Appendix A. Experimental Stimuli and best linear unbiased predictors (BLUPs).

Morals

Positive-consensus	ToM BLUPs	VMPFC BLUPs	Fact-like BLUPs	Moral-like BLUPs	Preference-like BLUPs
The deplorable conditions of Chinese electronics workers should not be ignored.	0.011	1.1269	3.0221	6.2564	3.8453
Driving after drinking heavily is a stupid and selfish way to behave.	-0.5251	0.5227	3.9144	5.9972	4.0798
Professors should not tolerate students cheating on their exams.	0.2148	0.2742	3.2157	5.9815	4.303
It is irresponsible for airlines to risk the safety of their passengers.	-0.7352	-0.0465	3.8749	6.2122	3.2389
Parents should be willing to make sacrifices for the benefit of their baby.	0.1883	1.5678	3.0313	5.7222	4.8858
The goal of sports should be to teach children that respect for others is more important than winning.	0.0277	0.9978	2.0062	4.5673	5.3466
No-consensus					
It is wrong to use animals as disposable space shuttle test pilots.	0.3939	1.7801	2.5669	6.14	4.6066
Dog racing is harmful and exploitative to the dogs being raced.	-0.2061	1.1528	3.3337	5.9298	4.1966
Destroying the habitats of owls through deforestation is deplorable.	-0.229	0.9305	2.497	6.1893	3.9591
It is unjust for businesses to allow apples to rot rather than giving them to the needy.	0.4754	0.8605	2.13	6.0441	4.4336
Music stores should prevent children from buying CDs with violent or sexist lyrics.	0.9306	2.5563	1.9186	5.1095	5.2449
Eating fish is acceptable if they were treated humanely when caught or raised.	-0.0027	0.4265	1.947	5.4241	5.0757
Good Americans buy American cars, such as Hummers.	0.371	2.2236	1.4816	2.6911	6.4033
It is wrong to knowingly buy sandals made using sweatshop labor.	0.0083	0.2	2.1262	6.1149	4.5567
People should help their elderly neighbors clear snow from their driveway.	1.0134	1.664	2.0092	5.6906	4.6681
Harry Potter should be banned from school libraries for idolizing witchcraft.	0.673	2.6323	1.2825	3.9891	5.6957
It is wrong to cheat when playing games such as Monopoly.	-0.0669	-0.0154	2.1859	6.1189	4.2002

It is unethical for businesses to promote sugary products to children.	0.0183	2.0761	2.1744	5.8517	4.3915
Negative-consensus					
It is wrong to harm cockroaches just because humans find them disgusting.	0.6888	1.2232	1.6661	5.3421	4.8423
Universal donors should be obligated to donate their blood.	0.4606	1.284	1.6699	4.8386	4.805
Sport fishing to kill and eat fish is barbaric and evil.	-0.07	1.4002	2.0179	5.5038	5.1079
Private beaches are immoral, as everyone should be able to share the space.	-0.0172	0.6177	1.7373	4.6917	5.2054
Child labor in coffee bean farming is acceptable because it lowers the market price.	-0.0665	1.8448	2.5858	5.7527	4.0372
It is fine for doctors to accidentally kill a small number of patients per year.	0.3621	-0.1727	1.5847	6.0397	3.8442
Preferences					
Positive-consensus					
Using touchscreens is a much more satisfying way to interact with computers.	-0.9189	0.1251	1.8004	1.1998	6.6632
Having a drink every now and then is a good way to relax.	0.6489	1.6982	1.9433	1.7279	6.4285
Professors who play videos make their classes more entertaining.	-0.1557	1.3853	1.9827	1.2996	6.4646
Going through airport security is an unpleasant experience.	0.3005	1.2072	2.372	1.3584	6.3961
Babies that are temperamental are aggravating to spend time around.	-0.5007	0.7931	1.971	1.4424	6.2982
Afterschool programs involving sports are more fun than most of the alternatives available to children.	-0.4045	1.4263	1.9626	1.3536	6.4338
No-consensus					
Gazing at planets through a telescope is a satisfying activity.	-0.1058	0.6164	1.8833	1.3411	6.5164
Dogs are not worth the stress and aggravation it takes to own them.	0.0287	0.648	1.5868	1.4261	6.5202
The "hoots" of owls in the woods make camping more enjoyable.	-0.9634	0.5373	1.6953	1.2657	6.5675
Green apples are too sour to be an enjoyable lunchtime snack.	-0.5099	0.4378	1.7474	1.1547	6.6263
Rock music is pleasing to the ear, and much more agreeable than rap music.	0.5904	1.8393	1.4535	1.195	6.5902
Sitting in a boat and fishing all day long is boring and a waste of time.	0.3776	1.0699	1.4302	1.2174	6.665

Nothing is more awesome than driving in a Hummer.	-0.573	1.0348	1.423	1.1657	6.7799
Because sandals have fewer styles, they are less fun to go shopping for.	-0.985	0.4179	1.8624	1.2667	6.5586
In the wintertime, it is fun to catch snowflakes on the tip of your tongue.	-0.5057	0.6827	1.93	1.2962	6.3668
The Harry Potter books are engaging and delightful to read, even for adults.	0.8586	1.1436	2.0413	1.3028	6.6514
Many games are better than Monopoly, which is incredibly boring.	-0.0913	1.3158	1.4839	1.2129	6.6934
Any ice cream flavor tastes better when served in a crunchy waffle cone.	-1.2491	-1.05	1.6208	1.2306	6.5764
Negative-consensus					
Cockroaches are delicious to eat because of their hard and crunchy shell.	-0.4857	-0.5003	1.6712	1.3205	6.6532
Having blood drawn is a pleasurable experience.	-0.4634	-0.2433	1.5413	1.3137	6.3696
Nothing is more appealing than the smell of rotting fish.	-1.413	-1.4443	1.4348	1.3312	6.5875
While at a hot beach, it is agonizing to dip your toes in the cool water.	-1.3909	-0.9091	1.8264	1.2195	6.301
Drinking coffee is a miserable experience when you are tired and need energy.	-0.2406	0.8269	1.5902	1.178	6.478
Having a doctor listen attentively to your medical concerns is awful.	-0.0912	1.4404	1.4676	1.5881	6.3524
Facts					
Positive-consensus					
Touchscreens are used in a variety of electronics, including smartphones.	-1.3451	-1.1744	6.689	1.1949	1.4992
A breathalyzer is used to determine whether a driver is intoxicated.	-1.3908	-1.3893	6.6404	1.5536	1.375
University professors teach classes but also conduct research.	-0.2091	0.5427	6.3226	1.2109	1.7311
Airplanes have wings that enable the plane to lift upwards.	-1.5847	-0.7983	6.6292	1.1615	1.277
In a full-term human pregnancy, babies spend nine months in a woman's womb.	-1.1268	-0.3927	6.7811	1.3927	1.206
In sports-based afterschool programs children participate in sports such as baseball or basketball to name a few.	-1.5024	-0.9464	6.2851	1.1473	2.0334

