Evidence for a Robust, Estradiol-Associated Sex Difference in Narrative-Writing Fluency

Oliver C. Schultheiss and Martin G. Köllner
Friedrich-Alexander University

Holger Busch and Jan Hofer
University of Trier

Objective: Despite evidence for an estradiol-linked sex difference in verbal fluency favoring women, recent reviews question this difference. We therefore examined the issue based on a narrative task that we have administered to different populations for over 20 years. *Method:* We meta-analyzed 98 studies ($N = 11,528$) conducted by our laboratories and that featured measures of biological sex and storytelling. We ran primary-data analyses ($N = 797$) on an overlapping subset of these studies that also included salivary hormone and digit ratio measures. *Results:* Women told longer stories than men, $d = 0.31$, 95% CI [0.24, 0.38], an effect that did not vary by geographic region but was moderated by cue type (verbal: $d = 0.57$, [0.44, 0.71]; pictures: $d = 0.29$, [0.22, 0.36]), response modality (oral: $d = -0.04$, [−0.18, 0.09]; handwriting: $d = 0.39$, [0.31, 0.47]; typing: $d = 0.31$, [0.21, 0.42]), and age (prepubertal children: $d = 0.13$, [−0.04, 0.30]; pubescents: $d = 0.48$, [0.23, 0.74]; premenopausal adults: $d = 0.36$, [0.29, 0.42]; postmenopausal adults: $d = -0.09$, [−0.35, 0.16]). Consistent with the age effect, estradiol, a sex-dimorphic hormone during the reproductive life stage, was a specific mediator of the sex difference in narrative-writing fluency. This mediation effect was moderated by prenatal hormone exposure, estimated via digit ratio. *Conclusions:* When verbal fluency is assessed through narrative writing, a robust female advantage becomes evident. It is associated with the reproductive life stage and variations in current estradiol concentrations, particularly in individuals prenatally exposed to relatively more estradiol than testosterone.

Key Points
Question: Are women verbally more fluent than men on a narrative test? If so, what are key moderators of this difference, and is it associated with hormonal differences? *Findings:* Using meta-analysis and primary-data analysis of studies from our laboratories, we found that women tell longer stories than men on a storytelling measure, that this effect is limited to written stories and the reproductive life stage, and that it is due to differences in estradiol, which was positively associated with narrative-writing fluency. *Importance:* On average, women outperform men on narrative writing, and this female advantage seems to be linked to circulating estradiol concentrations. *Next Steps:* Future research should examine the estradiol link more closely by, for instance, comparing populations with more extreme estradiol differences or using hormone-administration experiments.

Keywords: sex difference, verbal fluency, storytelling, estradiol, development

Supplemental materials: http://dx.doi.org/10.1037/neu0000706.supp

Are women more verbally able than men? Despite a clear tendency for answering this question in the affirmative in the popular literature (Brizendine, 2006; Gray, 1992; Pease & Pease, 2000), research findings on this issue are surprisingly inconsistent and complex, owing in part to the heterogeneity of what is meant by "verbal ability" and how it is measured. Early studies appeared to support the idea that women are more verbally fluent than men (Hampson & Kimura, 1992; Maccoby & Jacklin, 1974). Many of these studies typically used fluency tasks that assessed information retrieval from memory (J. Patterson, 2011) and required generating words fitting a certain category (semantic fluency) or finding words starting with a specific letter (phonemic fluency; Burton, Henninger, & Hafetz, 2005; Hirnstein, Coloma Andrews, & Hausmann, 2014; Weiss, Kemmler, Deisenhammer, Fleischhacker, & Delazer, 2003). Some research also provided evidence for a link between verbal fluency and gonadal steroid hormones. The evidence is particularly strong for estradiol, which is on average higher in normally cycling women than in men (Mazer, 2009) and for which positive associations with verbal fluency were reported (Kimura & Hampson, 1994). The link was established through studies looking at menstrual-cycle-related changes in verbal fluency in women (Hampson, 1990a), in studies assessing estradiol directly (Hampson, 1990b; Maki, Rich, & Rosenbaum, 2002; Wolf & Kirschbaum, 2002), and in research examining effects of hormone administration (Cherrier et al., 2005; Van Goozen, Cohen-Kettenis, Gooren, Frijda, & Van De Poll, 1995). The positive association between verbal fluency and estradiol was also found in other species capable of symbolic sign language (F. G. P. Patterson, Holts, & Saphire, 1991). Other research suggests that testosterone, which is higher in men than in women (Schultheiss, Dlugash, & Mehta, 2019), may play a role in verbal fluency, too, although results are more heterogeneous with regard to studied populations and observed effects (Cherrier et al., 2005; O'Connor, Archer, Hair, & Wu, 2001; Wolf et al., 2000).

