Supplementary Material

Table S1. Aggression, violence, and dominance (**MA:** meta-analysis; **S**: survey).

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
anger	**S:** Brebner (2003)	students from 41 countries: frequency intensity Australian students: frequency intensity	6868 2199 		 –0.08[a] –0.20[a] –0.09[a] –0.00[a]	 –0.12, –0.03 –0.25, –0.15 –0.17, –0.001 –0.06, 0.04
	MA: Archer (2004)	adults: self-reports	35558[b]	46	–0.003[*]	–0.03, 0.02
	MA: Else-Quest *et al*. (2006)	3 months to 13 years	3984	24	0.04	–0.04, 0.11
	MA: Chaplin & Aldao (2013)	infancy to adolescence (mostly preschool ages)	NA	77	0.10	0.03, 0.16
indirect	**MA:** Knight *et al*. (2002)	adults: various methods	NA	8	–0.07	–0.09, –0.07
	MA: Archer (2004)	adults: self-reports children: peer-reports	30920[b] 10148[b]	40 26	–0.02[*] –0.10	–0.07, 0.02 –0.14, –0.06
	MA: Card *et al*. (2008)	ages up to 18 years: various methods		107	–0.06	–0.11, –0.02
	S: Lansford *et al*. (2012)	ages 7–10 years (9 countries)	1410		0.08[d]	–0.02, 0.18[c]
verbal	**MA:** Hyde (1984, 1986)	adults: various methods	NA	6	0.43	NA
	MA: Eagly & Steffen (1986)	adults: experimental methods	NA	20	0.18	0.10, 0.25
	MA: Bettencourt & Miller (1996)	adults: experimental methods (1) neutral conditions (2) provocation	 NA NA	 13 20	 0.30 0.05	 0.12, 0.48 –0.08, 0.18
	MA: Knight *et al*. (1996)[e]	adults: various methods	NA	6	0.46	0.31, 0.61
	MA: Bettencourt & Kernahan (1997)	adults: experimental methods: (1) neutral conditions (2) violent cues, provocation	 31 71	 2 3	 0.39 –0.27	 –0.57, 1.35[c] –0.85, 0.31[c]
	MA: Knight *et al*. (2002)	adults: various methods	NA	22	0.28	0.26, 0.30
	MA: Archer (2004)	adults: self-reports children: observations children: peer reports children: teacher reports	52564[b] 1624[b] 5460[b] 4103[b]	68 29 14 11	0.30[*] 0.14 0.51 0.24	0.27, 0.33 0.02, 0.26 0.45, 0.56 0.13, 0.34
	MA: Card *et al*. (2008)	various methods: ages up to 18 years	NA	27	0.38	NA
physical	**MA:** Hyde (1984, 1986)	adults: various methods	NA	26	0.60	NA
	MA: Eagly & Steffen (1986)	adults: experimental methods	NA	30	0.40	0.33, 0.47

Significance of sex differences

Table S1 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
physical (contd)	**MA:** Bettencourt & Miller (1996)	adults: experimental methods (1) neutral conditions (2) provocation	 NA NA	 35 26	 0.36 0.30	 0.25, 0.47 0.17, 0.43
	MA: Knight *et al*. (1996)[e]	adults: various methods	NA	38	0.91	0.86, 0.96
	MA: Bettencourt & Kernahan (1997)	adults: experimental methods (1) neutral conditions (2) violent cues/ provocation	 198 218	 6 7	 0.43 0.32	 0.39. 0.47[c] 0.02, 0.62[c]
	S: Brener *et al*. (1999)	ages 15–18 years (YRBS): physical fight (over 4 years)	55734		0.41[f]	0.36, 0.47
	MA: Knight *et al*. (2002)	adults: various methods	NA	41	0.59	0.56, 0.61
	S: Nansel *et al*. (2003)	school age (HBSC): physical fights, over 4 in last year	15686		0.47[g]	0.41, 0.54
	MA: Archer (2004)	adults: self-reports children: observations children: peer reports children: teacher reports	85803[b] 2408[b] 8190[b] 4103[b]	63 43 21 11	0.59[h*] 0.53 0.84 0.40	0.56, 0.62 0.43, 0.62 0.80, 0.89 0.36, 0.45
	MA: Card *et al*. (2008)	various methods: up to 18 years	NA	27	0.73	NA
	S: Cross (2010)	online student sample, UK	3775		0.58	0.52, 0.65
	S: Lansford *et al*. (2012)	ages 7–10 years, 9 countries	1410		0.22[d]	0.12, 0.33[c]
weapon-carrying	**S:** Brener *et al*. (1999)	ages 15–18 years (YRBS): over 4 years	55734		0.91[f]	0.88, 0.94
	S: Durant (1999)	ages 11–15 years (YRBS): North Carolina schools	2227		0.71[i]	0.57, 0.84
	S: Nansel *et al*. (2003)	school age: HBSC	15686		0.77[g]	0.71, 0.82
weapon use	**S:** Singer & Flannery (2000)	mean age 16 years, Ohio/ Colorado: shot at person	3692		0.63[j]	-0.17, 1.43
	S: Marcus (2009)	NLSAH-W-III: 16–21years	14098		0.88[k*]	0.82, 0.95
violent crime	**S:** Yao *et al.* (2014)	adults (population data, official Swedish registers	4849478		1.11[g*]	1.10, 1.12 [c]
homicide	**S:** Daly & Wilson (1990)	Chicago 1965–1976: same-sex homicides	NA		2.91[l]	NA
	S: Fox & Zawitz (2012)	US, 1976–2004 (FBI): same-sex homicides	NA		2.54[l*]	NA
cyber-bullying	**MA:** Barlett & Coyne (2014)	ages 7–24 years	NA	122	0.08	0.07, 0.08
violent computer-game use	**S:** Exelmans *et al*. (2015)	ages 12–18 years, 129 Flemish schools	3372		1.41[m*]	1.33, 1.49
forgiveness	**MA:** Miller *et al*. (2008)	adults	15731	70	–0.28[*]	–0.36, –0.21

Table S1 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
revenge	**MA:** Miller *et al*. (2008)	adults	1453	6	0.83*	0.43, 1.24
social dominance orientation	**MA:** Lee *et al*. (2011)	adults: overall US studies non-US studies	52826 26255[n] 26572[n]	169 84 95	0.43* 0.51 0.36	0.39, 0.47 0.46, 0.57 0.30, 0.41
	S: Ho *et al*. (2015)	adults (US)	3107		0.30[o]	0.22, 0.37
competitive-ness	**S:** Ahlgren & Johnson (1979)	ages 7–18 years (US)	2130		0.10[p]	0.01, 0.18
	MA: Walters *et al*. (1998)	in negotiations, adults, US & Canada	NA	79	0.07*	0.02, 0.13

Note. *N* indicates total number of participants; *k* indicates number of samples in a meta-analysis; *d* indicates Cohen's *d*, which is positive if in the male direction and negative if in the female direction; *CI* indicates 95% confidence intervals.
* indicates studies that were selected for inclusion in Table 3. See footnote to Table 3 for selection criteria.
Comprehensive Meta-Analysis (CMA) was used for calculating effect sizes where these are not taken from the sources.

Abbreviations: FBI = US Federal Bureau of Investigation; HBSC = Health Behavior in School-aged Children, 1998 (World Health Organization survey); NA = not available; NLSAH-W-III = National Longitudinal Study of Adolescent Health Wave III; YRBS = Youth Risk Behavior Survey.

[a] Calculated from means, standard deviations and sample sizes in Table 1 of source (note that the values in the present table are lower than the ones in the source).
[b] Approximate values obtained by multiplying mean number by the number of studies.
[c] Standard error calculated from sample sizes and *d* values using formula from Hedges & Olkin (1985, p. 86); *CIs* calculated from standard error using equations from Borenstein *et al*. (2009, p. 52).
[d] Values supplied by the first author for each nation separately ranged from –0.12 (Sweden) to 0.22 (China and Thailand) for indirect aggression, and 0.03 (Kenya) to 0.64 (Jordan) for physical aggression.
[e] Reanalysis and extension of data in Hyde (1984, 1986).
[f] Calculated from proportions in Table 5 of source. The value for *N* is the total over the 4 years and the *d* values are the mean weighted *d*s for the 4 years of the study (random effects model).
[g] Calculated from proportions in Table 1 of source.
[h] Value with outliers removed; value for the whole sample, $k = 111$, was $d = 0.39$.
[i] Calculated from data in Tables 1 and 2 of source. The value presented here is derived from the composite of (1) gun-carrying and (2) carrying another weapon.
[j] Calculated from proportions in Table 2 of source: this is the mean value from a combination of (1) attack someone with a knife and (2) shot at someone.
[k] Calculated from proportions in text of source. The value presented here is the mean of four measures: (1) use of weapon in a physical fight; (2) gang fights with weapon; (3) pulling a knife or gun on someone; (4) shot or stabbed someone.
[l] Percentage of chance encounters in which male would be the killer in a representative sample of same-sex homicide perpetrators, converted into *d* values from Table 1 in Grissom (1994).
[m] Calculated from the means and standard deviations in the text of the source.
[n] Calculated from figures in Table 1 of source.
[o] Calculated from *t* values in Table 10 of source.
[p] Calculated from the one-way *F* value, obtained from Table 1 of source (i.e. recalculated using the data in this table).

Table S2. Sensation-seeking, risk-taking and impulsivity (**MA:** meta-analysis; **S**: survey).