No-consensus					
Saturn's moon, Titan, is the only moon known to have clouds.	-0.8656	-0.2341	6.7416	1.1645	1.2553
The dog breed, Basenji, is the world's only barkless dog breed.	-1.9031	-1.3962	6.7146	1.2293	1.3857
Of all types of birds, owls are the ones that can see the color blue.	-1.2773	-0.0252	6.5033	1.1704	1.221
Newtown Pippin was the first apple variety exported from the US.	-0.9766	0.02	6.5508	1.2877	1.3402
The first CD made for commercial release was the rock CD: "Born in the USA."	-0.7217	0.7578	6.3238	1.2247	1.4223
There are more fish species in the Amazon River than in the Atlantic Ocean.	-1.3589	-1.3786	6.2895	1.2685	1.2493
Hummer trucks were first marketed to civilians in 1990.	-0.9593	0.3892	6.6263	1.1715	1.3217
The oldest sandals in the world were found in Oregon's Paisley Caves.	-1.5318	-1.2682	6.5087	1.2753	1.4208
A town in North Dakota holds the world record for the tallest snowman.	-0.8145	-0.6858	6.6517	1.1539	1.2155
The author J.K. Rowling has two younger siblings, one brother and one sister.	-0.226	0.5963	6.6338	1.2753	1.1995
Monopoly pieces were made from wood, not metal, during WWI.	-0.699	0.4126	6.6678	1.1593	1.2906
The very first waffle cone was invented in Chicago, Illinois, at a state fair.	-0.8793	-0.0489	6.5895	1.2161	1.3635
Negative-consensus					
Cockroaches are a type of cold-blooded reptiles related to snakes.	-1.7082	-1.072	5.6398	1.2107	1.4144
In humans, the liver pumps blood throughout the body.	-1.8077	-1.992	5.6126	1.2347	1.2591
Fish are able to live outside of water for an extended time.	-1.1876	-0.5137	5.7055	1.4063	1.2456
The sand on beaches is usually transported there from nearby deserts.	-0.8853	0.0665	6.11	1.2452	1.2927
Coffee beans grow particularly well in freezing cold climates, such as Alaska and Russia.	-1.4569	-1.0238	5.8752	1.1975	1.4958
Medical students at hospitals are able to perform surgeries with little to no training.	-0.2611	0.9317	4.6601	1.734	1.6692

ToM BLUPs average estimates for DMPFC, PC, RTPJ, and LTPJ, as all by-stimulus random slopes were perfectly correlated. For model details, see Tables S4 and S6 of the supplemental online materials.

Appendix B. List of covariates, with descriptions.

Semantic/syntactic measures (2.2.6)		
Question name	**Source**	**Description**
Word count	Coh Metrix 3.0	Number of words in statement.
Flesch reading ease	Coh Metrix 3.0	Measures reading difficult through the average sentence length and number of syllables per word. Higher scores indicate more difficulty.
Anaphor reference	Coh Metrix 3.0	Measures the number of times a single idea is referenced by counting the use of anaphors (e.g. pronouns: he, she, it; ellipsis markers: did, was).
Intentional verb incidence	Coh Metrix 3.0	Measures intentional information by counting verbs categorized as intentional by Wordnet ratings (Fellbaum, 1998; Miller et al., 1990).
Causal verb incidence	Coh Metrix 3.0	Measures causal information by counting verbs categorized as causal by WordNet ratings.
Causal verb ratio	Coh Metrix 3.0	Measures the cohesion of causal events to actors through the ratio of causal particles (e.g. because, if) to causal verbs. Higher scores indicate increased cohesion and easier readability.
Noun concreteness	Coh Metrix 3.0	Measures concreteness of content words (e.g. chair is high in concreteness, democracy is low) using the mean concreteness ratings of content words, taken from human ratings in the MRC Psycholinguistics Database (Coltheart, 1981).
Noun familiarity	Coh Metrix 3.0	Measures the familiarity of content words using the mean familiarity ratings of all content words, taken from human ratings in the MRC Psycholinguistic Database.
Noun imageability	Coh Metrix 3.0	Measures the imageability of content words using the mean familiarity ratings of all content words, taken from human ratings in the MRC Psycholinguistic Database.
Negation density	Coh Metrix 3.0	Provides a measure of syntactic complexity (i.e. working memory load) through the count of negative expressions in the text (e.g. not, un-).
Number of modifiers	Coh Metrix 3.0	Provides a measure of syntactic complexity (i.e. working memory load) through the mean number of modifiers per noun phrase.
Left embeddedness	Coh Metrix 3.0	Provides a measure of syntactic complexity (i.e. working memory load) through the mean number of words before the main verb in a sentence.
Reaction time	In-scanner N = 25	The time from the appearance of the in-scanner agreement rating prompt to the input of a response by the participant.
Online Item Features		
Agreement	Study 1 (N = 49)	"To what extent do you agree / disagree with this statement?" (1-7; "strongly disagree"-"strongly agree").
Valence	Online sample (N = 42)	Valence was the difference between unipolar positive and negative ratings (Kron et al., 2013), described below: *Instructions:* "Please rate your feelings regarding this statement using the following two scales. An extreme