However, later studies failed to find consistent sex differences in verbal fluency measures (Mathuranath et al., 2003; Mulac & Lundell, 1994), particularly when verbal fluency was assessed in natural contexts and as spontaneously spoken language (Mehl, Vazire, Ramirez-Esparza, Slatcher, & Pennebaker, 2007). Other studies suggested that even with typical verbal-fluency tasks, sex differences can be observed only under some conditions (e.g., for affective words or when participants were instructed to vary consecutive words or under conditions of stereotype threat) but not others (e.g., for nonaffective words or when participants received no specific instructions or when no stereotype threat was present; Gawda & Szepietowska, 2013; Hirnstein et al., 2014; Scheuringer, Wittig, & Pletzer, 2017). Yet other studies have failed to find a direct link between estradiol and verbal fluency (Scheuringer & Pletzer, 2017; Schultheiss & Zimni, 2015; Taxel, Stevens, Trahiotis, Zimmerman, & Kaplan, 2004).

Consistent with these individual findings, meta-analytic studies provide little support for a verbal-fluency sex difference in favor of women (Hyde, 1981, 2005). Sex differences in verbal fluency conceived of as talkativeness are negligible in childhood and, for adults, even suggest a male advantage (Leaper & Ayres, 2007; Leaper & Smith, 2004). Hyde (2005) argued for gender similarity, not dissimilarity, for an array of verbal ability measures, basing this on an arbitrary cutoff of Cohen's $d = .35$, below which she classified differences as nil or small (see also Zell, Krizan, &

Teeter, 2015). In line with these meta-analytic findings, Wallentin (2009), in a narrative review of the neuropsychological evidence of sex differences in verbal fluency, concluded that "differences in language proficiency do not exist" (p. 175).

We were astonished when we read this conclusion and studied the supporting evidence because in our research, we routinely observe a sizable and very consistent verbal-fluency advantage for women (e.g., Pang & Schultheiss, 2005; Schultheiss & Brunstein, 2001). This sex difference becomes evident when research participants are prompted to tell imaginative stories in response to pictures—and sometimes also verbal cues—portraying individuals in ambiguous social situations. Typically, this test features four to eight cues and a 5-min limit for telling a story about each, usually in written form (Schultheiss & Pang, 2007). The test is termed the picture-story exercise (PSE; McClelland, Koestner, & Weinberger, 1989) and is a descendant of the Thematic Apperception Test (Morgan & Murray, 1935). In our research, we use it to assess motivational imagery from the stories to predict a variety of physiological, cognitive, affective, developmental, and behavioral criteria (Hofer, 2010; Hofer & Busch, 2017; Schultheiss, Frisch, et al., 2019). In the context of this research, we routinely determine the word count of the PSE stories—which we term *narrative fluency* in the context of the present research—and in doing so, we frequently find that women write stories that are approximately one tenth longer than men's. This may not appear to be a big difference at first blush. But it is large enough that we see it in almost every study with college students.

