

Supplemental Materials

Does Reading a Single Passage of Literary Fiction Really Improve Theory of Mind? An Attempt at Replication

by M. E. Panero et al., 2016, *Journal of Personality and Social Psychology*

<http://dx.doi.org/10.1037/pspa0000064>

Table S1 provides a complete list of all texts used, by experiment. Tables S2 and S3 include the mean scores and sample sizes for each text and research group for the RMET and ART, respectively. Table S4 lists all measures across the three research groups that were not reported in the main text (i.e., measures other than the RMET and the ART). Table S5 provides the number of participants excluded from each research group (and experiment), according to each exclusion criteria. Table S6 provides the number of participants excluded from each research group (and experiment), according to each reading condition.

Exclusion Procedures

Not completing the study. We excluded participants who did not complete both the RMET and the ART.

Reading times. We excluded participants with reading times greater than 3.5 standard deviations above and below the mean for each experiment. Participants with extremely long reading times might have simply left their browser open without reading the text. Those with extremely short reading times might not have actually read the assigned text.

RMET. We excluded participants with RMET scores greater than 3.5 standard deviations below the mean for each experiment.

ART. We excluded participants with guessing scores greater than 3.5 standard deviations above the mean for each experiment (i.e., selected an excessively high number of non-authors). It is important to note that this standardized exclusion procedure resulted in maintaining some

participants who did not correctly identify any authors and some who selected more non-authors than authors.

Other exclusions. Kidd and Castano (2013) additionally asked participants if they were native speakers of English and if they had participated in the study previously, and they excluded participants who were not native speakers and who had previously participated. Only Research group 1 included questions about these two issues; participants who said that they had previously participated or who reported not being native speakers were not allowed to continue with the procedure.

Research Group 1

Research Group 1 contributed four experiments and 540 participants to the replication effort. Although Research Group 1 replicated all of the five experiments in Kidd and Castano (2013) using the same texts, Experiment 2 ($n = 103$) was not included in the main analyses because it did not use the RMET. However, preliminary analyses of the experiment 2 measures did not reveal any significant findings. Additional measures used by Research Group 1 and Kidd and Castano (2013), but not included in the main analyses, are briefly described here.

The Vicki Test is a measure of cognitive Theory of Mind (Converse, Lin, Keysar, & Epley, 2008) in which participants are presented with a false belief scenario and are tested on their understanding of a character's false belief.

The Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988) consists of two 10-item mood scales that discriminate between positive or negative emotional experiences. Participants are given a list of 20 emotions and asked to indicate the extent to which they feel each emotion at the present moment.

The Narrative Transportation Scale (Green & Brock, 2000) assesses emotional involvement in the story, cognitive attention to the story, feelings of suspense, lack of awareness of surroundings, and mental imagery. Participants rate the extent to which they had the described experiences while reading the designated narrative.

The Diagnostic Analysis of Nonverbal Accuracy - Adult Faces Test (Nowicki, 2010) is an affective ToM task. Participants choose the emotion term corresponding to the emotion (anger, fear, sadness, or happiness) shown in a full face.

The Yoni task (Shamay-Tsoory & Aharon-Peretz, 2007) assesses both cognitive and affective ToM. For each trial, participants see a central character and four images. Participants identify which image is the one that the character is thinking about, wants, likes, loves, dislikes, etc.

Research Group 2

Research Group 2 contributed one experiment and 218 participants to the replication effort. All participants were recruited from mTurk as described in the main text, and participation was limited to the United States. An additional 131 individuals were not included in the analyses because they read narrative nonfiction texts not used in Kidd & Castano (2013). Unlike those used in the original study, these nonfiction texts were about people. These were an excerpt from the autobiographical book *The Kid* by Dan Savage and two excerpts from 1966 Esquire Magazine articles, “Frank Sinatra Has a Cold” (changing Sinatra’s name to “James Michaels”) and “M Company.”

Participants who read literary fiction or narrative nonfiction were assigned to one of four between-subjects conditions, constructed using a 2 (Actually Read: fiction, nonfiction) x 2 (Text Described As: fiction, nonfiction) design. Their instructions described their assigned text as

either an excerpt from a fiction text that was published as a story in a literary magazine or an excerpt from a nonfiction text that was published as an article in a news magazine. Preliminary analyses found no effect of these instructions on RMET or ART scores, and thus participants were included in the analyses in the main paper based on what they actually read, regardless of what they were told.