		unpleasant rating means you feel completely unpleasant, unhappy, annoyed, unsatisfied, melancholic, or despaired. An extreme pleasant rating means you feel completely pleased, happy, satisfied, content or hopeful.” *Ratings*: Negative valence (1-8; “no unpleasant feelings”-“strong unpleasant feelings”) and positive valence (1-8; “no pleasant feelings”-“strong pleasant feelings”).
Arousal	Online sample (N = 42)	Arousal was the sum of unipolar positive and negative ratings, described above. Recent work has demonstrated that summed unipolar valence ratings are highly correlated with physiological measures of arousal, and may be superior to separately measuring arousal (Kron et al., 2013).
Mental imagery	Online sample (N = 46)	“To what extent did you picture or imagine what the statements described as you read?” (1-7; “Very Little”-“Very Much”; Dodell-Feder et al., 2011).
Mental state	Online sample (N = 48)	“To what extent did this statement make you think about someone’s experiences, thoughts, beliefs and/or desires?” (1-7; “Very Little”-“Very Much”; Dodell-Feder et al., 2011).
Mental States (of Others)	Online sample (N = 44)	To what extent did this statement make you think about the experiences, thoughts, beliefs, and/or desires OF OTHER PEOPLE? (1-7; “Very Little”-“Very Much”)
Mental States (of Self)	Online sample (N = 46)	“To what extent did this statement make you think about YOUR OWN experiences, thoughts, beliefs, and/or desires?” (1-7; “Very Little”-“Very Much”)
Person present	Online sample (N = 48)	“Does this statement mention people or a person?” (“Yes” / “No”).

Coh Metrix ratings are calculated using an online tool at http://cohmetrix.com (Graesser et al., 2004; McNamara et al., 2014). In online samples, participants who did not correctly answer a catch question (asking them to describe any of the 72 statements they had read) were excluded from analysis. This caused some variability in N across covariates.

Supplemental Materials

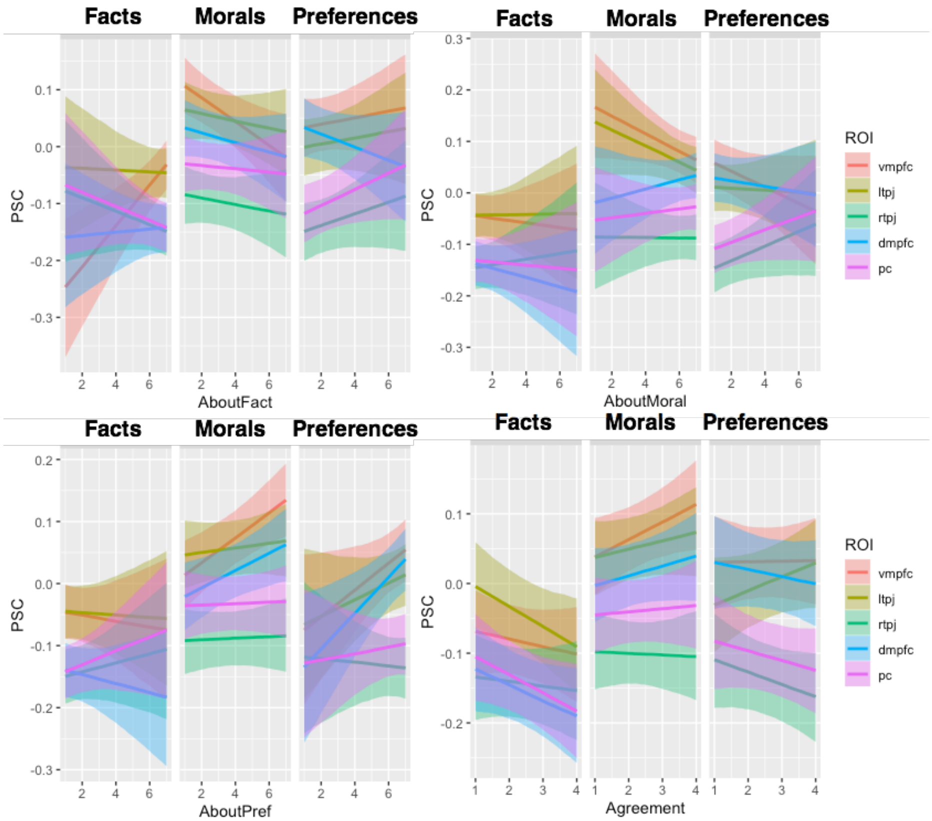

Figure S1. We ran a supplemental analysis, predicting percent signal change (PSC) in each ROI, evoked by fact/moral/preference statements using within-subject by-trial agreement ratings (collected in-scanner), and within-subject by-trial fact-/moral-/preference-like ratings (collected after the scan session). Due to the lack of data, and the ordinal nature of the unaveraged by-trial behavioral measures, we performed a multilevel Bayesian analysis, using *brms* in *R*. Marginal effects for each behavioral measure, within facts, morals, and preferences, and for each ROI are plotted below, and the models are available in our online data repository (https://osf.io/cx4dp/). On a visual inspection, the marginal effects within the moral domain appear either to be consistent with our reported findings, or to show no effect. In our paper, we found a negative association between ToMN activity and fact-/moral-like ratings, and a positive association with preference-like ratings (Section 3.2.3; Figure 3). In this analysis, for fact-like ratings (top left), VMPFC (red) shows a negative association, b = -0.005, 95%CI [-0.009, -0.001]. For moral-like ratings (top right) VMPFC (red) and LTPJ (yellow) both appear to show negative associations, although neither association excluded zero from a 95% credible interval: VMPFC, b = -0.004, 95%CI [-0.009, 0.0002]; LTPJ, b = -0.004, 95%CI [-0.009, 0.0005]. For preference-like ratings (bottom left), both VMPFC (red) and DMPFC (blue) appear to show a positive association with ToMN activity; and indeed, these associations both exclude zero from a 95% credible interval: VMPFC, b = 0.005, 95%CI [0.002, 0.008], DMPFC, b = 0.007, 95%CI [0.0008, 0.013]. The only discrepancy with the results reported in the main paper body was within agreement ratings (bottom right), which showed a positive association with VMPFC among morals, b = 0.006, 95%CI [0.00001, 0.012]—the opposite of what we observed in our reported results. We speculate that collecting measurements on a 4-point scale, or the collection of this data within the scanner settings (as opposed to outside of it) may account for this relationship. However, it is also worth noting that a concern in the main paper was that agreement and measures of objectivity were potentially confounded among morals, and for this reason, we find the opposite pattern of results observed here for agreement promising, rather than concerning.