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
reward sensitivity	**MA:** Miettunen *et al.* (2007)	TPQ-TCI reward dependence	21092	32	–0.63[a*]	–0.78, –0.49
	MA: Cross *et al.* (2011)	overall	5978	18	0.01[*]	–0.17, 0.19
		SPSRQ & GRAPES	3534	9	0.44	0.36, 0.53
		BAS reward subscale	2573	9	–0.27	–0.41, –0.13
		TPQ-TCI reward dependence	1278	4	–0.56	–0.68, –0.44
punishment sensitivity	**MA:** Miettunen *et al.* (2007)	TPQ-TCI harm avoidance	21092	32	–0.33[*]	–0.41, –0.24
	MA: Cross *et al.* (2011)	overall	6689	18	–0.32	–0.45, –0.19
		TPQ-TCI harm avoidance	2175	5	–0.43	–0.52, –0.33
		BIS	2223	8	–0.63	–0.74, –0.52
		harm avoidance	862	3	–0.78	–0.92, –0.64
sensation-seeking	**MA:** Cross *et al.* (2011)	overall	52316	130	0.39[*]	0.35, 0.43
		SSS	5635	22	0.48	0.41, 0.56
		UPPS sensation-seeking	3850	15	0.48	0.33, 0.63
	MA: Cross *et al.* (2013)	SSS overall	16936	67	0.46	0.41, 0.51
excitement-seeking[b]	**S:** Costa *et al.* (2001)	adults (3 samples)	22642	3	0.29[b*]	0.14, 0.44
risk-taking	**MA:** Byrnes *et al.* (1999)	driving, self-report	NA	21	0.29	0.26, 0.32
		driving, observations	NA	14	0.17	0.12, 0.22
		gambling	NA	33	0.21	0.14, 0.28
		physical (skills)	NA	7	0.43	0.38, 0.48
	MA: Cross *et al.* (2011)	risk-taking	7069	11	0.36	0.29, 0.44
		disinhibition	5293	15	0.52	0.40, 0.65
		thrill and adventure-seeking	6259	16	0.41	0.29, 0.54
		EVS	17996	49	0.49[*]	0.43, 0.56
impulsivity	**MA:** Else-Quest *et al.* (2006)	general: from 3 months to 13 years	2254	21	0.18	0.04, 0.33
	MA: Cross *et al.* (2011)	general	113233	206	0.07[*]	0.05, 0.10
		delayed discounting	1787	21	–0.08	–0.19, 0.02
		BART	576	10	0.30	0.11, 0.49
delay of gratification	**MA:** Silverman (2003*a*)	preschool to adulthood	5640	38	–0.12[c*]	–0.27, 0.03
resistance to temptation	**MA:** Silverman (2003*b*)	preschool to adulthood	NA	114	–0.06[c*]	–0.10, –0.02
effortful control	**MA:** Else–Quest *et al.* (2006)	from 3 months to 13 years	792	6	–1.01[*]	–1.37, –0.64
inhibitory control	**MA:** Else-Quest *et al.* (2006)	from 3 months to 13 years	2876	22	–0.41[*]	–0.61, –0.21
risky impulsivity	**S:** Cross (2010)	online student sample, UK (RIS)	3775		0.34[*]	0.28, 0.41[d]

Note. *N* indicates total number of participants; *k* indicates number of samples in a meta-analysis; *d* indicates Cohen's *d*, which is positive if in the male direction and negative if in the female direction; *CI* indicates 95% confidence intervals.

Table S2 contd.

* indicates studies that were selected for inclusion in Table 3. See footnote to Table 3 for selection criteria. Comprehensive Meta-Analysis (CMA) was used for calculating effect sizes where these are not taken from the sources.

Abbreviations: BART = Balloon Analogue Risk Test; BAS = Behavioral Activation System Scale (Carver & White, 1994); BIS = Behavioral Inhibition System Scale (Carver & White, 1994); EVS = Eysenck Venturesome Scale (Eysenck *et al*., 1985); GRAPES = Generalized Reward and Punishment Expectancy Scales (Ball & Zuckerman, 1990); NA = not available; RIS = Risky Impulsiveness Scale (Campbell & Muncer, 2009); SPSRQ = Sensitivity to Punishment and Sensitivity to Reward Questionnaire (Torrubia *et al*., 2001); SSS = Sensation Seeking Scale; TPQ-TCI = Tridimensional Personality Questionnaire – Temperament and Character Inventory (Cloninger, 1986); UPPS = Urgency, Premeditation, Perseverence and Sensation-seeking Questionnaire (Whiteside & Lynam, 2001).

[a] Described as a measure of sociability (Cloninger *et al*., 1994).

[b] Meta-analysis of sex differences for excitement-seeking (a facet of extraversion) for 3 samples (Costa *et al*., 2001), one of US students and two multi-national ones. The value presented is the mean, weighted by the reciprocal of the variance, using a random-effects model (CMA).

[c] Calculated from the author's *r* values and sample sizes using CMA.

[d] Calculated from standard error in source using equations from Borenstein *et al*. (2009, p. 52).

Table S3. Fearfulness (**MA:** meta-analysis; **S**: survey).

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
fear	**MA:** Peck (1999)	adults' reactions to fear-inducing stimuli in media	6486	94	–0.41*	–0.46, –0.36[a]
	S: Brebner (2003)	students from 41 countries: frequency intensity Australian students: frequency intensity	6868 2199 		 –0.25[b] –0.38[b] –0.21[b] –0.07[b]	 –0.30, –0.20 –0.43, –0.33 –0.29, –0.12 –0.15, 0.02
	MA: Else-Quest *et al*.(2006)	from 3 months to 13 years	4858	34	–0.12	–0.20, –0.05
	S: Burnham *et al.* (2013)	ages 7–17 years, schools in 2 US states	1033		–1.80[c]	NA
	MA: Chaplin & Aldao (2013)	infancy to adolescence, mostly preschool	NA	24	–0.10	–0.17, –0.03
fear in real-world situations	**S:** Campbell *et al*. (2016)	3 samples from UK and Romania: SFQ	869		–1.16*	–1.32, –1.01
pain perception	**MA:** Riley *et al*. (1998)	threshold tolerance	1696 41670	17 10	0.51* 1.17*	0.46, 0.56[d] 1.16, 1.18[d]

Note. *N* indicates total number of participants; *k* indicates number of samples in a meta-analysis; *d* indicates Cohen's *d*, which is positive if in the male direction and negative if in the female direction; *CI* indicates 95% confidence intervals. * indicates studies that were selected for inclusion in Table 3. See footnote to Table 3 for selection criteria. Comprehensive Meta-Analysis (CMA) was used for calculating effect sizes where these are not taken from the sources.

Abbreviation: SFQ = Situated Fear Questionnaire (Campbell *et al*., 2016).

[a] Calculated from standard error in source using equations from Borenstein *et al*. (2009, p. 52).
[b] Calculated from means, standard deviations and sample sizes in Table 1 of the source (note that the values in the present table are lower than the ones in the source).
[c] Calculated from percentage of cases correctly classified, using Table 1 in Coe (2002).
[d] Standard error calculated from sample sizes and *d* values using formula from Hedges & Olkin (1985, p. 86); *CIs* calculated from standard error using equations from Borenstein *et al*. (2009, p. 52).

Significance of sex differences

Table S4. Visuospatial and mathematical abilities (**MA:** meta-analysis; **S**: survey).

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
VISUOSPATIAL ABILITIES						
mental rotation	**MA:** Linn & Peterson (1985)	10–55 years	NA	29	0.73	0.50, 0.96
	MA: Druva-Roush & Wu (1989)	adults (US doctoral dissertations)	NA	43	0.43	NA
	MA: Voyer *et al*. (1995)	4–60 years over 18 years	NA NA	78 43	0.56 0.66[*]	0.32, 0.80[a] 0.52, 0.80[a]
	S: Peters *et al*. (2006)	mainly student samples (Canada, Germany & Japan)	3367		0.89	0.89, 0.89[a]
	S: Silverman *et al*. (2007)	adults: 40 nations (BBC internet survey)	244893		0.48	0.48, 0.48[b]
	S: Lippa *et al*. (2010)	adults: 53 nations (BBC internet survey)	255144		0.47	0.46, 0.48[b]
	MA: Voyer (2011)	adults: paper-and-pencil tests: overall short time limit no time limit	 NA NA NA	 36 7 23	 0.70 1.03 0.51	 0.65, 0.75[a] 0.91, 1.15[a] 0.44, 0.58[a]
	MA: Maeda & Yoon (2013)	adults (PSVT:R[c])	NA	70	0.57	0.50, 0.64[a]
visuospatial perception/ ability	**MA:** Hyde (1981)	adolescence to old age (US): visual-spatial ability rod-and-frame test	 NA NA	 13 20	 0.45 0.51	 NA NA
	MA: Linn & Peterson (1985)	school ages to over 18 years	NA	62	0.44	0.04, 0.84
	MA: Druva-Roush & Wu (1989)	adults (US doctoral dissertations)	NA	43	0.03	NA
	S: Hedges & Nowell (1995)	adolescents: 2 US national samples (1960–1992)	98494		0.19[d]	0.16, 0.22[a]
	MA: Voyer *et al*. (1995)	4–60 years over 18 years	NA NA	92 53	0.44 0.48[*]	0.06, 0.82[a] 0.10, 0.86[a]
	MA: Uttal *et al*. (2013)	adults	NA	79	0.29[e]	0.15, 0.43
spatial visualization	**MA:** Linn & Peterson (1985)	school ages to over 18 years	NA	81	0.13	-0.24, 0.50
	S: Feingold (1988)	DAT norms (US, 1947–1980)	193844	4	0.24	0.23, 0.25[b]
	MA: Druva-Roush & Wu (1989)	adults (US doctoral dissertations)	NA	43	0.17	NA
	MA: Voyer *et al.* (1995)	4–60 years over 18 years	NA NA	116 56	0.19 0.23[*]	-0.07, 0.45[a] 0.01, 0.46[a]
line angle judgment	**S:** Lippa *et al*. (2010)	adults: 53 nations (BBC internet survey)	255144		0.49[*]	0.48, 0.50[b]
object location memory	**MA:** Voyer *et al*. (2007)	over 18 years	NA	70	–0.29[f]	–0.34, –0.24[a]

Table S4 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
object location memory contd.	**S:** Silverman *et al*. (2007)	adults: 40 nations (BBC internet survey)	247516		–0.31*	–0.32, –0.30[b]
MATHEMATICAL ABILITIES (INCLUDING MECHANICAL ABILITIES AND SCIENCE)						
quantitative ability	**MA:** Hyde (1981)	adolescence to older adults (US)	NA	16	0.43	NA
	S: Feingold (1988)	DAT norms (US, 1947–1980)	193844		0.05	0.04, 0.06[b]
	S: Feingold (1992*b*)	2 US national samples (16–74 years): WAIS, 1955; WAIS-R, 1981 (arithmetic) 2 US national samples (6–14 years): CAT-2 (arithmetic)	3580 140000		0.34 0.00	0.27, 0.41[b] –0.01, 0.01
	S: Strand *et al*. (2006)	UK representative sample (11–12 years): CAT-3	320000		0.03*	0.03, 0.04[b]
mathematics	**MA:** Freeman (1985)	US school and university,	48648	35	0.14	0.12, 0.16[a]
	MA: Hyde *et al*. (1990)	over 5 years of age	3175188	254	0.15	0.15, 0.15[b]
	S: Feingold (1992*b*)	US college entrance: PSAT (1960–1983) SAT-M (1967)	 99654 NA		 0.31 0.42	 0.30, 0.32[b] NA
	S: Hedges & Nowell (1995)	adolescents (5 US national samples: 1960–1992) 17 years (US, NAEP, 1971–1992	151867 70000-100000 per year		0.16 0.16[d]	0.11, 0.21[a] 0.07, 0.25[a]
	S: Hyde *et al*. (2008)	7–17 years (school tests, 10 US states)	over 7000000		0.01	0.01, 0.01[b]
	MA: Else-Quest *et al*. (2010)	14–16 years: TIMSS 2003 (46 nations) PISA 2003 (46 nations)	 219612 273883	 46	 –0.01 0.11	 –0.05, 0.03 0.09, 0.13
	MA: Lindberg *et al*. (2010)	school, college, and adults worldwide samples	1286350	242	0.05	–0.05, 0.05[b]
	S: Lindberg *et al*. (2010)	13–18 years (US): NLSY-97 13–14 years (US): NELS-88 12–17 years (US): LSAY 9–17 years (US): NAEP	6044 23648 3065 various		0.08 0.10 -0.07 0.06[g]	NA NA NA NA
	S: Ball *et al*. (2013)	annual SAT-M data (US) 1996–2009	19000000		0.31[d]	0.31, 0.31[b]
	S: Stoet & Geary (2013)	15 years: 75-nations, PISA: 2000, 2003, 2006, 2009)[h]	1500000		0.09[d*]	–0.02, 0.19