Here, we aimed to quantify the size and variability of this sex difference in narrative fluency using meta-analytic techniques to synthesize the research from our two laboratories. We also examined effects of potential moderators, which of course included our respective laboratories but also geographic region, type and number of cues, response modality, and age. Analyses for geographic region allowed us to examine whether the sex difference depends on a specific cultural context or language; the absence of such an effect would argue against an effect of culture (Van de Vijver & Poortinga, 1982). Testing for an influence of type and number of cues enabled us to determine whether the sex difference depends on the presentation of pictorial or verbal cues and whether it changes with an increasing number of cues (see Schultheiss & Pang, 2007). Analyses for response modality tested the possibility that the sex difference depends on whether stories were told orally, written by hand, or typed on a keyboard (e.g., Schultheiss, Liening, & Schad, 2008; Turk, Brown, Symington, & Paul, 2010). Here, previous reports of small or reverse sex differences for speech or overall talkativeness (Leaper & Ayres, 2007; Leaper & Smith, 2004) suggested a weaker effect of spoken than written or typed responses. Finally, the exploration of the effect of age allowed us to revisit the idea that sex differences in verbal fluency may depend on circulating levels of estradiol (Hampson, 1990b), which are low and similar in boys and girls, are higher on average in postpubertal women compared to men, and become more similar to men's again after menopause (Ojeda, 2012). Thus, assuming a causal, positive effect of estradiol on narrative fluency, we expected a stronger sex difference after puberty and before menopause than in prepubertal childhood or after menopause. This prediction is also supported by the observation that sex differences in verbal fluency emerge only after childhood (Maccoby & Jacklin, 1974) and the meta-analytic finding of a decrease in verbal

fluency from before to after menopause (Weber, Maki, & McDermott, 2014). Moreover, it addresses a lacuna that Kolb and Whishaw (2015) identified in the research literature on sex differences in neurocognitive functions:

> The fact that sex hormones are important to cerebral function in adults leads to an interesting possibility: the cognitive functions of the two sexes may diverge functionally at puberty and begin to converge again in middle age as hormone levels drop. We are unaware of any direct test of this hypothesis. (p. 336)

To arrive at an estimate of the effect size of the sex difference in narrative fluency on the PSE and to be able to explore our moderators, we made unorthodox use of meta-analytic techniques (see also McShane & Böckenholt, 2017) to synthesize word count data from all studies run by our research teams between 1994 and the summer of 2015. We supplemented the data set with additional studies conducted with children and adolescents until 2018 to increase our ability to obtain precise effect size estimates for these age groups relative to the more prevalent adult samples.

Our data set also allowed us to test more directly if estradiol is associated with narrative fluency and mediates the effect of biological sex on this variable. We compiled primary data from 11 studies that featured salivary estradiol measurements in addition to measures of narrative fluency and sex, focusing on those research participants who did not alter their endogenous hormone release through the use of oral contraceptives. This enabled us to examine not only whether higher estradiol is linked to longer stories but also whether there is a sex difference in this association. If there is none, despite an overall sex difference in both narrative fluency and in estradiol, then this would suggest a hormonal mechanism that accounts for the sex difference in narrative fluency. Six of these studies also feature digit ratio measurements (i.e., the ratio of second to fourth digit length, assessed from hand scans), which are thought to represent a marker of the influence of prenatal steroids on the developing brain (Manning, 2002; Manning, Kilduff, Cook, Crewther, & Fink, 2014; Zheng & Cohn, 2011), with relatively higher estradiol being associated with a higher digit ratio and relatively higher testosterone being associated with a lower digit ratio. Prenatal hormone exposure as estimated via digit ratio has been shown to moderate the association between circulating gonadal steroid levels and specific cognitive functions in postnatal life (Donishi, Terada, & Kaneoke, 2018; Manning et al., 2014). We were thus able to explore whether the presumed mediating role of estradiol for the association between sex and narrative fluency depends on prenatal variations in brain development.

Meta-Analysis of Word Count Data

In the following, we first report a retrospective examination of studies from our laboratories in which we used meta-analytic techniques to efficiently summarize observations of a sex difference in narrative fluency. We also examined potential moderators of this difference.

Method

Study selection for pooled laboratory meta-analysis. We compiled sample-level data from all studies conducted by our research teams (i.e., research efforts led or supervised by the first or last author) between the start of our research activities in the mid-1990s and the summer of 2015. Publication status was recorded but was not a criterion for the selection of studies into the data set. This initially led to the identification of 79 suitable studies that fulfilled the dual criteria of mixed-sex composition, with information on participant age and self-identified biological sex (male/female) available and word count data resulting from a picture-story procedure with either pictorial or verbal cues (three additional studies that were originally included were dropped because they did not fulfill the second criterion and featured an episodic-recall task instead). Because scrutiny of the age distribution of these data revealed that children and adolescents were underrepresented in our overall data set, we added sample-level data from another 19 studies conducted between 2015 and 2018 to specifically increase testing power in the low age range (i.e., all samples met the requirement of participants being < 18 years old). This led to a final data set of $k = 98$ studies and associated effect sizes.