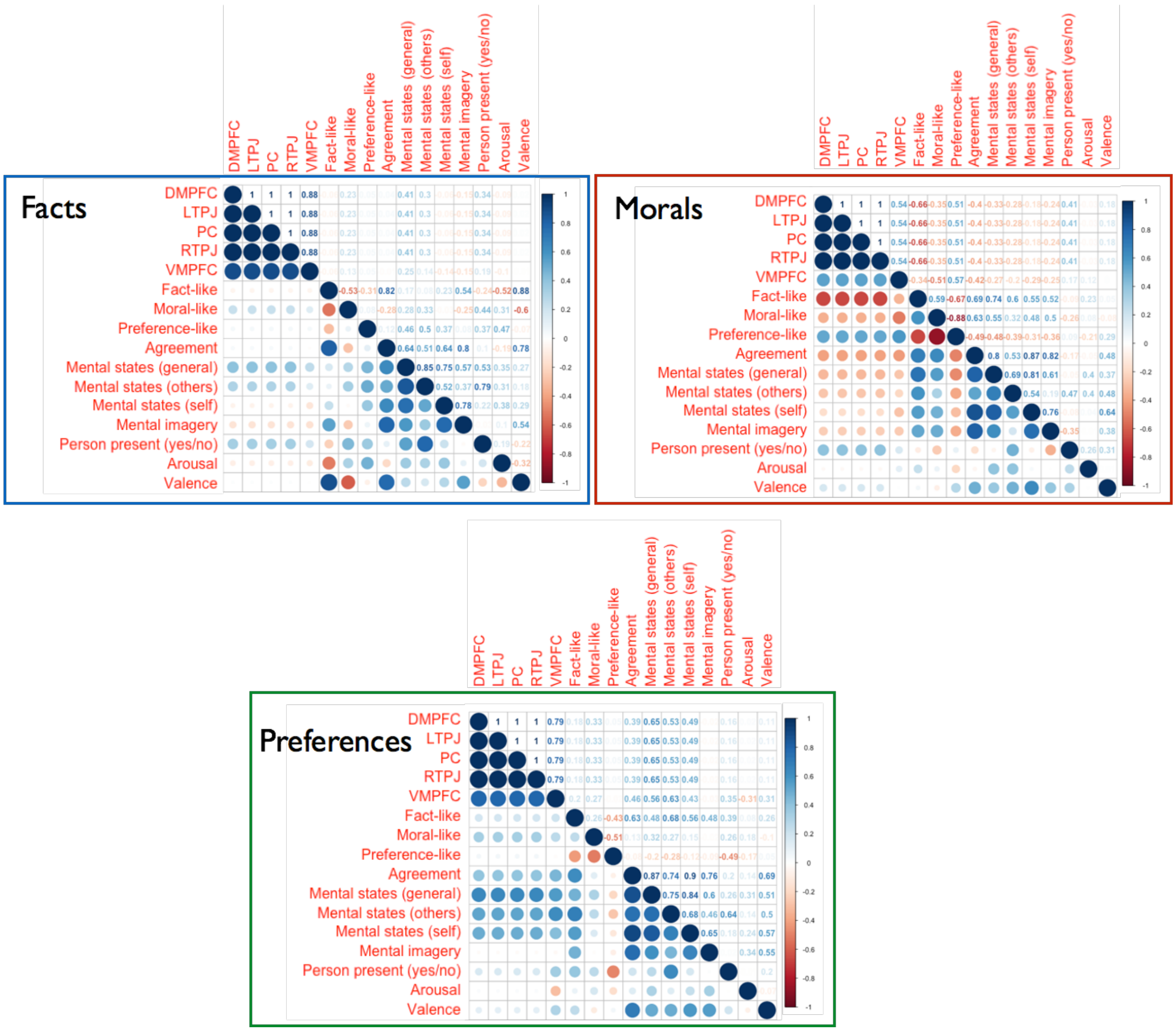

Figure S2. Correlation table of by-stimulus estimates.

Table S1. Study 1 condition means.

Metaethical judgments

Model:
Rating ~ Category * Rating-type * Consensus +
(1+ Category * Rating-type | ID) +
(1 + Rating-type | Item)

		Consensus		
		Positive-consensus	**No-consensus**	**Negative-consensus**
Category	**Rating-type**	Mean (SE)	Mean (SE)	Mean (SE)
Morals	**About Facts**	3.24 (0.21) ^{A}	2.13 (0.18) ^{B}	1.83 (0.21) ^{B}
	About Morals	5.78 (0.28)	5.46 (0.23)	5.33 (0.28)
	About Preferences	4.28 (0.28)	4.81 (0.25)	4.59 (0.28)
Facts	**About Facts**	6.59 (0.17) ^{A}	6.59 (0.17) ^{B}	5.53 (0.20) ^{B}
	About Morals	1.31 (0.24)	1.19 (0.18)	1.35 (0.24)
	About Preferences	1.58 (0.21)	1.29 (0.16)	1.36 (0.21)
Preferences	**About Facts**	2.03 (0.20	1.68 (0.17)	1.57 (0.20)
	About Morals	1.40 (0.24)	1.26 (0.18)	1.32 (0.24)
	About Preferences	6.46 (0.22)	6.60 (0.17)	6.43 (0.22)

Agreement ratings

Model:
Agreement ~ Category * Consensus +
(1+ Category | ID) +
(1 | Item)

	Consensus		
	Positive-consensus	**No-consensus**	**Negative-consensus**
Category	Mean (SE)	Mean (SE)	Mean (SE)
Fact	6.41 (0.40)	4.76 (0.31)	2.46 (0.42)
Moral	6.05 (0.40)	4.90 (0.29)	2.66 (0.39)
Preference	4.98 (0.40)	4.08 (0.28)	1.58 (0.40)

Mean estimates and standard errors are derived from contrasts within the models described in 2.2.1 and 2.2.2. Superscripts denote significant differences within each row (p values corrected for 27 comparisons; single-step method; $\alpha_{familywise} = .05$; single-step method).

Table S2. Study 2 in-scanner agreement ratings.