Table S4 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
mathematics (contd.)	**MA:** Voyer & Voyer (2014)	marks from elementary school to university (70% US)	NA	93	–0.07	–0.12, –0.01
	MA: Reilly *et al*. (2015)	17 years (US): NAEP	104900	6	0.10	0.08, 0.12
	S: Pargulski & Reynolds (2017)	4 to 17 years (US): WAIT–III	2580		0.14	0.06, 0.22
mechanical reasoning	**S:** Feingold (1988)	4 DAT norms (US, 1947–1980)	193844		0.98*	0.97, 0.98[b]
	S: Hedges & Nowell (1995)	adolescents: 2 US national samples (1960–1992)	85339		0.78[d]	0.70, 0.86[b]
science	**S:** Hedges & Nowell (1995)	adolescents: 3 US national samples (1960–1992)	109938		0.33[d]	0.29, 0.37[b]
	S: Hedges & Nowell (1995)	17 years: US national samples (NAEP: 1971–1992)	70000 to 100000 per year		0.28[d*]	0.19, 0.37[b]
	MA: Voyer & Voyer (2014)	marks from elementary school to university (70% US)	NA	31	–0.15	–0.23, –0.08
	MA: Reilly *et al*. (2015)	17 years: US national samples (NAEP)	56437	6	0.13	0.09, 0.18

Note. *N* = total number of participants; *k* = number of samples in a meta-analysis; *d* indicates Cohen's *d*, which is positive if in the male direction and negative if in the female direction; *CI* indicates 95% confidence intervals. * indicates studies that were selected for inclusion in Table 3. See footnote to Table 3 for selection criteria. Comprehensive Meta-Analysis was used for calculating effect sizes where these are not taken from the sources.

Abbreviations: BBC = British Broadcasting Corporation; CAT-3 = Cognitive Abilities Test; CAT-2 = California Achievement Tests; DAT = Differential Aptitude Test; LSAY = The longitudinal Study of American Youth; NA = not available; NAEP = National Assessment of Education (annual national probability samples); NELS-88 = National Education Longitudinal Study of 1988; NLSY-97 = National Longitudinal Surveys of Youth (data for initial year, 1997); PISA = Programme for International Student Assessment; PSAT = Preliminary Scholastic Aptitude Test; PSVT:R = Purdue Spatial Visualization Tests: Visualization of Rotations; SAT-M = Scholastic Aptitude Test-math; TIMSS = Trends in International Mathematics and Science Study; WAIS = Wechsler Adult Intelligence Scale; WAIS-R = Wechsler Adult Intelligence Scale-Revised; WAIT-III = Wechsler Individual Achievement Test – Third Edition.

[a] Calculated from standard error or standard deviation or *Z* score in source using equations from Borenstein *et al*. (2009, p. 52).

[b] Standard error calculated from sample sizes and *d* values using formula from Hedges & Olkin (1985, p. 86); *CIs* calculated from standard error using equations from Borenstein *et al*. (2009, p. 52).

[c] This test was not included in the previous two meta-analyses involving mental rotation (Linn & Petersen, 1985; Voyer *et al*., 1995).

[d] Means of annual values shown in Table 2 or 3 of Hedges & Nowell (1995) or Table S2 in Stoet & Geary (2013) or Table 2 of Ball *et al*. (2013).

[e] The mean value for control and treatment conditions and pre- and post-training in studies involving training of spatial skills, taken from Table 5: treatment effects were very similar for the two sexes.

[f] Values calculated from data in Tables 2, 3 and 4 of source.

Table S4 contd.

[g] Mean weighted *d* value across all 18 of the main assessments (from source).

[h] There are other analyses of parts of these PISA data: for 2009, covering 65 nations, Reilly (2012) gives values of *d* = –0.44 for reading and *d* = 0.22 for maths, both higher than the values for this year from Stoet & Geary (2013) for 75 nations.

Table S5. Object-centred orientation (**MA:** meta-analysis; **S**: survey).

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
systemizing (SQ)	**S:** Manning *et al*. (2010)	adults: BBC internet survey (53 nations)	170164		1.21*	1.20, 1.22[a]
	S: Svedholm-Häkkinen & Lindeman (2016)	Finnish adults	3084		1.04	0.96, 1.11
occupational interests[b]	**S:** Lippa (2008, 2010)	adults: BBC internet survey (53 nations)	200000		1.39*	1.38, 1.40[a]
engineering interests	**MA:** Su *et al*. (2009)	ages 12–42 years, interest inventories (US)	503188	45	1.11*	1.01, 1.20
interest in "things"	**MA:** Woodcock *et al*. (2013)	Thing Orientation[c]	7450	15	0.97[d]*	0.80, 1.15

Note. *N* indicates total number of participants; *k* indicates number of samples in a meta-analysis; *d* indicates Cohen's *d*, which is positive if in the male direction and negative if in the female direction; *CI* indicates 95% confidence intervals. * indicates studies that were selected for inclusion in Table 3. See footnote to Table 3 for selection criteria. Comprehensive Meta-Analysis (CMA) was used for calculating effect sizes where these are not taken from the sources.

Abbreviations: BBC = British Broadcasting Corporation; SQ = Systemizing Quotient (Ling *et al*., 2009).

[a] Standard error calculated from sample sizes and *d* values using formula from Hedges & Olkin (1985, p86); *CIs* calculated from standard error using equations from Borenstein *et al*. (2009, p. 52).
[b] Interest in car mechanic, builder, carpenter, electrical engineer, and inventor, compared with interest in costume designer, dance-teacher, school-teacher, florist, and social worker.
[c] A measure of Person-Thing Orientation that has separate items for Person and Thing Orientation (Graziano *et al*., 2011).
[d] Calculated from the means, standard deviations, and sample sizes for the 15 samples presented in Table 1 of the source, using CMA (random effects model): some of the resulting *d* values were discrepant from those shown in the source table. The samples were mainly from the US, with two small-sample studies from Greece and Turkey. Some are likely to have been selected, for example, university students studying specific subjects, such as psychology and engineering. The specific values for Person Orientation varied from –0.27 (Greek sample) to –0.82 (sixth-grade US students), and those for Thing Orientation from 0.50 (US third-grade students) to 1.34 (US Introductory Psychology students).

Table S6. Characteristics directly associated with social relations (**MA:** meta-analysis; **S**: survey).

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
SOCIAL SKILLS AND INTERESTS						
people–things dimension	**S:** Lippa (1998)	twins (high school)	1678		–1.27[a]	–1.43, –1.10
	MA: Su *et al*. (2009)	ages 12 to 42 years, interest inventories (US)	503188	79	–0.93[*]	–0.99, –0.87
	MA: Woodcock *et al*. (2013)	Person Orientation[b]	7450	15	–0.48[c]	–0.53, –0.44
social interests	**S:** Lippa (1998)	twins (high school)	1678		–1.14[a]	–1.35, –0.93
	MA: Su *et al*. (2009)	ages 12 to 42 years, interest inventories (US)	503188	80	–0.68[*]	–0.74, –0.62
emotional intelligence	**MA:** Joseph & Newman (2010)	adults: performance-based tasks	2216	14	–0.47[*]	–0.72, –0.24
face recognition	**S:** Herlitz *et al*. (1997)	adult Swedish sample	1000	43	–0.60[d]	–0.66, –0.53
	MA: Herlitz & Lovén (2013)	children and adults	NA		–0.36[*]	–0.44, –0.29
decoding non-verbal cues	**MA:** Hall (1978)	preschool to adults (US)	10244	46	–0.34	–0.52, –0.16
	MA: McClure (2000)	facial expression processing: infants children and adolescents	NA 23 80	 –0.18 –0.13	 –0.33, –0.03 –0.19, –0.07
	S: Sasson *et al*. (2010)	over 18 years, online sample: facial affect recognition	7320		–0.35[e]	–0.40,–0.30[f]
	MA: Thompson & Voyer (2014)	children & adults: emotion recognition	NA	404	–0.27[g*]	–0.32, –0.23
SOCIAL INTERACTIONS IN DYADS OR SMALL GROUPS						
peer attachment	**MA:** Gorrese & Ruggieri (2012)	adolescents and young adults, degree of attachment	21052		–0.51[*]	–0.59, –0.42
implicit affiliation motivation	**MA:** Drescher & Schultheiss (2016)	adults	5962	33	–0.45[*]	–0.53, –0.37
smiling	**MA:** La France *et al*. (2003)	adolescents and adults	109654	448	–0.41[*]	–0.42, –0.39
seek emotional social support	**MA:** Tamres *et al*. (2002)	adults, English-speaking nations	2171	12	–0.41[*]	–0.49, –0.32
disclosure	**MA:** Belk (1991)	adults (US) self-disclosure	NA	76	–0.27	–0.73, 0.19[h]
	MA: Dindia & Allen (1992)	to same sex to opposite sex overall	6264 7320 23702	66 66 205	–0.37[*] –0.13 –0.18	–0.42, –0.32 –0.09, –0.18, –0.21, –0.16

