of illness <10 years v. 30+years*	0.002	0.947	-0.011	0.894	0.023	0.544	0.004	0.951	0.012	0.760
10-19 years v. 30+ years*	0.005	0.881	0.022	0.765	-0.025	0.400	0.026	0.703	-0.027	0.387
20-29 years v. 30+ years*	-0.009	0.755	0.005	0.936	0.003	0.912	-0.009	0.879	-0.010	0.769
Adjusted mean (95% CI) RMET-CV-R score computed from regression										
UT	36.3% (33.8%-38.8%)		42.7% (38.6%-46.9%)		30.9% (28.0%-33.7%)		41.3% (37.5%-45.2%)		30.0% (27.0%-33.1%)	
TC	40.5% (38.4%-42.7%)		49.4% (46.1%-52.6%)		32.6% (29.5%-35.7%)		47.1% (44.0%-50.2%)		32.1% (28.5%-35.6%)	

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes

statistically significant results are **bolded**

* denotes the reference group

[a] One case (in treated control group) was excluded from the analysis as an outlier because it is the only case in the education group '11 years or above'

Supplementary Table 2B. Multivariate linear regression comparing RMET-CV-R scores (percent correct responses about emotions in 70 pictures) between untreated cases (UT) and treated controls (TC) in the full sample and in different subsamples, including duration of illness and PANSS total score as covariates

| | | All respondents who successfully completed RMET-CV-R

n=224 | | Respondents who successfully self-completed RMET-CV-R (self-completion subsample)

n=105 | | Respondents who successfully completed RMET-CV-R administered RMET-CV-R by interviewers (interviewer-completion subsample)

n=118[a] | | Respondents who successfully completed RMET-CV-R who were with ≥ 5 years of schooling and <60 (educated-younger subsample)

n=125 | | Respondents who successfully completed RMET-CV-R with <5 years of schooling or ≥ 60 (undereducated-older subsample)

n=99 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Coef | P value | Coef | P value | Coef | P value | Coef | P value | Coef | P value |
| Group | TC v. UT* | 0.031 | 0.094 | **0.064** | **0.022** | 0.005 | 0.845 | **0.053** | **0.047** | 0.007 | 0.798 |
| Age in years | | **-0.002** | **0.049** | -0.0004 | 0.852 | -0.002 | 0.058 | -0.001 | 0.705 | **-0.003** | **0.038** |
| Gender (1=female*, 2=male) | | 0.036 | 0.051 | 0.017 | 0.558 | 0.045 | 0.078 | 0.031 | 0.232 | 0.049 | 0.095 |
| Years of schooling | 0 years | Reference group | | ---- | ---- | Reference group | | ---- | ---- | Reference group | |
| | 1-4 years | **0.077** | **<0.001** | ---- | ---- | **0.070** | **0.001** | ---- | ---- | **0.066** | **0.002** |
| | 5-7 years | **0.126** | **<0.001** | Reference group | | **0.072** | **0.043** | Reference group | | 0.062 | 0.273 |
| | 8-10 years | **0.147** | **<0.001** | 0.037 | 0.257 | 0.057 | 0.095 | 0.015 | 0.614 | **0.096** | **0.037** |
| | 11+ years | **0.165** | **<0.001** | 0.038 | 0.407 | No observation | | 0.035 | 0.439 | No observation | |
| Residence (1=urban*, 2=rural) | | **-0.078** | **0.004** | **-0.089** | **0.018** | -0.067 | 0.057 | **-0.082** | **0.027** | **-0.092** | **0.018** |
| Ethnicity (1=Han*, 2=minority group) | | -0.028 | 0.105 | -0.010 | 0.719 | -0.035 | 0.098 | -0.006 | 0.827 | **-0.049** | **0.032** |
| Duration of illness | <10 years v. 30+years* | 0.005 | 0.889 | -0.010 | 0.904 | 0.026 | 0.514 | 0.006 | 0.937 | 0.013 | 0.755 |
| | 10-19 years v. 30+ years* | 0.008 | 0.788 | 0.022 | 0.764 | -0.022 | 0.477 | 0.027 | 0.694 | -0.024 | 0.463 |
| | 20-29 years v. 30+ years* | -0.001 | 0.961 | 0.007 | 0.926 | 0.011 | 0.729 | -0.007 | 0.909 | -0.001 | 0.978 |
| PANSS total score | | -0.001 | 0.182 | -0.0001 | 0.905 | -0.001 | 0.157 | -0.0003 | 0.747 | -0.001 | 0.141 |
| ***Adjusted mean (95% CI) RMET-CV-R score computed from regression*** | | | | | | | | | | | |
| UT | | 36.9% (34.3%-39.5%) | | 42.8% (38.7%-47.0%) | | 31.5% (28.4%-34.5%) | | 41.6% (37.7%-45.4%) | | 30.7% (27.4%-33.9%) | |
| TC | | 40.0% (37.7%-42.3%) | | 49.3% (45.8%-52.7%) | | 31.9% (28.8%-35.0%) | | 46.9% (43.5%-50.2%) | | 31.3% (27.9%-34.8%) | |

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes; PANSS, Positive and Negative Syndrome Scale

statistically significant results are **bolded**

* denotes the reference group

[a] One case (in treated control group) was excluded from the analysis as an outlier because it is the only case in the education group '11 years or above'