Effect size measures. Our meta-analytic effect size measure was the standardized difference in narrative fluency, assessed as the total number of words written by a person, between women and men, as expressed by Cohen's d for each study, which was calculated as follows (Borenstein, Hedges, Higgins, & Rothstein, 2009; Lipsey & Wilson, 2001):

$$d = \frac{M_{Female} - M_{Male}}{SD_{pooled}}$$

SD_{pooled} was calculated as

$$SD_{pooled} = \sqrt{\frac{(n_{Female} - 1)SD^2_{Female} + (n_{Male} - 1)SD^2_{Male}}{n_{Female} + n_{Male} - 2}}$$

As a measure of precision, we also calculated for each study the standard error of d as follows:

$$SE = \sqrt{\frac{n_{Female} + n_{Male}}{n_{Female}n_{Male}} + \frac{d^2}{2(n_{Female} + n_{Male})}}$$

Moderators. We coded each study for the following moderators: number of cues, coded as the number of pictorial or verbal cues to collect stories in each study; cue type, coded as pictorial when picture cues were used and coded as verbal when verbal cues were used to elicit stories; response modality, coded as oral when spoken stories were recorded and later transcribed, coded as handwritten when participants wrote their stories using paper and pencil, and coded as typed when participants wrote their stories on a keyboard; geographic region, reflecting whether stories were collected in North America (United States, testing language: English), Central America (Costa Rica, testing language: Spanish), Europe (Germany, testing language: German; Czech Republic, testing language: Czech; Luxembourg, testing language: Luxembourgish, French, and German), Africa (Cameroon and Zambia, testing language: both English), or Asia (China and Hong Kong, testing language: English and Mandarin); laboratory, indicating whether the data were collected in Hofer's or Schultheiss' laboratory; and age, coded as the average age of each sample in years. For follow-up analyses on the last moderator, we also created subgroups such that samples with an average age < 9 years were classified as prepubertal children (lowest sample-level mean age:

5.38 years),[1] samples between 9 and 17 years were classified as pubertal based on the onset of Tanner Stage II and the approximate end of Stage V, samples between 18 and 50 years were classified as adult men and premenopausal women, and samples older than (and including) 50 years were classified as adult men and postmenopausal women based on the average age of menopause onset (Collaborative Group on Hormonal Factors in Breast Cancer, 2012; Lee, Guo, & Kulin, 2001; highest sample-level mean age: 85.45 years). Because our laboratories had access to the original study sampling plans, data collection materials, and data, we could code each moderator directly from the information available for each study.

Statistical methods. We prepared data and calculated effect size and standard error measures in SYSTAT 13. We then analyzed data using the meta-analysis module of JASP 0.9.2, an R-based open-source GUI package developed at the University of Amsterdam (JASP Team, 2019). We used restricted maximum likelihood random-effects models in all analyses.

Results

A restricted maximum likelihood random-effects model across all $k = 98$ studies in our data set, representing 11,528 research participants, yielded a small-to-medium-sized sex difference in overall narrative fluency, Cohen's $d = 0.31$, 95% CI [0.24, 0.38], $Z = 9.13$, $p < .001$, favoring women (see Figure 1). Sample-size weighted averages showed that across all studies, women ($n = 6,209$) wrote 93.09 words per cue and men ($n = 5,319$) wrote 86.15 words per cue; thus, women wrote stories that were 8% longer than men's. Supplemental analyses for the data of one study in which we had collected data on typing speed in addition to data on narrative-writing fluency suggested that although women were slightly faster typers than men and typing speed and narrative-writing fluency were positively correlated, controlling for typing speed did not account for or substantially alter the size of the sex difference effect on narrative-writing fluency (see Section 4 in the online supplemental materials). Meta-analytic effect size heterogeneity was substantial ($I^2 = 64.53\%$, $Q_{residual} = 290.93$, $p < .001$) and justified testing moderators in subsequent mixed-effects models. As shown in Table 1, effect sizes were similar across all geographic regions our research teams had obtained data from and did not systematically differ between our laboratories. However, they differed according to cue type, with verbal cues eliciting a larger effect size than pictures. They also differed according to response modality, with handwritten or typed stories yielding sex differences of medium size and orally narrated stories showing no sex difference. Metaregression results indicated that the number of cues was not associated with the effect size of the sex difference, $B = 0.001$, $[-0.041, 0.044]$, $SE = 0.022$, $z = 0.06$, $p = .95$. Finally, age group also turned out to be a significant moderator, with sex effects close to zero in the prepubertal group and in the group of adult men and postmenopausal women and indicating a female advantage in narrative fluency in the pubertal group and the group of adult men and premenopausal women. The age effect also emerged when, instead of using cutoffs for age groups, we added sample age as a continuous variable in a metaregression: Here, the cubic term was significant, $B = 0.0000157$, $[0.00000697, 0.0000245]$, $SE = 0.0000045$, $z = 3.52$, $p < .001$, and captured the