Table S6 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
disclosure (contd.)	**MA:** Lundquist (1993)	adults, self-disclosure	NA	25	–0.20	–0.27, –0.14
agreeableness	**MA:** J. Archer (unpublished data)	variety of adult samples[i]	572222	10	–0.29[i*]	–0.39, –0.19
friendship expectations	**MA:** Hall (2011)	adolescents and young adults: overall communion agency	 8825 8245 3470	 36 31 9	 –0.17[*] –0.39 0.34	 –0.22, –0.12 –0.44, –0.34 0.26, 0.42
intimacy in relationships	**MA:** Lundquist (1993)	adolescence to old age	46594	257	–0.11[*]	–0.13, –0.09[h]
touch initiation	**MA:** Stier &.Hall (1984)	overall	NA	6	–0.09[*]	NA
personal space	**MA:** Daigle (1996)	children and adults	NA	92	0.08[*]	0.04, 0.11
adult attachment style	**S:** Noftle & Shaver (2006)	college students, US: anxious avoidant	8318		 –0.02 0.00	 –0.06, 0.02 –0.04, 0.04
	MA: del Giudice (2011)	worldwide sample: anxious avoidant bivariate D for two styles	65047	112	 –0.04[*] 0.02[*] 0.05	 –0.07, –0.01 –0.01, 0.05 NA
	S: del Giudice (2016)	college students, US: self-reliance closeness discomfort preoccupation neediness reject desire closeness	8829		 0.15 –0.07 –0.13 –0.14 0.31	 0.11, 0.19[h] –0.11, –0.03[h] –0.17, –0.09[h] –0.18, –0.10[h] 0.27, 0.35[h]
EMPATHY						
empathy (various measures)	**MA:** Eisenberg & Lennon (1983)	US adults: questionnaires US children: picture/story measures US infants: reflexive crying	4085 1282 339	17 22 7	–0.91[*] –0.11 –0.34	–0.97, –0.85[h] –0.22, 0.00[h] –0.55, –0.13[h]
EQ	**S:** Manning *et al*. (2010)	adults, 53 nations, BBC internet survey	170227		–0.87[*]	–0.88, –0.86[h]
	S: Svedholm-Häkkinen & Lindeman (2016)	Finnish adults	3084		–0.59	–0.67, –0.51[h]
reading the mind in the eyes	**MA:** Kirkland *et al*. (2013)	adults	NA	40	–0.18*	–0.24, –0.12

Table S6 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
HELPING BEHAVIOUR/COOPERATION						
helping behaviour	**MA:** Eagly & Crowley (1986)	adults, US and Canada	37308	99	–0.34[j*]	–0.36, –0.32
cooperation	**S:** Ahlgren & Johnson (1979)	ages 7 to 18, US	2130		–0.10[k]	–0.19, –0.02
	MA: Balliet *et al.* (2011)	adults, mostly US: overall mixed-sex interactions same-sex interactions	 30000 NA NA	 272 90 58	 –0.05[*] –0.22 0.16	 –0.11, 0.001 –0.29, –0.15 0.06, 0.25
MORALITY						
moral norms *v.* consequences	**MA:** Friesdorf *et al*. (2015)	young adults	6100	40	–0.57[*]	–0.62, –0.52
moral self-esteem	**MA:** Gentile *et al.* (2009)	children and adults	NA	15	–0.38[*]	–0.48, –0.29
cheating	**MA:** Whitley *et al.* (1999)	US college students: attitudes behaviour	 6292 26262	 14 52	 0.35[l*] 0.17[l]	 0.32, 0.37 0.16, 0.18
moral orientation	**MA:** Jaffe & Hyde (2000)	children and adults: care orientation justice orientation	 12437 8138	 160 95	 –0.28[*] 0.19	 –0.32, –0.25[h] 0.15, 0.23[h]
moral sensitivity	**MA:** You *et al*. (2011)	adults	4408	20	–0.24[*]	–0.34, –0.14[f]
justice-based moral reasoning	**MA:** Thoma (1986)	children and young adults	6863	56	–0.21[*]	–0.26, –0.16
LEADERSHIP						
emergence of leaders in groups	**MA:** Eagly & Karau (1991)	US and Canada: task unspecified no task (social leadership)	 NA NA NA	 34 29 15	 0.41[*] 0.29 –0.18[*]	 0.34, 0.48 0.24, 0.34 –0.29, –0.06
leadership effectiveness	**MA:** Eagly *et al*. (1995)	US and Canada: overall self-ratings subordinates' ratings	 NA NA	 76 34 40	 –0.02 0.14 –0.12	 –0.05, 0.02 0.08. 0.20 –0.16, –0.07
	MA: Paustian-Underdahl *et al*. (2014)	self-ratings other ratings overall	4711 96893 101676	19 78 99	0.21 –0.12 –0.05	0.09, 0.31 –0.18, –0.06 –0.10, –0.00

Table S6 contd

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
leadership style	**MA:** Eagly & Johnson (1990)	adults, US and Canada: democratic vs autocratic style	NA	23	–0.22[m*]	–0.29, –0.15
	MA: Eagly *et al*. (2003)	adults: transformational[n]	NA	44	–0.10	–0.13, –0.08
		laissez-faire[n]	NA	16	0.16	0.14, 0.19
		transactional, contingent reward[n]	NA	21	–0.13	–0.17, –0.10
		active management by exception[n]	NA	12	0.12	0.08, 0.16
		passive management by exception[n]	NA	18	0.27	0.23, 0.30
	MA: Hallinger *et al*. (2016)	instructional (perceptions of school principals by teachers and principals	2000+	40	–0.30	–0.37, –0.21
negotiation outcome	**MA:** Stuhlmacher & Walters (1999)	adults, US and Canada	3496	53	0.09	0.02, 0.16
	MA: Mazei *et al*. (2015)	adults	10888	123	0.20	0.11, 0.28
influencing others	**MA:** Eagly & Carli (1981)	adults, US and Canada	12856	90	0.26[o*]	0.19, 0.33

Note. *N* = total number of participants; *k* = number of samples in a meta-analysis; *d* indicates Cohen's *d*, which is positive if in the male direction and negative if in the female direction; *CI* indicates 95% confidence intervals. * indicates studies that were selected for inclusion in Table 3. Comprehensive Meta-Analysis (CMA) was used for calculating effect sizes where these are not taken from the sources.

Abbreviations: BBC = British Broadcasting Corporation; EQ = Empathy Quotient (Baron-Cohen & Wheelwright, 2004; Muncer & Ling, 2006); NA = not available.

[a] Calculated from the two *r* values in Lippa (1998, p.1005), which were combined using CMA (random effects model).

[b] A measure of Person–Thing Orientation that has separate items for Person and Thing Orientation (Graziano *et al*., 2011).

[c] Calculated from the means, standard deviations, and sample sizes for the 15 samples presented in Table 1 of the source, using CMA (random effects model): (some of the resulting *d* values were discrepant from those shown in this table). The samples were mainly from the US, with two small-sample studies from Greece and Turkey. Some are likely to have been selected, for example, university students studying specific subjects, such as psychology and engineering. The specific values for Person Orientation varied from –0.27 (Greek sample) to –0.82 (sixth-grade US students), and those for Thing Orientation from 0.50 (US third-grade students) to 1.34 (US Introductory Psychology students).

[d] This was the value calculated from the means and standard deviations in Table 2 of the source. It differs considerably from the value in this table (*d* = –0.27).

[e] Calculated from the means and standard deviations in Table 6 of the source. This *d* value is lower than that in the text (*d* = –0.41), as this value excluded neutral faces (N.J. Sasson, personal communication, 14 January 2018).

[f] Calculated from standard error or *Z* score in source using equations from Borenstein *et al*. (2009, p. 52).

[g] Mean weighted *d* value for samples where it was possible to compute a *d* value: for all reports including where "no difference" was coded as zero, *d* = 0.17.

Table S6 contd.

[h] Standard error calculated from sample sizes and *d* values using formula from Hedges & Olkin (1985, p. 86); *CIs* calculated from standard error using equations from Borenstein *et al*. (2009, p. 52).
[i] Meta-analysis of 10 studies, one a meta-analysis (Guo *et al*., 1995), and nine large-sample studies, including multi-national ones (Costa *et al*., 2001, Studies 2 and 3: 26 nations; Lippa, 2008: 53 nations; Schmitt *et al*., 2008: 55-nations; Vianello *et al.*, 2013: 127 nations, although 80% participants from the US); other studies from the US (Costa *et al*., 2001, Study 1; Noftle & Shaver. 2006), Germany (Rammstedt, 2007), and English-speaking nations (Soto & John, 2017). Values presented here are mean values weighted by the reciprocal of the variance, random-effects model (Comprehensive Meta-Analysis). The values used for Costa *et al*. (2001) were the means for the six facets listed in their Table 2. See Table S13 for details of these and other sex differences in personality.
[j] Mean weighted *d* value for 99 samples where it was possible to compute an effect size; for all reports including those where it was only stated that there was "no difference" (k = 181), the mean was 0.07 (unweighted as variance values were unavailable).
[k] Calculated from the one-way *F* value, from Table 1 of the source (i.e. recalculated using the data in this table).
[l] A positive value indicates more positive attitudes to cheating, or more cheating behaviour, in males.
[m] *d* value for samples where it was possible to compute an effect size; for all reports including those where it was only stated that there was "no difference" (k = 28), the mean was –0.34.
[n] Leadership styles: (1) transformational = a role model who gains the trust and confidence of subordinates (this can be achieved in several ways); (2) laissez-faire = not taking responsibility for managing; (3) transactional, contingent reward = rewarding satisfactory performance; (4) transactional, active management by exception = attends to subordinates' errors to reach standards; (5) transactional, passive management by exception = intervenes only after subordinates' errors become severe.
[o] *d* value for 90 samples where it was possible to compute an effect size; for all reports including those where it was only stated that there was "no difference" (k = 148), the mean was –0.16 (neither *d*s were weighted means). The *N* was calculated from 90/148 of the total *N* for all effect sizes in Table 2 of source.