	Consensus		
	Positive-consensus	**No-consensus**	**Negative-consensus**
Category	Mean (SD)	Mean (SD)	Mean (SD)
Facts	3.64 (0.06)	2.68 (0.16)	1.47 (0.11)
Morals	3.64 (0.07)	2.81 (0.06)	1.89 (0.11)
Preferences	2.93 (0.07)	2.56 (0.06)	1.46 (0.12)

Mean and standard error (across participants). All comparisons, within content categories, were significant at $p < .001$.

Table S3. Theory of Mind network ROI coordinates.

Region	x	y	z	T score	K
R Temporoparietal Junction (RTPJ)	52	-60	24	10.55	3398
L Temporoparietal Junction (LTPJ)	-56	-56	28	9.69	3210
Precuneus (PC)	0	-52	40	10.81	2263
Ventromedial Prefrontal Cortex (VMPFC)	0	44	-20	7.69	533
Dorsomedial Prefrontal Cortex (DMPFC)	0	58	22	5.62	480
Additional significant clusters from localizer contrast					
Region	**x**	**y**	**z**	**T score**	**K**
R Cerebellum	32	-80	-38	7.42	425
R Inferior Frontal Gyrus	58	22	12	6.31	85
L Cerebellum	-18	-74	-34	6.30	597
L Thalamus	-8	-12	8	5.77	78
L Cerebellum	-6	-50	-50	5.70	54
R Superior Frontal Gyrus	8	38	54	5.64	60
R Middle Frontal Gyrus	30	30	52	5.60	158
R Occipital Pole	10	-96	4	5.57	626
L Middle Frontal Gyrus	-40	6	52	5.31	52
R Middle Frontal Gyrus	46	8	50	5.23	48
L Inferior Frontal Gyrus	-52	22	10	4.93	26
L Superior Frontal Gyrus	-10	36	56	4.92	61
R Lingual Gyrus	8	-48	-46	4.83	33
R Frontal Pole	16	52	36	4.74	75
L Frontal Pole	-2	62	4	4.72	19
Cingulate Gyrus	2	-16	40	4.61	17
L Occipital Fusiform Gyrus	-28	-88	-14	4.47	18
L Frontal Orbital Cortex	-42	22	-14	4.40	50
L Occipital Fusiform Gyrus	-12	-86	-16	4.30	32
Subcallosal Cortex	-4	18	-6	4.16	33
L Lateral Ventricle	-4	4	2	3.81	12

ROIs were a 9mm sphere around the reported coordinates. T scores represent difference scores in the false belief > false photograph contrast, in a random effects analysis across all subjects (*df* = 24). All coordinates are reported in MNI space. SPM contrast file is available at https://osf.io/cx4dp/.

Table S4. Study 1 behavioral rating model details.

ToM network activity

Model:
Rating ~ Category * Rating-type +
(1+ Category * Rating-type | ID) +
(1 + Rating-type | Item)

REML criterion at convergence: 32392.5

Dummy coded control conditions: Facts (category) & Fact-like ratings (rating-type)

Random effects structure (by-subject)

	Variance	St.Dev	Correlations							
			Intercept	Moral	Pref	Moral-like	Pref-like	M* M-like	M* P-like	P* Ml-like
Intercept	0.08	0.89								
Moral	2.52	1.59	-0.86							
Preference	2.37	1.54	-0.89	0.94						
Moral-like	1.94	1.39	-0.95	0.84	0.90					
Preference-like	2.24	1.50	-0.97	0.88	0.94	0.98				
M*Moral-like	7.23	2.69	0.84	-0.89	-0.87	-0.88	-0.87			
M*Pref-like	6.85	2.62	0.73	-0.82	-0.84	-0.77	-0.77	0.74		
P*Moral-like	2.34	1.53	0.91	-0.94	-1.00	-0.91	-0.95	0.89	0.85	
P*Pref-like	7.70	2.78	0.89	-0.90	-0.98	-0.94	-0.94	0.88	0.86	0.98

Random effects structure (by-stimulus)

	Variance	St.Dev	Correlations	
			Intercept	Moral-like
Intercept	.301	.549		
Moral-like	.349	.590	-0.58	
Pref-like	.750	.866	-0.90	0.27

Residual

Variance	St.Dev
1.10	1.05

Fixed Effects

Name	B (SE)	t(df)	p
Intercept	6.32 (0.17)	t(101.1) = 36.85	< .001 ***
Moral	-3.99 (0.28)	t(88.2) = 14.25	< .001 ***
Preference	-4.58 (0.27)	t(90.0) = 16.70	< .001 ***
Moral-like	-5.06 (0.24)	t(79.0) = 21.38	< .001 ***
Preference-like	-4.94 (0.28)	t(98.6) = 17.59	< .001 ***
Moral*Moral-like	8.24 (0.42)	t(66.1) = 19.40	< .001 ***
Preference*Moral-like	4.63 (0.28)	t(91.9) = 16.32	< .001 ***
Moral*Preference-like	7.23 (0.45)	t(86.2) = 15.93	< .001 ***
Preference*Preference-like	9.72 (0.47)	t(82.8) = 20.57	< .001 ***

St.Dev = standard deviation. *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .10$

Table S5. Study 2 behavioral rating means.

		Consensus		
		Positive-consensus	**No-consensus**	**Negative-consensus**
Category	**Rating-type**	Mean (SE)	Mean (SE)	Mean (SE)
Morals	**About Facts**	3.14 (0.24) [A]	2.30 (0.22) [B]	2.01 (0.24) [B]
	About Morals	6.35 (0.18) [A]	5.88 (0.17) [B]	5.87 (0.18) [B]
	About Preferences	4.23 (0.31)	4.31 (0.30)	4.34 (0.31)
Facts	**About Facts**	6.59 (0.19)	6.65 (0.18)	6.38 (0.19)
	About Morals	1.55 (0.21)	1.29 (0.20)	1.57 (0.21)
	About Preferences	1.85 (0.22) [A]	1.37 (0.21) [B]	1.68 (0.22) [AB]
Preferences	**About Facts**	2.52 (0.23)	1.95 (0.21)	2.14 (0.23)
	About Morals	2.11 (0.25)	1.74 (0.24)	1.90 (0.25)
	About Preferences	6.51 (0.19) [A]	6.58 (0.18) [B]	6.32 (0.19) [AB]

Mean and standard error are estimated using contrasts within the model defined in 3.2.1. Superscripts denote significant differences within each row (*p* values corrected for 27 comparisons; single-step method; $\alpha_{familywise} = .05$; single-step method).