nonlinear association between age and effect size, depicted in Figure 2, rather well.

Primary Data Analysis

Next, we conducted a primary data analysis on a subset of 10 studies that were also included in the previous meta-analysis, plus one additional study with female participants only that was not, to determine whether estradiol and other hormones, measured in participants' saliva, mediate the association between sex and narrative-writing fluency. We also examined whether such a mediating effect of estradiol depends on variations in prenatal estradiol exposure, estimated via measurements of digit length.

Method

Study selection for primary data analysis. We identified all studies containing data on narrative fluency and salivary estradiol and run until 2015 by the first author's laboratory. We integrated them into one data set, which consisted of 797 individuals (271 women and 526 men) and was based on 10 studies in which salivary estradiol measures and narrative fluency using a picture-story method were assessed in men and women and one additional study in which these variables were assessed in women only. These studies had received institutional review board approval by the University of Michigan and Friedrich-Alexander University. All participants provided informed consent prior to testing and were treated according to the principles expressed in the Declaration of Helsinki. Salivary estradiol was assessed using the radioimmunoassay previously used by and following validated protocols developed in the first author's laboratory (Oxford, Tiedtke, Ossmann, Özbe, & Schultheiss, 2017; Schultheiss, Dargel, & Rohde, 2003) and log transformed to reduce skew. Although these studies originally also included women who self-identified as using hormonal contraceptives, we excluded them from all further analyses because (a) this type of birth control exerts a profound suppressive effect on naturally occurring estradiol and (b) estrogens are sometimes contained in the contraceptives and may add to, or interact with, endogenous estradiol, influencing cognitive functions in complex ways (Beltz, Hampson, & Berenbaum, 2015; Griksiene & Ruksenas, 2011). Six of these studies ($N = 555$; Oxford et al., 2017; Schultheiss, Frisch, et al., 2019; Schultheiss & Zimni, 2015) also featured scans of participants' hands, from which the length measurements of the second and fourth digits were taken and then converted to overall digit ratio scores by dividing the second digit length by the fourth digit length and averaging ratio scores across hands. The procedures employed in these measurements are characterized by high intercoder reliability in the first author's laboratory (i.e., $rs > .96$; Schultheiss, Frisch, et al., 2019).

Statistical methods. We used JASP 0.10.0 to run Bayesian linear regressions (Wagenmakers, Love, et al., 2018; Wagenmakers, Marsman, et al., 2018) testing associations between sex, salivary estradiol, and narrative-writing fluency. All hypotheses were