Table S7. Language and related attributes (**MA:** meta-analysis; **S**: survey)

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
overall verbal abilities	**MA:** Hyde (1981)	adolescence to older adults US)	NA	27	–0.27*	NA
	MA: Hyde & Linn (1988)	from 3 years of age	NA	119	–0.11	NA
	S: Feingold (1992*b*)	college entrance, US (PSAT) college entrance, US (PSAT)	99654 NA		–0.03 0.01	–0.04, –0.00[a] NA
language ability	**S:** Feingold (1988)	4 DAT norms: US, 1947–1980	193844		–0.43	–0.44, –0.43[a]
	S: Feingold (1992*b*)	2 US national samples, 6–14 years: CAT-2	140000		–0.39	–0.40, –0.38[a]
	MA: Voyer & Voyer (2014)	marks from elementary school to university (70% US)	NA	81	–0.37*	–0.43, –0.32
speech	**MA:** Hyde & Linn (1988)	from 3 years of age: speech production	NA	12	–0.33*	–0.46, –0.20
	MA: Leaper & Ayres (2007)	adults: talkativeness adults: affiliative speech adults: assertive speech	4385 2781 2541	70 54 50	0.14* –0.12* –0.09	0.08, 0.19 –0.18, –0.06 –0.15, –0.02
	MA: Leaper & Smith (2004)	children: talkativeness children: affiliative speech children: assertive speech	3303 2694 3495	73 46 75	–0.11 –0.26 –0.11	–0.17, –0.05 –0.33, –0.19 –0.17, –0.05
	MA: Leaper & Robnett (2011)	adults: tentative speech	3502	39	–0.23*	–0.32, –0.13
	MA: Anderson & Leaper (1998)	adults: interrupting adults intrusive interruptions	3058 3058	53 17	0.15* 0.33	0.07, 0.23[b] 0.17, 0.49[b]
vocabulary	**MA:** Hyde & Linn (1988)	from 3 years of age	NA	40	–0.02	–0.06, 0.02
	S: Feingold (1992*b*)	US national samples, 6–14 years: WAIS/ WAIS-R CAT-2	3580 140000		–0.04 –0.10*	–0.07, –0.01[a] –0.11, –0.09[a]
	S: Hedges & Nowell (1995)	adolescents (4 US national)	127268		0.06[c]	0.02, 0.10[b]
verbal reasoning	**S:** Feingold (1988)	US (1947–1980): 4 DAT norms	193844		0.05	0.04, 0.06[a]
	S: Strand *et al*. (2006)	UK, representative sample, 11–12 years, CAT-1	320000		–0.15*	–0.16, –0.14[a]
reading	**MA:** Hyde & Linn (1988)	from 3 years	NA	18	–0.03	–0.04, –0.01
	S: Feingold (1992*b*)	US national samples, 6-14 years: CAT-2	140000		–0.21	–0.32, –0.10[a]
	S: Hedges & Nowell (1995)	adolescents (5 US national) US, national probability (NAEP), 17 years	151867 70000–100000/ year		–0.09[c] –0.25[c]	–0.13, –0.05 –0.33, –0.17

Table S7 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
reading contd.	**S:** Stoet & Geary (2013)	75-nations, 15 years, PISA: 2000, 2003, 2006, 2009	1.5 million		–0.36[c*]	–0.48, –0.24
	S: Pargulski & Reynolds (2017)	US, 4–17 years, WAIT-III	2345		–0.04	–0.12, 0.05
writing	**MA:** Hyde & Linn (1988)	from 3 years	NA	5	–0.09	–0.10, –0.08
	S: Hedges & Nowell (1995)	17 years, US, national probability: NAEP: 1971–92	70000–100000/ year		–0.57[c*]	–0.72, –0.42[b]
	S: Pargulski & Reynolds (2017)	US, 4–17 years, WAIT-III: written expression	2580		–0.25	–0.33, –0.18
spelling	**S:** Feingold (1988)	US (1947-1980): 4 DAT norms	193844		–0.50[*]	–0.51, –0.49[a]
	S: Feingold (1992*b*)	US national samples, 6–14 years: CAT-2	140000		–0.42	–0.43, –0.41[a]

Note. *N* = total number of participants; *k* = number of samples in a meta-analysis; *d* indicates Cohen's *d*, which is positive if in the male direction and negative if in the female direction; *CI* indicates 95% confidence intervals. [*] indicates studies that were selected for inclusion in Table 3. Comprehensive Meta-Analysis (CMA) was used for calculating effect sizes where these are not taken from the sources.

Abbreviations: CAT-1 = Cognitive Abilities Test; CAT-2 = California Achievement Tests; DAT = Differential Aptitude Test; NA = not available; NAEP = National Assessment of Education (annual national probability samples); PSAT = Preliminary Scholastic Aptitude Test; PISA = Programme for International Student Assessment; SAT = Scholastic Aptitude Test; WAIS = Wechsler Adult Intelligence Scale; WAIS-R = Wechsler Adult Intelligence Scale-Revised; WAIT-III = Wechsler Individual Achievement Test – Third Edition.

[a] Standard error calculated from sample sizes and *d* values using formula from Hedges & Olkin (1985, p. 86); *CIs* calculated from standard error using equations from Borenstein *et al*. (2009, p. 52).
[b] Calculated from standard error or *Z* score in source using equations from Borenstein *et al*. (2009, p. 52).
[c] Means of annual values (Table 2 or 3: Hedges & Nowell, 1995; or Table S3: Stoet & Geary, 2013). Another analysis of the PISA data for 2009, covering 65 nations, gives a value of *d* = –0.44 for reading (Reilly, 2012) for 75 nations.

Significance of sex differences

Table S8. Depression and negative emotions (**MA:** meta-analysis; **S**: survey).

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
depressive symptoms	**S:** Meddin (1986)	US, adults: QAL, 1978	2114		–0.50[a]	–0.84, –0.17
	S: Kessler *et al*. (1994)	NCS, major depression: 12-month prevalence lifetime prevalence	8098		 –0.32[b] –0.34[b]	 –0.40, –0.23 –0.41, –0.28
	S: Gater *et al*. (1998)	WHO study, 15 nations, ICD-10, current status	5438		–0.35[bc]	–0.45, –0.24
	S: Nolen-Hoeksema *et al*. (1999)	California, 25 to 75 years	1132		–0.16[d]	–0.29, –0.04
	S: Daradkeh *et al*. (2002)	UAE community study: lifetime prevalence	1336		 –0.76[b]	 –1.05, –0.48
	MA: Twenge & Nolen-Hoeksema (2002)	US, school ages: 8 to 12 years 13 to 16 years	43916 NA NA	251 86 49	–0.02 0.04 –0.16	–0.04, 0.00[e] 0.01, 0.07 –0.20, –0.12
	MA: Wang *et al*. (2016)	Non-clinical samples (BDI)	53049	91	–0.19	–0.22, –0.16[e]
	MA: Salk *et al*. (2017)	representative national samples from 90 nations	1922064	180	–0.27[*]	–0.29,–0.26
Major Depression diagnosis	**MA:** Salk *et al*. (2017)	representative national samples from 90 nations	1716195	149	–0.34[f*]	–0.35, –0.34[f]
sadness	**S:** Brebner (2003)	students from 41 countries: frequency intensity Australian students: frequency intensity	6868 2199 		 –0.23[g*] –0.40[g] –0.24[g] –0.27[g]	 –0.28, –0.18 –0.45, –0.35 –0.32, –0.15 –0.36, –0.19
	MA: Chaplin & Aldao (2013)	infancy to adolescence mostly preschool ages	NA	69	–0.06	–0.12, 0.004
rumination	**S:** Nolen-Hoeksema *et al*. (1999)	California, 25 to 75 years	1132		–0.26[h]	–0.38, –0.13
	MA: Tamres *et al*. (2002)	adults, English-speaking nations	2014	10	–0.39[i]	–0.48, –0.30
	MA: Rood *et al*. (2009)	children adolescents	1580[j] 3188[j]	9[j] 10[j]	–0.14 –0.36	–0.15, –0.13[e] –0.44, –0.28[e]
	MA: Johnson & Whisman (2013)	adults	14321	59	–0.24[*]	–0.27, –0.21[e]
anxiety	**S:** Kessler *et al*. (1994)	NCS: 12-month prevalence NCS: lifetime prevalence	8098 8098		–0.43[b] –0.34[b]	–0.50, –0.36 –0.40, –0.28
	S: Lewinsohn *et al*. (1998)	age 16.6 years, OADP: presence/absence disorder anxiety symptom score	1221		 –0.47[b] –0.36[k]	 –0.67, –0.26 –0.47, –0.25
	MA: Abdel-Khalek & Alansari (2004)	students from 10 Arab countries: KUAS	3064	10	–0.35[l]	–0.49, –0.20

Table S8 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
anxiety contd.	**S:** Vesga-López *et al*. (2008)	US representative sample, NESARC 12-month prevalence lifetime prevalence	43093		 –0.59[b*] –0.37[b]	 –0.67, –0.51 –0.47, –0.27
	MA: Chaplin & Aldao (2013)	infancy to adolescence, mostly preschool ages	NA	33	–0.01	–0.09, 0.07
social anxiety	**S:** Lustig & Anderson (1990)	US students, 40 universities	2735		–0.16[k]	–0.23, –0.18
	MA: Lustig & Anderson (1990)	from previous reviews	NA	22	–0.19[m]	–0.27, –0.12
	S: Caballo *et al*. (2014)	multinational Spanish–speaking sample: SAQ-A	31112		–0.36[*]	–0.38, –0.33[e]
neuroticism	**MA:** J. Archer (unpublished data)	variety of adult samples[n]	572222	10	–0.31[n*]	–0.38, –0.24
negative emotions overall	**S:** Brebner (2003)	students from 41 countries: frequency intensity Australian students: frequency intensity	6868 2199 		 –0.21[g] –0.37[g] –0.19[g] –0.16[g]	 –0.26, –0.16 –0.41, –0.32 –0.27, –0.10 –0.25, –0.08
	MA: Chaplin & Aldao (2013)	infancy to adolescence, mostly preschool	NA	111	0.03[*]	–0.03, 0.08
shame	**MA:** Else-Quest *et al*. (2012)	mostly adults early adulthood	NA NA	232 116	–0.29[*] –0.34	–0.34, –0.24 NA
	MA: Chaplin & Aldao (2013)	infancy to adolescence, mostly preschool	NA	6	–0.56	–1.01, –0.11
guilt	**S:** Brebner (2003)	students from 41 countries: frequency intensity Australian students: frequency intensity	6868 2199 		 –0.03 –0.11 –0.04 –0.04	 –0.07, 0.02 –0.16, –0.06 –0.12, 0.05 –0.12, 0.05
	MA: Else-Quest *et al*. (2012)	mostly adults early adulthood	NA NA	307 143	–0.27[*] –0.32	–0.32, –0.23 NA

Note. *N* = total number of participants; *k* = number of samples in a meta-analysis; *d* indicates Cohen's *d*, which is positive if in the male direction and negative if in the female direction; *CI* indicates 95% confidence intervals. [*] indicates studies that were selected for inclusion in Table 3. Unless otherwise stated, Comprehensive Meta-Analysis (CMA) was used for calculating effect sizes where these are not taken from the sources.

Abbreviations: BDI = Beck Depression Inventory; ICD-10 = International Statistical Classification of Diseases, 10th revision; KUAS = Kuwait University Anxiety Scale (Abdel-Khalek, 2000); NA = not available; NCS = US National

Table S8 contd.

Comorbidity Study; NESARC = National Epidemiologic Survey on Alcohol and Related Conditions; OADP = Oregon Adolescent Depression Project (Lewinsohn *et al*., 1998); QAL, 1978 = Quality of American Life Survey, 1978, Institute of Social Research, University of Michigan; SAQ-A = Social Anxiety Scale for Adults (Caballo *et al*., 2010); UAE = United Arab Emirates; WHO = World Health Organization.

[a] Calculated from *d* values shown in Table 1 of source: values for those in traditional and non-traditional divisions of labor were combined using a random-effects model.