Immediately following the instructions, participants were asked to report whether their text was fiction or nonfiction, as a manipulation check. After reading, participants completed five measures in addition to the RMET and ART. First, in a test of higher-order cognitive theory of mind (Kinderman, Dunbar & Bentall, 1998), participants read a vignette about several characters' interactions and responded to a series of true/false questions about characters' beliefs. Second, the Positive and Negative Affect Schedule (described above) measured current affect (i.e. happy, sad, angry, scared, surprised, or disgusted) on a 5-point scale. Third, participants who read narrative texts also completed the Narrative Transportation Scale (described above). Fourth, participants who read texts responded to three multiple-choice memory questions. Fifth, participants who read texts were asked to report whether their text was fiction or nonfiction and to judge how certain they were on a 5-point scale.

Research Group 3

Research Group 3 contributed one experiment and 34 participants to the replication effort. An additional 34 participants were excluded from the analyses reported in the paper due to having read texts not used in Kidd and Castano (2013). These were "Puppy" (Saunders, 2007; literary fiction) and "The CIA's Most Highly-Trained Spies Weren't Even Human" (Vanderbilt, 2013; nonfiction). Before reading, participants were told that they would be asked to write a short summary of the story. Immediately after reading, they completed the RMET, then wrote

the aforementioned summary, then completed the Narrative Transportation Scale (described above) with reference to the story they had just read. Finally, they completed the ART.

Table S1

Text and Reading Condition Per Research Group and Experiment

Text	Reading Condition	Research Group	Experiment
Blind Date (Davis, 2009)	Literary Fiction	1, KC	1
Chameleon (Chekhov, 1979)	Literary Fiction	1, KC	1
Corrie (Munro, in Furman, 2012)	Literary Fiction	1, 2, KC	3, 4, 5
Leak (Ruddick, in Furman, 2012)	Literary Fiction	1, KC	3
Nothing Living Lives Alone (“NLLA”; Berry, in Furman, 2012)	Literary Fiction	1, KC	3, 4
The Round House (Erdrich, 2012)	Literary Fiction	1, KC	2
The Runner (Delillo, 1988)	Literary Fiction	1, 3, KC	1
Salvage the Bones (Ward, 2011)	Literary Fiction	1, KC	2
The Tiger’s Wife (Obreht, 2011)	Literary Fiction	1, KC	2
Uncle Rock (Gilb, in Furman, 2012)	Literary Fiction	1, 2, KC	4, 5
The Vandercook (Mattinson, in Furman, 2012)	Literary Fiction	1, 2, KC	5
Cross Roads (Young, 2012)	Popular Fiction	1, KC	2
Gone Girl (Flynn, 2012)	Popular Fiction	1, KC	2
Jane (Rinehart, in Hoppenstand, 1998)	Popular Fiction	1, KC	4, 5
Lalla (Pilcher in Hoppenstand, 1998)	Popular Fiction	1, KC	3
The Sins of the Mother (Steel, 2012)	Popular Fiction	1, KC	2
Space Jockey (Heinlein in Hoppenstand, 1998)	Popular Fiction	1, KC	3, 4, 5
Too Many Have Lived (“TMHL”; Hammett in Hoppenstand, 1998)	Popular Fiction	1, KC	3, 4, 5
Bamboo Steps Up (“Bamboo”; Gandel, 2008)	Nonfiction	1, 2, KC	1

How the Potato Changed the World ("Potato"; Mann, 2011)	Nonfiction	1, 2, KC	1
The Story of the Most Common Bird in the World ("Bird"; Dunn, 2012)	Nonfiction	1,2,3, KC	1

Note. Kidd and Castano (2013) denoted as KC. Abbreviations used for long titles in parentheses.

Table S2

Mean RMET Scores, Standard Deviations, and Sample Sizes for Each Text, by Research Group (Experiment)

Text	Research Group (Experiment)	<i>M</i>	<i>SD</i>	<i>N</i>
Literary Fiction				
Blind Date	1 (1)	27.57	6.11	14
Blind Date	KC (1)	27.29	3.54	14
Chameleon	1 (1)	27.44	4.16	18
Chameleon	KC (1)	24.92	6.21	13
Corrie	1 (3)	27.67	5.45	15
Corrie	1 (4)	24.84	8.34	19
Corrie	1 (5)	28.13	4.50	15
Corrie	2	26.79	4.26	43
Corrie	KC (3)	25.09	4.93	11
Corrie	KC (4)	27.00	4.59	14
Corrie	KC (5)	26.73	3.54	33
Leak	1 (3)	26.07	6.94	14
Leak	KC (3)	25.33	3.90	15
NLLA	1 (3)	25.09	7.74	11
NLLA	1 (4)	21.44	8.61	16
NLLA	KC (3)	28.29	1.98	7
NLLA	KC (4)	23.78	6.74	9
The Runner	1 (1)	27.85	3.08	13
The Runner	3	26.94	3.80	17
The Runner	KC (1)	25.69	3.07	16
Uncle Rock	1 (4)	24.56	6.21	25
Uncle Rock	1 (5)	26.56	6.19	18
Uncle Rock	2	26.89	4.57	45
Uncle Rock	KC (4)	26.80	5.18	10
Uncle Rock	KC (5)	25.91	3.44	44
Vandercook	1 (5)	28.17	3.68	18
Vandercook	2	24.90	6.10	41
Vandercook	KC (5)	26.31	3.84	39
Popular Fiction				