Table S6. Study 2 ToM network mixed effects model.

ToM network activity

Model:
PSC ~ Category * ROI +
(1+ Moral + Preference + VMPFC + PC + RTPJ + LTPJ + Moral*(VMFPC+LTPJ) | ID) +
(1 + VMPFC | Item)

REML criterion at convergence: 2271

Dummy coded control conditions: Facts (category) & DMPFC (ROI)

Random effects structure (by-subject)

	Variance	St.Dev	Correlations							
			Intercept	**VMPFC**	**PC**	**RTPJ**	**LTPJ**	**Moral**	**Pref**	**M* VMPFC**
Intercept	.008	.091								
VMPFC	.010	.101	0.01							
PC	.008	.090	-0.60	-0.13						
RTPJ	.011	.105	-0.57	-0.20	0.71					
LTPJ	.007	.085	-0.46	-0.22	0.54	0.53				
Moral	.003	.053	-0.27	0.24	-0.34	-0.30	-0.41			
Pref	.002	.042	-0.07	0.18	-0.14	0.07	-0.11	0.52		
M*VMPFC	.002	.042	-0.14	-0.08	-0.43	-0.18	-0.41	0.62	0.66	
M*LTPJ	.002	.041	-0.33	0.30	0.59	0.35	0.64	-0.08	0.37	-0.30

Random effects structure (by-stimulus)

	Variance	St.Dev	Correlations
			Intercept
Intercept	.002	.048	
VMPFC	.004	.064	.03

Residual

Variance	St.Dev
.071	.027

Fixed Effects

Name	B (SE)	t(df)	p
Intercept	-0.147 (0.023)	*t*(46) = 6.27	< .001 ***
Moral	0.166 (0.023)	*t*(128) = 7.05	< .001 ***
Preference	0.160 (0.022)	*t*(169) = 7.11	<. 001 ***
VMPFC	0.084 (0.029)	*t*(51) = 2.93	.005 **
PC	0.005 (0.024)	*t*(47) = 0.20	.846
RTPJ	0.005 (0.026)	*t*(42) = 0.18	.858
LTPJ	0.105 (0.023)	*t*(45) = 4.55	< .001 ***
Moral*VMPFC	-0.035 (0.030)	*t*(97) = 1.18	.241
Moral*PC	-0.059 (0.022)	*t*(844) = 2.68	.007 **
Moral*RTPJ	-0.113 (0.022)	*t*(844) = 5.12	< .001 ***
Moral*LTPJ	-0.068 (0.024)	*t*(162) = 2.88	.004 **
Pref*VMPFC	-0.074 (0.029)	*t*(114) = 2.56	.012 *
Pref*PC	-0.123 (0.022)	*t*(844) = 5.56	< .001 ***
Pref*RTPJ	-0.152 (0.022	*t*(844) = 6.89	< .001 ***
Pref*LTPJ	-0.11 (0.022)	*t*(844) = 5.05	< .001 ***

St.Dev = standard deviation. *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .10$

Table S7. Study 2 ToM–behavioral analysis.

DV: Metaethical judgments (Study 1 BLUPs)			
Term		***F statistic***	***p***
PSC x 2 (ROI: ToM/VMPFC) x 3 (rating-type (fact-/moral-/preference-like) x 3 (category: fact/moral/preference)		$F(2, 405) = 5.73$	< .001 ***
PSC x 2 ROI x 3 rating-type x 2 (fact/preference)		$F(2, 270) = 0.01$	.986
PSC x 2 ROI x 3 rating-type x 2 (fact/moral)		$F(2, 270) = 7.42$	< .001 ***
PSC x 2 ROI x 3 rating-type x 2 (moral/preference)		$F(2, 270) = 6.64$	.002 **
Within moral statements			
PSC x 2 ROI x 3 rating-type		$F(2, 135) = 6.03$	.003**
PSC x 2 ROI x 2 (fact-like/preference-like)		$F(1, 90) = 13.77$	< .001 ***
PSC x 2 ROI x 2 (fact-like/moral-like)		$F(1, 90) = 0.81$	.371
PSC x 2 ROI x 2 (moral-like/preference-like)		$F(1, 90) = 5.47$	.022 *
Within preference-like ratings			
PSC x ROI		$F(1, 45) = 4.57$	.038 *
Within fact-/moral-like ratings			
PSC x ROI		$F(1, 92) = 8.80$	.004 **
Model: rating-type + (PSC x ROI) + (PSC x ROI x preference-like)	***B (SE)***	***t statistic***	**p**
Intercept (Fact-like rating)	2.57 (0.11)	$t(137) = 22.80$	< .001 ***
Moral-like	3.17 (0.14)	$t(137) = 23.00$	< .001 ***
Preference-like	1.80 (0.17)	$t(137) = 10.62$	< .001 ***
PSC (within fact-like/moral-like)	-1.01 (0.23)	$t(137) = 4.32$	< .001 ***
PSC x preference-like (within ToM)	1.94 (0.40)	$t(137) = 4.81$	< .001 ***
PSC x ROI (interaction for VMFPC, within fact-/moral-like)	0.72 (0.23)	$t(137) = 3.11$	.002 **
PSC x preference-like x ROI (interaction for VMPFC, for preference-like)	-1.35 (0.40)	$t(137) = 3.35$	.001 **
Contrasts:	***B (SE)***	***t statistic***	**p**
Fact-/moral-like–ToM relationship	-1.01 (0.23)	$t(140) = 4.32$	< .001 ***
Preference-like–ToM relationship	0.94 (0.33)	$t(140) = 2.85$	.020 *
Fact-/moral-like-VMPFC relationship	-0.28 (0.09)	$t(140) = 3.18$	.007 **
Preference-like–VMPFC relationship	0.31 (0.12)	$t(140) = 2.48$	.055 †

Contrast p values corrected for 4 comparisons; single-step method; $\alpha_{familywise} = .05$; single-step method. *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .10$

Table S8. Whole brain correlation peak coordinates.