[b] Calculated from proportions in source.

[c] An identical effect size was derived from the report by Maier *et al*. (1999) based on the same survey.

[d] Calculated from means and standard deviations in Table 1 of source: values are for T2 (T1 values: $d = –0.17$ for depression and $d = –0.23$ for rumination).

[e] Standard error calculated from sample sizes and *d* values using formula from Hedges & Olkin (1985, p86); *CI*s calculated from standard error using equations from Borenstein *et al*. (2009, p. 52).

[f] Calculated from the Odds Ratio in the source.

[g] Calculated from means, standard deviations and sample sizes in Table 1 (note that the values in the present table are lower than the ones in the source).

[h] Calculated from means and standard deviations in Table 1 of source: values are for T2 (T1 values: $d = –0.17$ for depression and $d = –0.23$ for rumination).

[i] Calculated from *r* value in source.

[j] Values calculated from figures in Table 1 of source.

[k] Calculated from *t*-values in source.

[l] Mean weighted value calculated from the 10 individual means and standard deviations in Table 1 of source (random effects model).

[m] Calculated from R^2 value in source.

[n] See Table S6 footnote i for details of this analysis, and Table S13 for details of these and other sex differences in personality.

Significance of sex differences

Table S9. Sexuality (**MA:** meta-analysis; **S**: survey).

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
sociosexuality	**S:** Schmitt (2005)	48 nations, ISDP (SOI)	14059		0.80[a]	0.76, 0.83[b]
	S: Lippa (2009)	53 nations, BBC internet survey	200000		0.74[a*]	0.74, 0.75[c]
	S: Penke & Asendorpf (2008)	18 to 50 years, online sample, Germany (SOI-R)	2708		0.61[d]	0.53, 0.69[c]
	S: Cross (2010)	online student sample, UK	3775		0.58	0.52, 0.64[c]
	S: Kennair & Bendixen (2012)	mean age 17 years: online survey, 9 high schools, Norway	1199		0.79[e]	0.67, 0.91[c]
	S: Zheng *et al.* (2014)	representative sample, China	4645		0.64	0.58, 0.70[c]
	S: Kandrik *et al.* (2015)	online student sample, US	4453		0.52[f]	0.46, 0.59
pornography use	**MA:** Petersen & Hyde (2010)	1993–2007	NA	25	0.63[*]	0.39, 0.85
	S: Petersen & Hyde (2010)	mean from 3 large-scale studies (US & Australia)	23794[g]		0.46	0.44, 0.49
	S: Kennair & Bendixen (2012)	mean age 17 years: online survey, 9 high schools, Norway	1199		1.79[h]	1.65, 1.93[c]
sex drive/ arousal	**MA:** Murnen & Stockton (1997)	young adults: self-reported arousal to sexual stimuli	8076	62	0.31	0.27, 0.35[b]
	S: Lippa (2006*a*)	self-reported sex drive: US students US adults (internet survey)	 1622 1519		 0.58 0.82	 0.48, 0.68[c] 0.71, 0.93[c]
	S: Lippa (2009)	53 nations, BBC internet survey	200000		0.62[*]	0.62, 0.63[c]
masturbation	**MA:** Oliver & Hyde (1993)	US & Canada (1966–1990), mostly young adults	NA	26	0.96	0.92, 1.00
	MA: Petersen & Hyde (2010)	1993–2007	NA	66	0.53[*]	0.51, 0.55
	S: Petersen & Hyde (2010)	mean from 3 large-scale studies (US & Australia)	22301[g]		0.58	0.55, 0.61
casual sex	**MA:** Oliver & Hyde (1993)	US & Canada (1966–1990), mostly young adults	NA	10	0.81	0.75, 0.87
	S: Schmitt & ISDP (2003)	ISDP, 52 nations: actively seeking short-term mates	16288		0.49[i*]	0.42, 0.56
	MA: Petersen & Hyde (2010)	1993–2007: incidence attitudes	NA NA	69 10	0.38 0.45	0.37, 0.39 0.39, 0.50
	S: Petersen & Hyde (2010)	mean from 3 large-scale studies (US & Australia) UK National surveys: attitudes 1990 & 2000	20794[g] 12110 18876		0.18 0.45 0.42	0.16, 0.21 0.41, 0.49[c] 0.39, 0.45[c]
	S: Kennair & Bendixen (2012)	mean age 17 years: online survey, 9 high schools, Norway: attitudes	1199		0.79	0.76, 0.82[c]

Table S9 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
number of partners	**MA:** Oliver & Hyde (1993)	US & Canada (1966–1990), mostly young adults	NA	12	0.25	0.19, 0.32
	S: Schmitt & ISDP (2003)	ISDP, 52 nations: mean number desired over time periods from next month to 30 years	16288		0.46[i*]	0.43, 0.49
	S: Petersen & Hyde (2010)	1993–2007 mean from 7 large-scale studies (US, UK & Australia)	NA 78683[g]	256	0.36 0.15	0.35, 0.36 0.14, 0.17
perceptions of sexual interest	**MA:** La France *et al*. (2009)	seductiveness promiscuousness flirtatiousness	3631 3631 3631	28 28 28	0.41[j*] 0.32[j*] 0.18[j*]	0.26, 0.54 0.24, 0.39 0.06, 0.30
mate poaching	**S:** Schmitt & ISDP (2004)	ISDP, 52 nations: short-term attempts long-term attempts	16954		 0.46[i*] 0.36[i*]	 0.40, 0.52 0.26, 0.45

Note. *N* = total number of participants; *k* = number of samples in a meta-analysis; *d* indicates Cohen's *d*, which is positive if in the male direction and negative if in the female direction; *CI* indicates 95% confidence intervals. * indicates studies that were selected for inclusion in Table 3. Comprehensive Meta-Analysis (CMA) was used for calculating effect sizes and *CIs* where these are not taken from the sources.

Abbreviations: BBC, British Broadcasting Corporation; ISDP = International Sexuality Description Project; NA = not available; SOI = Sociosexual Orientation Inventory (Simpson & Gangestad, 1991); SOI-R = Revised Sociosexual Orientation Inventory (Penke & Asendorpf, 2008).

[a] Value calculated from means and standard deviations in Table 6 of source; note that this is different from the value shown there ($d = 0.74$).
[b] Calculated from standard error (or standard deviation or Z score) in source using equations from Borenstein *et al*. (2009, p. 52).
[c] Standard error calculated from sample sizes and *d* values using formula from Hedges & Olkin (1985, p, 86); *CIs* calculated from standard error using equations from Borenstein *et al*. (2009, p. 52).
[d] This is the overall value for the SOI-R. The values for its three subscales were: $d = 0.06$ (behaviour); $d = 0.43$ (attitude); $d = 0.86$ (desire). The value for the original SOI in this study was $d = 0.27$.
[e] This is the value for attitudes. As in the study by Penke & Asendorpf (2008) the components of the SOI covering behaviour had a very small sex difference.
[f] Calculated from *t* value in source.
[g] Calculated from information in source (Table 5, p. 31).
[h] This is the value for frequent exposure to pornography.
[i] Mean of *d* values shown in Table 2 or Table 8 of Schmitt & ISDP (2003) or Table 3 of Schmitt & ISDP (2004), weighted by sample size. Note that the values calculated are slightly different from the composite in the second source.
[j] Calculated from the *r* values in source.

Table S10. Mate choice criteria (**MA:** meta-analysis; **S**: survey).

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
age difference	**MA:** Buss (1989)	37 cultures study	10047	37	–2.00[a*]	–2.05, –1.95
	S: Eastwick *et al*. (2006)	9-nation student sample	3682		–1.74	–1.82, –1.66[b]
	S: Zentner & Mitura (2012)	10 nations differing in gender parity	3177		–1.43[c]	–1.79, –1.06
	S: Souza *et al*. (2016)	Brazilian young adults	1186		–1.07	–1.19, –0.95
financial prospects	**MA:** Buss (1989)	37 cultures study	10047	37	–0.76[a*]	–0.81, –0.72
	S: Eastwick *et al*. (2006)	9-nation student sample	3682		–0.48	–0.55, –0.41[b]
	S: Zentner & Mitura (2012)	10 nations differing in gender parity	3177		–0.81[c]	–1.13, –0.49
	S: Schwarz & Hassebrauck (2012)	ages 18–65 years, online sample, Germany	21245		–0.79[d]	–0.82, –0.76[b]
	S: Souza *et al*. (2016)	Brazilian young adults	1186		–0.59	–0.71, –0.47[b]
good looks	**MA:** Buss (1989)	37 cultures study	10047	37	0.59[a]	0.55, 0.63
	MA: Feingold (1990)	Questionnaires Personal ads	19541 2247	28 6	0.54 0.47	0.51, 0.57[b] 0.39, 0.55[b]
	S: Lippa (2007)	BBC internet survey	200000	53	0.55[*]	0.55, 0.56[b]
	S: Zentner & Mitura (2012)	10 nations differing in gender parity	3177		0.39[c]	0.22, 0.56
	S: Schwarz & Hassebrauck (2012)	ages 18–65 years, online sample, Germany	21245		0.34	0.31, 0.37[b]
	S: Souza *et al*. (2016)	Brazilian young adults	1186		0.32	0.20. 0.44[b]
ambition/ industriousness	**MA:** Buss (1989)	37 cultures study	10047	37	–0.50[a*]	–0.54, –0.46
	MA: Feingold (1992*a*)	Questionnaires	3174	10	–0.67	–0.74, –0.59
	S: Zentner & Mitura (2012)	10 nations differing in gender parity	3177		–0.56[c]	–0.66, –0.45
	S: Souza *et al*. (2016)	Brazilian young adults	1186		–0.21	–0.33, –0.09[b]
social status/ dominance	**MA:** Buss *et al*. (1990)	37 cultures study	9474	37	–0.34[e*]	–0.38, –0.30[b]
	MA: Feingold (1992*a*)	Questionnaires Personal ads	6830 3089	15 8	–0.69 –0.57	–0.74, –0.64 –0.64,–0.50[b]
	S: Zentner & Mitura (2012)	10 nations differing in gender parity	3177		–0.47[c]	–0.70, –0.24
	S: Schwarz & Hassebrauck (2012)	ages 18–65 years, online sample, Germany	21245		–0.52	–0.55, –0.49[b]
	S: Souza *et al*. (2016)	Brazilian young adults	1186		–0.25	–0.37, –0.13[b]
cook–housekeeper	**MA:** Buss *et al*. (1990)	37 cultures study	9474	37	0.56[e*]	0.52, 0.60[b]
	S: Eastwick *et al*. (2006)	9-nation student sample	3682		0.14	0.07, 0.21[b]

Table S10 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
cook-housekeeper contd.	**S:** Zentner & Mitura (2012)	10 nations differing in gender parity	3177		0.09[c]	0.01, 0.16
	S: Schwarz & Hassebrauck (2012)	ages 18–65 years, online sample, Germany	21245		0.17[f]	0.14, 0.20[b]
	S: Souza *et al*. (2016)	Brazilian young adults	1186		0.06	–0.06, 0.18[b]
chastity	**MA:** Buss (1989)	37 cultures study	10047	37	0.30[a*]	0.25, 0.34
	S: Zentner & Mitura (2012)	10 nations differing in gender parity	3177		0.28[c]	0.09, 0.48[b]
	S: Souza *et al*. (2016)	Brazilian young adults	1186		0.11	–0.01, 0.23[b]

Note. *N* = total number of participants; *k* = number of samples in a meta-analysis; *d* indicates Cohen's *d*, which is positive if in the male direction and negative if in the female direction; *CI* indicates 95% confidence intervals. * indicates studies that were selected for inclusion in Table 3. Comprehensive Meta-Analysis (CMA) was used for calculating effect sizes and *CIs* where these are not taken from the sources.