---

[1] For a description of the collection of oral picture stories in kindergarten children, please see Raihala and Hansen (2018); for a description of the collection of written stories in school-age children, please see Raihala and Kranz (2019).
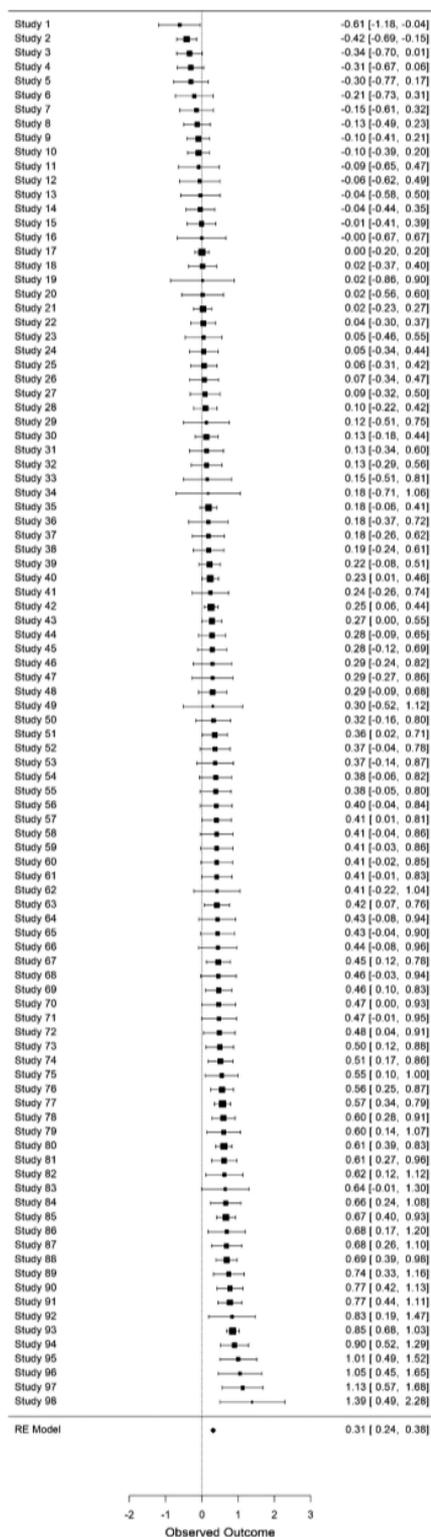
Figure 1. Forest plot for rank-ordered effect sizes (Cohen's *d*, 95% lower and upper limits) of all studies included in the restricted maximum likelihood random-effects (RE) meta-analysis. For further details on the studies, please see https://osf.io/zr9wn/. Figure available at https://osf.io/zr9wn/ under a CC-BY4.0 license.

tested against a null model containing dummy variables coding for the 11 studies included in the overall sample and, in the case of the association between salivary estradiol and narrative-writing fluency, sex as an additional covariate. We used default settings (Jeffreys-Zellner-Siow prior and a default beta binomial model prior with $a = 1$ and $b = 1$) for the priors and Bayesian adaptive sampling with 1,000 samples for estimating credible intervals. We report regression parameters from the posterior summaries of coefficients. All hypothesis tests and parameter estimates were two-sided, with the null hypothesis being represented by both a constant and slope of zero. In interpreting the Bayes factor (*BF*), which we always calculated such that values > 1 represent evidence in favor of the alternative hypothesis (i.e., BF_{10}), we followed the conventional interpretation of its magnitude (i.e., 1 to 2.99 = anecdotal, 3 to 9.99 = moderate, 10 to 29.99 strong, 30 to 100 = very strong, > 100 = extreme; Quintana & Williams, 2018). We used the "mediation analysis" option in JASP's structural equation model (SEM) module to test simple and moderated mediation models, again controlling for between-study differences via dummy variables for all studies included and using standardized sample scores for the mediator (estradiol) and its moderator (digit ratio). For moderated mediation, we also used Hayes and Preacher's (2013) PROCESS module (Model 14) in SPSS to calculate conditional indirect effects after residualizing all zero-order variables in the mediation model for the dummy variables coding between-study differences. All data files, analysis output, and supplemental results are available from https://osf.io/zr9wn/.

Results

Evidence for sex as a positive predictor of both narrative fluency ($B = 8.57$, $SD = 2.02$, 95% credible interval [4.61, 12.54], $BF_{10} = 2,084$) and salivary estradiol ($B = 0.197$, $SD = 0.0265$, [0.145, 0.249], $BF_{10} = 7.5605e + 10$) was extreme. Moreover, salivary estradiol was a positive predictor after controlling for sex, $B = 5.69$, $SD = 2.67$, [0.45, 10.93], although the evidence in favor of this effect over the null hypothesis was only in the anecdotal range (Quintana & Williams, 2018), $BF_{10} = 2.60$. We found no evidence for a difference between women ($\beta = .068$, $p = .29$, $n = 271$) and men ($\beta = .119$, $p = .02$, $n = 526$) in the association between salivary estradiol and narrative fluency; differences in statistical significance of this association in men and women were driven primarily by differences in sample size. A model testing all meaningful combinations of main and interaction effects of sex and salivary estradiol on narrative fluency after controlling for potential between-study differences (i.e., the null model) indicated that (a) the model including two main effects of estradiol and sex outperforms the null model best ($BF_{10} = 5,411$) and that (b) its Bayes factor is about threefold larger than the one for the model that includes the additional interaction term ($BF_{10} = 1,711$). The latter finding suggests that the observed data are 3 times more likely under the two-main-effects model than under the main-effects-plus-interaction model.