Contrast	Name	Cluster Size	Peak T	x	y	z
Fact-like rating (negative)	M Superior frontal gyrus	1037	7.60	-8	24	58
			6.76	-18	20	62
			5.40	16	44	48
	L Middle frontal gyrus	658	7.22	-28	10	50
			6.24	-42	12	44
			5.33	-48	28	24
	L Angular gyrus	398	6.53	-44	-62	-48
			5.60	-38	-72	48
			5.23	-34	-62	48
	L Middle temporal gyrus	142	6.51	-58	-32	-16
			5.57	-66	-26	-12
	R Angular gyrus	412	6.43	44	-68	46
			5.76	50	-56	32
			5.74	52	-62	40
	R Middle frontal gyrus	239	5.58	40	20	44
			5.14	32	28	42
	L Medial caudate nucleus	132	5.18	-14	14	10
			5.16	-12	6	14
Fact-like rating (positive)	M Parietooccipital sulcus	145	4.85	8	-78	36
			4.83	-6	-80	18
			4.14	6	-84	44
Preference-like rating (positive)	R Angular gyrus	124	5.12	54	-60	34
			4.84	50	-62	42
			3.95	44	-68	46
	L Angular gyrus	102	4.98	-36	-70	48
			4.36	-30	-64	54
			4.14	-30	-76	46

First level models produced a beta map for each item, for each participant. For each participant, 3 models, predicting by-stimulus estimates were created, with fact-like, moral-like, and preference-like ratings as respective predictors. Beta maps for ratings from each model were entered into a random effects analysis across all participants. Permutation tests (5000 samples) were used to achieve a cluster-corrected familywise error rate of $\alpha = .05$ in each contrast, while thresholding voxels at $p < .001$. Permutation testing was performed using SnPM 13 (http://warwick.ac.uk/snpm; Nichols & Holmes, 2002). All coordinates reported in MNI space.

Table S9. Study 2 ToM–behavioral analysis: controlling for semantic/syntactic features

Model:
PSC ~ Category*ROI*NounConcreteness + ROI*LeftEmbeddedness + NounFamiliarity +
(1+ Moral + Preference + VMPFC + PC + RTPJ + LTPJ + Moral*(VMFPC+LTPJ) | ID) +
(1 + VMPFC | Item)

REML criterion at convergence: 2378.2

Dummy coded control conditions: Facts (category) & DMPFC (ROI)

DV: Metaethical judgments (Study 1 BLUPs)
PSC corrected for syntactic/semantic features

Term		*F statistic*	*p*
PSC x 2 (ROI: ToM/VMPFC) x 3 (rating-type (fact-/moral-/preference-like) x 3 (category: fact/moral/preference)		$F(4, 405) = 0.86$	.486
PSC averaged across ROI			
PSC x 3 rating-type x category		$F(4, 198) = 9.54$	<.001 ***
PSC x 3 rating-type x category (fact/preference)		$F(2, 132) = 0.61$	.542
PSC x 3 rating-type x category (fact/moral)		$F(2, 132) = 9.81$	<.001 **
PSC x 3 rating-type x category (moral/preference)		$F(2, 132) = 14.17$	<.001 **
Within moral statements			
PSC x 3 rating-type		$F(2, 66) = 9.89$	<.001 ***
PSC x 2 (fact-like/preference-like)		$F(1, 44) = 18.20$	<.001 ***
PSC x 2 (fact-like/moral-like)		$F(1, 44) = 0.001$	.972
PSC x 2 (moral-like/preference-like)		$F(1, 44) = 13.97$	< .001 ***
Model: rating-type + PSC + (PSC x preference-like)	***B (SE)***	***t statistic***	**p**
Intercept (Fact-like rating)	2.18 (0.14)	$t(67) = 15.31$	< .001 ***
Moral-like	3.18 (0.19)	$t(67) = 16.50$	< .001 ***
Preference-like	2.61 (0.21)	$t(67) = 12.72$	< .001 ***
PSC (within fact-like/moral-like)	-0.97 (0.26)	$t(67) = 3.79$	< .001 ***
PSC x preference-like	1.99 (0.44)	$t(67) = 4.48$	< .001 ***
Contrasts:	***B (SE)***	***t statistic***	**p**
Fact-/moral-like–ToM relationship	-0.97 (0.26)	$t(70) = 3.79$	< .001 ***
Preference-like–ToM relationship	1.02 (0.36)	$t(70) = 2.81$	.013 *

Contrast *p* values corrected for 2 comparisons; single-step method; $\alpha_{familywise} = .05$; single-step method. *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .10$

Table S10. Model simplification for item features.