[a] Values taken from a meta-analysis of Buss' data by Archer & Mehdikhani (2003).

[b] Standard error calculated from sample sizes and *d* values using formula from Hedges & Olkin (1985, p. 86); *CIs* calculated from standard error using equations from Borenstein *et al*. (2009, p. 52).

[c] Weighted mean for all 10 nations, calculated from the three effect sizes provided for high, medium and low gender-parity nations in Tables 2, 3 and 4 of source (random-effects model).

[d] The wording was "wealthy and generous".

[e] Calculated from the *r* values in source.

[f] The wording was "creative and homely".

Table S11. Sexual conflict (**MA:** meta-analysis; **S**: survey).

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
(1) SEXUAL VIOLENCE (based on victimization reports unless stated otherwise)						
rape	**S:** Koss *et al*. (1987)	students at 32 US colleges	6159		0.76[a]	0.65, 0.87
	S: Tjaden & Thoennes (1998)	US, NVAWS: past year lifetime	16000		 0.61[a] 1.07[a]	 0.17, 1.05 0.99, 1.14
	S: Walby & Allen (2004)	UK, BCS: by partner past year since age 16 years	19951		 2.32[a*] 1.23[a]	 0.77, 3.86 1.14, 1.31
	S: Basile *et al*. (2007)	US, ICARIS-2: lifetime	9684		0.94[a]	0.82, 1.06
	S: Black *et al*. (2011)	US, NISVS: past year lifetime	16507		 2.82[a] 1.52[a]	 1.29, 4.35[b] 0.41, 1.63[b]
	S: Krahé & Berger (2013)	Germany, 10 universities since age 14 years (online)	2013		0.35[a]	0.18, 0.52
	S: Macdowall *et al*. (2013)	UK, ages 16–74 years, Natsal-3, since age 13 years	14283		1.12[a]	1.00, 1.25[b]
sexual assault/ aggression	**S:** Koss *et al*.(1987)	students at 32 US colleges	6159		0.31[a]	0.21, 0.40
	S: Davidson *et al*. (2004)	Honolulu, 1990-2001, SATC	NA		1.81[c]	NA
	S: Walby & Allen (2004)	UK, BCS: by partner past year since age 16 years	19951		 1.31[a] 1.23[a]	 1.04, 1.57 1.14, 1.31
	S: Sundaram *et al*. (2008)	Danish national samples: 16–39 years 14–16 years	 3932 6185		 1.09[d] 0.81[d]	 0.90, 1.28 0.57, 1.04
	S: Basile *et al*. (2007)	US, ICARIS-2: past year	9684		0.57[a]	0.38, 0.77
	S: Jansson (2007)	UK, BCS: by partner year since age 16 years	 24459 24751		 0.62[a*] 1.12[a]	 0.56, 0.67[b] 1.06, 1.17[b]
	S: Black *et al*. (2011)	US, NISVS: past year lifetime	16489		 0.03 0.57[a]	 0.00, 0.06[e] 0.54, 0.61[b]
	S: Krahé & Berger (2013)	Germany, 10 universities since age 14 years (online): overall heterosexual only	 2121 1722		 0.47[a] 0.49[a]	 0.35, 0.58[b] 0.36, 0.62[b]
	S: Krahé *et al*. (2014)	Nationally representative sample, Czech Republic: lifetime	2000		 1.06[e]	 0.64, 1.47

Table S11 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
perceptions of sexual harassment	**MA:** Rotundo *et al*. (2001)	overall	33164	66	–0.30*	–0.30, –0.30[b]
convictions in rape and child abuse cases	**MA:** Schutte & Hosch (1997)	simulated studies	9813	36	–0.32*	–0.36, –0.28
(2) AGGRESSION TO PARTNERS						
verbal	**MA:** Stockdale *et al*. (2013)	self-reports	NA	20	–0.25*	NA
physical (any)	**MA:** Archer (2000)	self–reports partner-reports	52993 57970	81 75	–0.12 –0.02*	–0.14, –0.10 –0.03, –0.00
physical (severe)	**MA:** Archer (2000)	causing injuries	14542	17	0.35[f]*	0.13, 0.56
	MA: Archer (2002)	beat up choke/strangle	13853 6645	34 15	0.39[g]* 0.52[g]*	0.35, 0.42 0.47, 0.58
homicide	**S:** Fox & Zawitz (2012)	US, 1976–2004 (FBI)	NA		0.65[h]	NA
	S: Smith *et al*. (2014)	US, 2003–2009, 16 states	NA		1.06[h]*	NA
(3) SEXUAL JEALOUSY						
overall	**MA:** Archer (2013)		12540	11	–0.20[i]*	–0.36, –0.04
greater upset from sexual *vs* emotional	**MA:** Dreznick (2003)	forced-choice	NA	37	0.49	0.45, 0.53
	MA: Harris (2003)	forced-choice	NA	29	0.60[j]	0.51, 0.70[g]
	MA: Sagarin *et al*. (2012)	continuous choice	NA	45	0.31	0.23, 0.39
	MA: Carpenter (2012)	forced-choice	10632	37	0.87[k]*	0.83, 0.91
	S: Bendixen *et al*. (2015)	Continuous forced-choice	1074 1074		0.86 0.89[l]	0.80, 0.93[i] 0.76, 1.02[j]

Note. *N* = total number of participants; *k* = number of samples in a meta-analysis; *d* indicates Cohen's *d*, which is positive if in the male direction and negative if in the female direction; *CI* indicates 95% confidence intervals. * indicates studies that were selected for inclusion in Table 3. Comprehensive Meta-Analysis (CMA) was used for calculating effect sizes where these are not taken from the sources.

Abbreviations: BCS = British Crime Survey; FBI = US Federal Bureau of Investigation; ICARIS-2 = Second Injury Control and Risk Survey (Nationally-representative Health Survey, Center for Disease Control and Prevention); NA = not available; Natsal-3 = third National Survey of Sexual Attitudes and Lifestyles (UK); NISVS = National Intimate Partner and Sexual Violence Survey (National Center for Injury Prevention and Control); NVAWS = National Violence Against Women Survey; SATC = Sex Abuse Treatment Center, Honolulu, Hawaii.

Table S11 contd.

[a] *d* values calculated from proportions in source. In the case of Jansson (2007), this was from Figs 3.1 and 3.2.

[b] Calculated from standard error in source using equations from Borenstein *et al*. (2009, p. 52).

[c] Percentage of Chance Encounters in which victim would be female for all victims who received treatment at the centre, converted into a *d* value from Table 1 in Grissom (1994).

[d] Calculated from incidences and sample sizes in source: this produced more accurate figures than the percentages reported as these were expressed to the nearest whole number.

[e] Standard error calculated from sample sizes and *d* values using formula from Hedges & Olkin (1985, p86); *CI*s calculated from standard error using equations from Borenstein *et al*. (2009, p. 52).

[f] This is not the value presented in Archer (2000), which was obtained from proportions *via* DSTAT (*Software for the Meta-analytic Review of Research Literatures:* Johnson, 1989);, since this is likely to be an underestimate (Haddock *et al*., 1998). The value in this table is the result of a re-analysis using CMA for the same 17 studies shown in shown in Table 5 of Archer (2000).

[g] Recognizing the underestimates in the values calculated from proportions in DSTAT (see above), this source presents several *d* values calculated in different ways (Archer, 2002, Table 3). The one presented is from log Odds Ratios calculated from the study-level proportions, from victim reports, using the formula of Hasselblad & Hedges (1995). This corresponds to the computations used in CMA. Sample sizes are taken from Table 5 of source.

[h] Percentage of Chance Encounters in which male would be the killer in a representative sample of opposite-sex partner homicide perpetrators, converted into *d* values from Table 1 in Grissom (1994).

[i] Weighted mean shown in Table 12 of source.

[j] Log Odds Ratios from p. 105 of source, for the 29 studies that involved heterosexual samples, converted into *d* values and *CI*s using the formula of Hasselblad & Hedges (1995).

[k] Calculated from the *r* values and sample size in source.

[l] Calculated from the means and standard deviations. This *d* value is slightly higher than that in the source.

Significance of sex differences

Table S12. Evidence for sex differences and similarities in other attributes (**MA**: meta–analysis; **S**: survey).