Jane	1 (4)	28.16	4.97	19
Jane	1 (5)	27.17	5.90	18
Jane	KC (4)	26.27	2.87	11
Jane	KC (5)	25.55	3.95	31
Lalla	1 (3)	24.69	7.40	13
Lalla	KC (3)	24.46	4.31	13
Space Jockey	1 (3)	25.29	7.16	14
Space Jockey	1 (4)	20.47	10.01	17
Space Jockey	1 (5)	27.71	6.44	17
Space Jockey	KC (3)	21.63	8.19	8
Space Jockey	KC (4)	20.08	6.08	12
Space Jockey	KC (5)	25.24	4.41	42
TMHL	1 (3)	26.80	4.97	15
TMHL	1 (4)	25.73	7.19	22
TMHL	1 (5)	27.88	5.93	17
TMHL	KC (3)	23.00	6.53	15
TMHL	KC (4)	24.81	4.02	16
TMHL	KC (5)	24.80	4.48	35
Nonfiction				
Bamboo	1 (1)	27.94	4.14	16
Bamboo	2	25.15	7.66	13
Bamboo	KC (1)	20.30	5.52	10
Bird	1 (1)	28.06	3.61	16
Bird	2	26.12	4.89	17
Bird	3	27.71	2.64	17
Bird	KC (1)	24.61	5.34	18
Potato	1 (1)	28.83	5.17	12
Potato	2	25.94	6.57	18
Potato	KC (1)	23.87	4.15	15
Control				
No-reading	1 (5)	26.41	5.80	148
No-reading	2	24.80	6.15	41
No-reading	KC (5)	25.09	4.70	133

Table S3

Mean ART Scores, Standard Deviations, and Sample Sizes for Each Text, by Research Group (Experiment)

Text	Research group (Experiment)	<i>M</i>	<i>SD</i>	<i>N</i>
Literary Fiction				
Blind Date	1 (1)	15.50	13.02	14
Blind Date	KC (1)	20.07	12.69	14
Chameleon	1 (1)	24.61	16.05	18
Chameleon	KC (1)	21.54	15.63	13
Corrie	1 (3)	22.13	15.16	15
Corrie	1 (4)	17.95	11.62	19
Corrie	1 (5)	19.60	10.03	15
Corrie	2	19.98	16.29	43
Corrie	KC (3)	25.73	14.32	11
Corrie	KC (4)	18.79	14.06	14
Corrie	KC (5)	21.76	15.13	33
Leak	1 (3)	22.29	14.92	14
Leak	KC (3)	34.67	15.13	15
NLLA	1 (3)	18.64	14.11	11
NLLA	1 (4)	9.25	10.82	16
NLLA	KC (3)	32.00	15.22	7
NLLA	KC (4)	15.33	10.87	9
The Round House	1 (2)	22.08	13.40	12
The Round House	KC (2)	13.08	8.04	12
The Runner	1 (1)	16.77	12.06	13
The Runner	3	8.24	4.41	17
The Runner	KC (1)	17.94	14.73	16
Salvage the Bones	1 (2)	22.89	11.63	9
Salvage the Bones	KC (2)	19.46	13.14	13

The Tiger's Wife	1 (2)	19.38	13.40	12
The Tiger's Wife	KC (2)	21.23	13.40	13
Uncle Rock	1 (4)	11.12	11.19	25
Uncle Rock	1 (5)	19.28	13.54	18
Uncle Rock	2	20.96	14.10	45
Uncle Rock	KC (4)	22.70	18.87	10
Uncle Rock	KC (5)	24.57	14.79	44
Vandercook	1 (5)	17.06	10.33	18
Vandercook	2	18.49	15.76	41
Vandercook	KC (5)	21.28	11.36	39
<hr/> Popular Fiction <hr/>				
Cross Roads	1 (2)	28.00	14.29	11
Cross Roads	KC (2)	19.00	10.91	10
Gone Girl	1 (2)	19.64	12.64	11
Gone Girl	KC (2)	20.15	9.41	13
Jane	1 (4)	16.11	17.71	19
Jane	1 (5)	14.72	12.44	18
Jane	KC (4)	20.36	13.32	11
Jane	KC (5)	25.42	14.24	31
Lalla	1 (3)	20.38	15.74	13
Lalla	KC (3)	20.15	12.94	13
The Sins of the Mother	1 (2)	25.29	14.97	14
The Sins of the Mother	KC (2)	18.27	12.92	15
Space Jockey	1 (3)	20.29	12.98	14
Space Jockey	1 (4)	9.18	9.84	17
Space Jockey	1 (5)	15.71	11.12	17
Space Jockey	KC (3)	24.00	9.83	8
Space Jockey	KC (4)	20.33	18.28	12
Space Jockey	KC (5)	20.21	13.29	42