DV: Agreement (Study 1 BLUPs)			
Term		***F statistic***	***p***
PSC x 3 (category: fact/moral/preference) x 2 (ROI: ToM/VMPFC)		$F(2, 135) = 1.27$	.284
PSC x 2 (ROI)		$F(1, 137) = 0.30$	.587
PSC averaged across ROI			
PSC x 3 (category)		$F(2, 66) = 5.92$	.004 **
PSC x 2 (category: moral/fact)		$F(1, 44) = 3.71$	.061 †
PSC x 2 (category: moral/preference)		$F(1, 44) = 11.78$	.001 **
PSC x 2 (category: fact/preference)		$F(1, 44) = 2.32$	.135
Model: PSC x 3 (category)	***B (SE)***	***t statistic***	**p**
Contrast: PSC *within Facts*	0.02 (0.50)	$t(69) = 0.04$	1.00
Contrast: PSC *within Morals*	-1.46 (0.59)	$t(69) = 2.59$	.037 *
Contrast: PSC *within Preferences*	1.01 (0.44)	$t(69) =2.30$	.072 †
DV: Mental State General (Online Sample BLUPs)			
Term		***F statistic***	***p***
PSC x 3 (category: fact/moral/preference) x 2 (ROI: ToM/VMPFC)		$F(2, 135) = 1.36$	.261
PSC x 2 (ROI)		$F(2, 137) = 0.04$	.844
PSC averaged across ROI			
PSC x 3 (category)		$F(2, 66) = 5.92$	.004 **
PSC x 2 (category: moral/fact)		$F(1, 44) – 5.07$	.029 *
PSC x 2 (category: moral/preference)		$F(1, 44) = 12.33$	.001 **
PSC x 2 (category: fact/preference)		$F(1, 44) = 1.06$	.309
Model: PSC x 3 (category)	***B (SE)***	***t statistic***	**p**
Contrast: PSC *within Facts*	0.20 (0.12)	$t(66) = 1.65$	.275
Contrast: PSC *within Morals*	-0.24 (0.14)	$t(66) = 1.70$	.253
Contrast: PSC *within Preferences*	0.37 (0.11)	$t(66) = 3.39$	004 **
DV: Mental State Self-Oriented (Online Sample BLUPs)			
Term		***F statistic***	***p***
PSC x 3 (category: fact/moral/preference) x 2 (ROI: ToM/VMPFC)		$F(2, 135) = 0.43$	.689
PSC x 2 (ROI)		$F(1, 137) = 0.12$	.720
PSC averaged across ROI			
PSC x 3 (category)		$F(2, 66) = 3.96$	.024 *
PSC x 2 (category: moral/fact)		$F(1, 44) = 0.44$	.511
PSC x 2 (category: moral/preference)		$F(1, 44) = 6.96$	.011 *
PSC x 2 (category: fact/preference)		$F(1, 44) = 4.16$	.047 *
Model: PSC x 3 (category)	***B (SE)***	***t statistic***	**p**
Contrast: PSC *within Facts*	-0.12 (0.23)	$t(66) = 0.55$	.928
Contrast: PSC *within Morals*	-0.37 (0.27)	$t(66) = 1.37$	.436
Contrast: PSC *within Preferences*	0.51 (0.21)	$t(66) = 2.48$	.047 *
DV: Mental State Other-Oriented (Online Sample BLUPs)			
Term		***F statistic***	***p***
PSC x 3 (category: fact/moral/preference) x 2 (ROI: ToM/VMPFC)		$F(2, 135) = 0.52$	.596
PSC x 2 (ROI)		$F(1, 137) = 0.09$	.763
PSC averaged across ROI			
PSC x 3 (category)		$F(2, 66) = 3.48$	.037 *
PSC x 2 (category: moral/fact)		$F(1, 44) = 2.34$	.133
PSC x 2 (category: moral/preference)		$F(1, 44) = 9.30$	.004 **
PSC x 2 (category: fact/preference)		$F(1, 44) = 0.83$	.367
Model: PSC x 3 (category)	***B (SE)***	***t statistic***	**p**
Contrast: PSC *within Facts*	0.23 (0.19)	$t(66) = 1.20$	.547
Contrast: PSC *within Morals*	-0.26 (0.22)	$t(66) = 1.19$	.552
Contrast: PSC *within Preferences*	0.47 (0.17)	$t(66) = 2.77$	.022 *
DV: Person Present (Online Sample BLUPs)			
Term		***F statistic***	***p***
PSC x 3 (category: fact/moral/preference) x 2 (ROI: ToM/VMPFC)		$F(2, 135) = 0.61$	.546
PSC x 2 (ROI)		$F(1, 137) = 0.22$	.637
PSC averaged across ROI			
PSC x 3 (category)		$F(2, 66) = .04$	.962
Model: DV ~ category + PSC	***B (SE)***	***t statistic***	**p**

Main effect: PSC (*within* moral/preference/fact)	1.49 (0.65)	$t(68) = 2.28$	.026 *
DV: Mental Imagery (Online Sample BLUPs)			
Term		***F statistic***	***p***
PSC x 3 (category: fact/moral/preference) x 2 (ROI: ToM/VMPFC)		$F(2, 135) = 0.31$	.732
PSC x 2 (ROI)		$F(1, 137) = 0.07$	.797
PSC averaged across ROI			
PSC x 3 (category)		$F(2, 66) = 0.39$	.680
Model: DV ~ category + PSC	***B (SE)***	***t statistic***	**p**
Main effect: PSC (*within* moral/preference/fact)	-0.13 (0.11)	$t(68) = 1.15$	.253
DV: Arousal (Online Sample BLUPs)			
Term		***F statistic***	***p***
PSC x 3 (category: fact/moral/preference) x 2 (ROI: ToM/VMPFC)		$F(2, 135) = 0.11$	.892
PSC x 2 (ROI)		$F(1, 137) = 0.26$	.611
PSC averaged across ROI			
PSC x 3 (category)		$F(2, 66) = 0.35$	.703
Model: DV ~ category + PSC	***B (SE)***	***t statistic***	**p**
Main effect: PSC (*within* moral/preference/fact)	-0.05 (0.06)	$t(68) = 0.77$	.443
DV: Valence (Online Sample BLUPs)			
Term	***F statistic***		***p***
PSC x 3 (category: fact/moral/preference) x 2 (ROI: ToM/VMPFC)		$F(2, 135) = 0.13$	.868
PSC x 2 (ROI)		$F(1, 137) = 0.03$	.857
PSC averaged across ROI			
PSC x 3 (category)		$F(2, 66) = 0.03$	.857
Model: DV ~ category + PSC	***B (SE)***	***t statistic***	**p**
Main effect: PSC (*within* moral/preference/fact)	0.51 (0.47)	$t(68) = 1.09$	.281

*** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .10$

References

Coltheart, M. (1981). The MRC Psycholinguistic Database. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 497–505. https://doi.org/10.1080/14640748108400805

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). FMRI item analysis in a theory of mind task. *NeuroImage*, *55*(2), 705–712. https://doi.org/10.1016/j.neuroimage.2010.12.040

Fellbaum, C. (1998). *Wordnet: An electronic lexical database.* MIT Press.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 193–202. https://doi.org/10.3758/BF03195564

Kron, A., Goldstein, A., Lee, D. H.-J., Gardhouse, K., & Anderson, A. K. (2013). How Are You Feeling? Revisiting the Quantification of Emotional Qualia. *Psychological Science*, *24*(8), 1503–1511. https://doi.org/10.1177/0956797613475456

McNamara, D. S., Louwerse, M. M., Cai, Z., & Graesser, A. (2014). *Coh-Metrix version 3.0.* http://cohmetrix.com.

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An On-line Lexical Database*. *International Journal of Lexicography*, *3*(4), 235–244. https://doi.org/10.1093/ijl/3.4.235

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, *15*(1), 1–25. https://doi.org/10.1002/hbm.1058