An SEM using bias-corrected percentile scores from 5,000 bootstrap samples indicated that the indirect effect of sex on narrative fluency with salivary estradiol as mediator was greater than zero, $B = 1.25$, $SE = 0.58$, 95% CI [0.22, 2.50], $z = 2.15$, $p = .032$. However, the direct effect of sex on narrative fluency re-

Table 1
Results of Moderator Analyses

Moderator	k	N	d	95% CI	df	Q_W	Q_B	I^2 (%)
Age group					3, 94		20.75***	56.17
<9 years (prepubertal)	12	804	0.131	[−0.040, 0.301]	1, 11	2.26		28.30
9–17 years (pubertal)	7	1,386	0.483	[0.227, 0.739]	1, 6	13.65***		77.87
18–50 years (adult men and premenopausal women)	72	8,297	0.356	[0.288, 0.424]	1, 71	106.10***		52.79
>50 years (adult men and postmenopausal women)	7	1,041	−0.092	[−0.346, 0.161]	1, 6	0.51		71.26
Geographic region					4, 93		2.97	64.10
Africa	11	1,848	0.195	[−0.039, 0.429]	1, 10	2.67		82.98
Asia	6	748	0.213	[−0.035, 0.461]	1, 5	2.84		61.11
Europe	55	6,423	0.352	[0.264, 0.440]	1, 54	61.28***		63.05
Central America	5	470	0.336	[−0.049, 0.721]	1, 4	2.93		72.25
North America	21	2,039	0.310	[0.193, 0.426]	1, 20	27.05***		35.21
Cue type					1, 96		4.74*	62.80
Pictorial	92	10,601	0.290	[0.221, 0.359]	1, 91	67.41***		64.18
Verbal	6	927	0.571	[0.435, 0.707]	1, 5	67.83***		0
Response modality					2, 95		20.51***	55.98
Oral	14	1,542	−0.044	[−0.175, 0.087]	1, 13	0.44		31.42
Handwritten	57	7,015	0.389	[0.301, 0.472]	1, 56	84.22***		61.76
Typed	27	2,971	0.312	[0.205, 0.420]	1, 26	32.57***		48.15
Laboratory					1, 96		0.16	64.62
Hofer	52	6,512	0.299	[0.201, 0.397]	1, 51	35.83***		71.12
Schultheiss	46	5,016	0.328	[0.244, 0.412]	1, 45	58.62***		48.29

Note. CI = confidence interval.
* $p < .05$.   *** $p < .001$.

mained significant after inclusion of the mediator in the model, $B = 7.99$, $SE = 2.15$, [3.44, 12.47], $z = 3.72$, $p = .0002$.

We therefore also explored to what extent the mediating effect of estradiol depends on variations in prenatal hormone exposure, as reflected by digit ratio measurements. When we repeated the SEM described above with digit ratio (z scores) and its interaction term with salivary estradiol (z scores) as additional mediators ($N = 555$), the total indirect effect of sex on narrative fluency was again greater than zero, $B = 2.11$, $SE = 1.02$, 95% CI [0.0467, 4.36], $z = 2.06$, $p = .0392$. The 95% confidence interval for the specific

Figure 2. Relationship between average sample age in years and effect size, with a LOESS smoother adjusted for individual samples' standard error. Symbol size represents sample weight (1/*SE*). Figure available at https://osf.io/zr9wn/ under a CC-BY4.0 license. See the online article for the color version of this figure.

effect of the Digit Ratio × Estradiol interaction term—0.034 to 2.084—excluded zero ($B = 0.698$, $SE = 0.391$, $z = 1.78$, $p = .074$; note that the significance test reflects the result obtained under normal-theory assumptions and differs from the confidence interval results). Supplemental analyses showed that for individuals with a low digit ratio (−1 *SD*), estradiol had no mediating effect, [−2.80, 2.34], whereas for individuals with a high digit ratio (+1 *SD*), it was a statistically reliable mediator, [0.91, 5.48]. For individuals with average digit ratios, the effect estimate still included zero, [−0.06, 3.41]. These findings indicate that the mediating role of estradiol for the sex → narrative fluency path depends on prenatal hormone exposure, as indexed by digit ratio, and that the mediating effect of estradiol is present only for those individuals who have been exposed to relatively high estradiol and/or relatively low testosterone prenatally.