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
schizotypy	**MA:** Miettunen & Jääskeläinen (2010)	healthy adults:	41003			
		physical anhedonia		23	0.59*	0.52, 0.66
		social anhedonia		14	0.44*	0.37, 0.50
		magical ideation		29	–0.01*	–0.08, 0.06
		perceptual aberration		22	–0.08*	–0.17, –0.00
body image	**MA:** Feingold & Mazzella (1998)	body image	81172	144	0.52*	0.51, 0.53[a]
		self-rated attractiveness	40807	102	0.17	0.15, 0.19[a]
		physical attractiveness	9187	68	–0.26*	–0.28, –0.24[a]
simple reaction time	**MA:** Silverman (2006)	adults: wide age range	15314	73	0.35[b]*	0.31, 0.38
positive emotions	**S:** Brebner (2003)	students from 41 countries:	6868			
		frequency			–0.20[c]*	–0.29, –0.11
		intensity			–0.23[c]	–0.39, –0.29
		Australian students:	2199			
		frequency			–0.06[c]	–0.13, 0.01
		intensity			–0.02[c]	–0.09, 0.05
	MA: Chaplin & Aldao (2013)	infancy to adolescence mostly preschool ages	NA	146	–0.08	–0.14, –0.03
positive well-being	**MA:** Wood *et al*. (1989)	adults, mostly US	NA	85	–0.01	–0.02, 0.00
dream recall	**MA:** Schredl & Reinhard (2008)	children	4834	15	–0.10	–0.16, –0.04
		adolescents	5725	18	–0.36	–0.46, –0.27
		young adults	15781	81	–0.24*	–0.29, –0.19
		middle-aged adults	16250	36	–0.27	–0.33, –0.21
		older adults	1795	13	–0.24	–0.36, –0.13
nightmare frequency	**MA:** Schredl & Reinhard (2011)	children	42119	29	–0.03	–0.07, 0.01
		adolescents	35333	20	–0.22	–0.28, –0.16
		young adults	40162	42	–0.26*	–0.32, –0.21
		middle-aged adults	61174	19	–0.15	–0.19, –0.10
		older adults	8351	8	–0.10	–0.26, 0.07
narcissism	**MA:** Grijalva *et al*. (2015)	overall	470841	355	0.26*	0.23, 0.28
		entitlement	44108	44	0.29	0.26, 0.32
		leadership/authority	44739	40	0.20	0.16, 0.24
		grandiose/exhibitionism	42460	39	0.04	0.01, 0.08
coping styles	**MA:** Tamres *et al*. (2002)	adults: wishful thinking	1512	11	–0.26[b]*	–0.42, –0.10[a]
		problem-focussed, active	6036	22	–0.26[b]*	–0.29, –0.23[a]
		avoidance	5383	26	–0.06[b]	–0.11, –0.01[a]
self–esteem	**MA:** Lirgg (1991)	children & adults: athletic self-confidence	NA	46	0.40	NA
	MA: Kling *et al*. (1999)	children & adults: overall	97121	216	0.21	0.19, 0.22
	S: Kling *et al*. (1999)	4 large US samples: NCES	48000		0.10[d]	0.04, 0.16
	MA: Major *et al*. (1999)	children & adults: overall	82569	226	0.14	0.13, 0.15
	MA: Pinquart & Sörensen (2001)	older adults: overall	32098	59	–0.08	–0.10, –0.06

Table S12 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
self–esteem contd.	**MA:** Sahlstein & Allen (2002)	adults: overall social physical appearance	379217 39487 15399	124 51 28	–0.02[b] 0.06[b] 0.41[b*]	–0.02, –0.01 0.04, 0.08 0.38, 0.45
	MA: Gentile *et al*. (2009)	school ages to adults: physical appearance athletic personal behavioural conduct	 NA NA NA NA	 76 68 9 56	 0.35 0.41 0.28 –0.17	 0.31, 0.40 0.36, 0.46 0.11, 0.45 –0.28, –0.06
	S: Bleidorn *et al*. (2016)	48 nations: overall	985937	48	0.25[e*]	0.25, 0.25[a]
	MA: Zuckerman *et al*. (2016)	overall	1170935	1148	0.11[*]	0.10, 0.13
self or life satisfaction	**MA:** Pinquart & Sörensen (2001)	older adults (life)	79367	176	–0.08	–0.09, –0.06
	MA: Gentile *et al*. (2009)	school ages to adults (self)	NA	10	0.33	0.18, 0.49
academic self–efficacy	**MA:** Huang (2013)	mostly 11–22 years adults (over 23 years)	68429 NA	247 21	0.08 0.23	0.03, 0.13 0.11, 0.34
self–concept	**MA:** Wilgenbusch & Merrell (1999)	school ages, overall: academic/scholastic physical appearance social athletic	11918 10752 7873 7545 8290	19 17 17 11 16	0.24 0.11 0.37 –0.04 0.39	0.20, 0.28[a] 0.07, 0.15[a] 0.33, 0.41[a] –0.09, 0.01[a] 0.35, 0.43[a]
episodic memory	**S:** Herlitz *et al*. (1997)	ages 35–80 years (Sweden): word recall newly acquired facts activities free recall cued recall	1000		 –0.23[f*] –0.33[*] –0.19[*] –0.29[*] –0.22[*]	 –0.36, –0.10 –0.60, –0.06 –0.48, 0.10 –0.52, –0.06 –0.46, 0.02
working memory	**S:** Redick *et al*. (2012)	southern US: operation symmetry (spatial task) reading	5767 5549 5068		0.09[*] 0.26[*] 0.06	0.04, 0.14[a] 0.21, 0.31[a] 0.00, 0.12[a]
learning orientation	**MA:** Severiens & ten Dam (1998)	young adults: achievement reproduction non-academic	NA NA NA	22 22 22	0.19[*] –0.18[g] 0.13	NA NA NA
vividness of visual imagery	**MA:** Richardson (1995)	adults, VVIQ	NA	11	–0.16[*]	–0.31, –0.02
scholastic achievement	**MA:** Voyer & Voyer (2014)	elementary school to university ages (70% US)	NA	454	–0.25[b*]	–0.28, –0.23
abstract reasoning	**MA:** Lynn & Irwing (2004)	RPM, 15 nations: adults children	9631 60168	10 15	0.33 0.21	0.28, 0.37 0.19, 0.23
	MA: Irwing & Lynn (2005)	RPM, students	20432	22	0.15[h*]	0.12, 0.18[i]

Table 12 contd.

Variable	Study	Sample/measures	*N*	*k*	*d*	*CI*
abstract reasoning contd	**S:** Lynn & Kanazawa (2011)	NCDS: age 7 years 11 years 16 years	14407 14095 11919		-0.08 –0.06 0.11	–0.11, –0.05 –0.09, –0.03[a] 0.07, 0.15[a]
general knowledge	**S:** Herlitz *et al*. (1997)	Sweden, ages 35–80 years	1000		0.04*	–0.12, 0.20[c]
time judgment	**MA:** Block *et al*. (2000)	children and adults	NA	51	–0.06j*	–0.11, –0.01
occupational stress	**MA:** Martocchio & O'Leary (1989)	adults	11364	19	–0.02[b]*	–0.06, 0.02[a]
morningness–eveningness	**MA:** Randler (2007)	healthy people (morningness)	NA	52	–0.08*	–0.14, –0.03

Note. *N* = total number of participants; *k* = number of samples in a meta–analysis; *d* indicates Cohen's *d*, which is positive if in the male direction and negative if in the female direction; *CI* indicates 95% confidence intervals. * indicates studies that were selected for inclusion in Table 3. Comprehensive Meta-Analysis (CMA) was used for calculating effect sizes where these are not taken from the sources.

Abbreviations: NA = not available; NCDS = National Child Development Study (UK longitudinal study, 50 year follow-up); NCES = National Center for Educational Statistics (US nationally representative); RPM = Raven's Progressive Matrices; VVIQ Vividness of Visual Imagery Questionnaire.

[a] Standard error calculated from sample sizes and *d* values using formula from Hedges & Olkin (1985, p. 86); *CI*s calculated from standard error using equations from Borenstein *et al*. (2009, p. 52).

[b] Calculated from *r* value in source using CMA.

[c] Calculated from means, standard deviations and sample sizes in Table 1 (note that the values in the present table are lower than the ones in the source).

[d] Mean of the four values for 17 year olds shown in Table 1 of source.

[e] A single-item scale was used in this study.

[f] Mean of the four values shown in Table 2 of source.

[g] Reproduction orientation involves such characteristics as reliance on memorising, relying on staff to define tasks, over-cautious reliance on details and fear of failure (Severiens & ten Dam, 1998, Table 2)

[h] Overall value: a range of values are presented in Table 2 of the source. Examining a funnel plot of studies included in this analysis indicates the likelihood that the value was an overestimate in the male direction owing to the absence of small-sample studies in the female direction.

[i] Recalculated as the values in Table 1 of the source appear to be incorrect.

[j] *d* value for samples where it was possible to compute an effect size; for all reports including those where it was only stated that there was "no difference" (*k* = 87), the mean was –0.03.

Table S13. Personality factors: *d* values for sex differences in the Big Five from meta-analysis and large samples.

FACTORS	1	2	3	4	5	6	7	8	9	10
AGREEABLENESS	–0.34	–0.35[a]	–0.20[a]	–0.24[a]	–0.22	–0.18	–0.15	–0.56	–0.33	–0.35
NEUROTICISM	–0.25	–0.28[a]	–0.22[a]	–0.27[a]	–0.49	–0.26	–0.40	–0.41	–0.20	–0.27
EXTRAVERSION	0.10	–0.07[a]	–0.08[a]	0.01[a]	–0.16	–0.07	–0.10	–0.15	–0.22	–0.21
OPENNESS	0.12	0.04[a]	–0.16[a]	0.12[a]	0.08	–0.17	0.05	–0.04		0.02
CONSCIENTIOUSNESS	–0.10	0.04[a]	–0.05[a]	–0.01[a]	–0.28	–0.23	–0.12	–0.30		–0.20

Samples:
1 = Guo *et al*. (1995): meta-analysis of studies involving US participants over 16 years of age (agreeableness: *k* = 226, *N* = 309266; neuroticism: *k* = 231, *N* = 290441; extraversion: *k* = 156, *N* = 227559; openness: *k* = 138, *N* = 275184; conscientiousness: *k* =127, *N* = 209980). Values are from Table 1 in source. Various measures were used in the analysis.
2 = Costa *et al*. (2001): US adults (*N* = 1000). From Table 2 in source; measure: NEO-PI-R.
3 = Costa *et al*. (2001): college ages from 26 other cultures (*N* = 10952). From Table 2 in source; measure: NEO-PI-R.
4 = Costa *et al*. (2001): adults from 26 other cultures (*N* = 10690). From Table 2 in source; measure: NEO-PI-R.
5 = Noftle & Shaver (2006): college students, internet survey (*N* = 8318); measure: NEO-PI-R. *d*s are converted from *r* values presented in Table 2 of source, using CMA.
6 = Rammstedt (2007): German population sample (*N* = 2550); measure BFI-10: values calculated from means and standard deviations in Table 2 of source, using CMA.
7 = Schmitt *et al*. (2008): ISDP, 55 nations (*N* = 17637); measure BFI.
8 = Lippa (2008): BBC internet survey (53 nations; *N* = 196461: this is the lower vale taken from footnote 1); measure IPIP (40 items).
9 = Vianello *et al*. (2013): online survey (127 nations, 80% of participants from the US; *N* = 14348) measure IPIP (50 items). Different values were found using an implicit test.
10 = Soto & John (2017): online sample from English-speaking nations (N = 1000): measure BFI-2.

Abbreviations: BBC = British Broadcasting Corporation; BFI = Big Five Inventory (44 items); BFI-2 = Big Five Inventory-2 (60 items); BFI-10 = 10–item BFI; IPIP = International Personality Item Pool; CMA = Comprehensive Meta-Analysis; ISDP = International Sexuality Description Project; NEO-PI-R = Revised NEO Personality Inventory.

[a] Mean values for the six facets listed in Costa *et al*. (2001), Table 2.