TMHL	1 (3)	20.33	11.17	15
TMHL	1 (4)	17.41	13.43	22
TMHL	1 (5)	22.59	13.51	17
TMHL	KC (3)	25.00	12.41	15
TMHL	KC (4)	17.94	10.40	16
TMHL	KC (5)	22.60	14.34	35
Nonfiction				
Bamboo	1 (1)	21.75	13.65	16
Bamboo	2	15.00	11.45	13
Bamboo	KC (1)	15.10	10.35	10
Bird	1 (1)	27.31	15.20	16
Bird	2	24.06	14.83	17
Bird	3	9.82	9.01	17
Bird	KC (1)	22.72	15.77	18
Potato	1 (1)	21.50	11.91	12
Potato	2	14.89	10.99	18
Potato	KC (1)	16.80	12.87	15
Control				
No-reading	1 (5)	14.31	11.60	148
No-reading	2	15.98	12.75	41
No-reading	KC (5)	20.30	13.27	133

Table S4

Additional Measures Administered per Research Group and Experiment

Measure	Research Group(s)	Experiment
Vicki Test	1, KC	1, 2
Positive and Negative Affect Schedule	1, 2, KC	1, 3, 5
Narrative Transportation Scale	1, 2, 3, KC	1, 2, 5
Diagnostic Analysis of Nonverbal Accuracy - adult faces test	1, KC	2
Yoni Task	1, KC	4, 5
A single item about current sadness	1, KC	3
Two items assessing current feelings of sadness and happiness	1, KC	5
Two items assessing enjoyment of text and agreement that it was “an example of excellent literature”	1, KC	5
Higher-order theory of mind reasoning	2	
Three multiple-choice memory questions	2	
Report whether text read was fiction or nonfiction, and certainty of judgement	2	
Write short summary of the assigned text	3	

Table S5

Number of participants in research group and experiment per exclusion criteria

Research group (Experiment)	Participants Recruited	Dropped out before being assigned text	Excluded for reading a story not in KC	Excluded for not completing the study	Excluded for short or long reading time	Excluded for low outlier on the RMET	Excluded for high rate of guessing on ART	Final <i>n</i>
1 (1)	115	4	0	20	1	0	1	89
1 (3)	123	4	0	37	0	0	0	82
1 (4)	204	21	0	63	0	0	2	118
1 (5)	371	0	0	116	0	0	4	251
2	420	17	131	48	2	1	3	218
3	68	0	34	0	0	0	0	34
Total	1,301	46	165	284	3	1	10	792

Table S6

Number of participants in research group and experiment excluded per reading condition

Research group (Experiment)	Nonfiction	Literary Fiction	Popular Fiction	No Reading
1 (1)	12	10	n/a	n/a
1 (3)	n/a	19	18	n/a
1 (4)	n/a	36	29	n/a
1 (5)	n/a	23	37	60
2	15	33	n/a	6
3	17	17	n/a	n/a
Total	44	138	84	66

References

- Converse, B. A., Lin, S., Keysar, B., & Epley, N. (2008). In the mood to get over yourself: Mood affects theory-of-mind use. *Emotion* 8, 725–730.
- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology* 79, 701–721.
- Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, 342(6156), 377–380.
- Kinderman, P., Dunbar, R., & Bentall, R. P. (1998). Theory-of-mind deficits and causal attributions. *British Journal of Psychology*, 89(2), 191–204. doi:10.1111/j.2044-8295.1998.tb02680.x
- Nowicki, S. (2010). Manual for the Receptive Tests of the Diagnostic Analysis of Nonverbal Accuracy 2 (Department of Psychology, Emory University, Atlanta, GA).
- Shamay-Tsoory, S. G. (2008). Recognition of ‘fortune of others’ emotions in Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders* 38, 1451–1461.
- 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, 1063–1070.