We also conducted exploratory mediation analyses adding salivary testosterone (men > women) or progesterone (women > men) to the model or replacing salivary estradiol with these variables but failed to find specific associations between these hormone measures and narrative-writing fluency (see online supplemental materials). Thus, our mediation analysis findings described above are specific to estradiol.

Discussion

Our meta-analytic synthesis of studies from our laboratories shows that, overall, there is a small-to-medium sex difference in narrative fluency, with women producing stories that are 8% longer than men's, regardless of how many picture cues were presented. To put this finding in perspective, our observed percentage difference between women and men is very similar to the percentage difference in median height in adults (women: 163.33 cm, men: 176.85 cm; difference = 8.3%; National Center for Health Statistics, 2000), although the variability for the former difference is greater than for the latter. However,
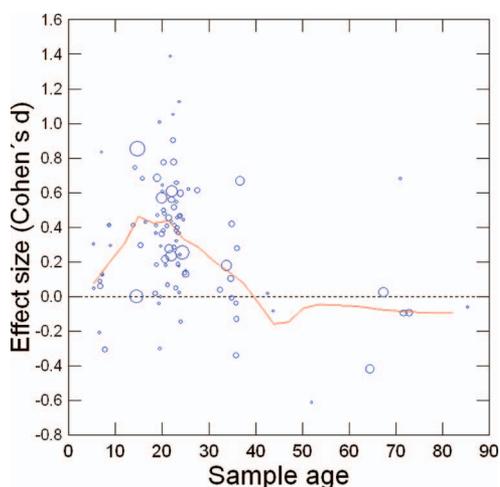
confirming Kolb and Whishaw's (2015) hypothesis that some cognitive sex differences should covary with age-related differences in hormone levels, moderator analyses indicated that the sex difference in narrative fluency was stronger (i.e., medium sized) when participants belonged to the pubertal and adult, premenopausal age groups. Conversely, males and females showed similar narrative fluency when studies tested prepubertal children or adults in the postmenopausal age range, although we acknowledge that these latter age brackets were represented by relatively fewer data points and that the attenuated gender difference in prepubertal children may reflect a floor effect due to limited vocabulary.[2] Consistent with a biological explanation of this result pattern, our findings also suggest that the female advantage for narrative fluency does not depend on samples' cultural, ethnic, or language background: Across all geographic regions in which we had sampled picture stories—North America, Central America, Europe, Africa, and Asia—and the diverse races, cultures, and languages they represent, we obtained a sex difference of comparable size.

Our findings also indicate that the sex difference is stronger when instead of the more typical picture cues, verbal cues are used to elicit storytelling, although it remained robust for pictorial cues, too, and we acknowledge that results for the verbal cues are based on only six studies. This finding may reflect compatibility between the processing format of the stimulus and the response, with the sex difference becoming more pronounced if both are consistently represented in a verbal code (e.g., Paivio, 1986). Intriguingly, other studies show that high estradiol is associated with a shift from nonverbal processing in the right hemisphere to verbal processing bilaterally or in the left hemisphere (Hausmann, Becker, Gather, & Güntürkün, 2002; Hjelmervik et al., 2012; Weis et al., 2008). When these perspectives are combined, the marked sex difference in narrative fluency in response to verbal cues may be taken to suggest that estradiol facilitates associations within the verbal system more than between verbal and nonverbal processing systems.

We followed up on the meta-analytic finding of a greater sex difference between puberty and menopause—that is, during a life phase in which women's and men's hormone levels differ considerably—with primary data analysis in a subset of studies for which we had salivary estradiol measurements available. Restricting our analysis to those individuals with natural hormonal release patterns (i.e., excluding oral-contraceptive users), we were able to confirm the robust difference in narrative-writing fluency between women and men. Furthermore, women also had robustly higher salivary estradiol levels than men. Overall, estradiol was positively associated with narrative-writing fluency even after controlling for sex in addition to between-study differences, and there was no compelling evidence to suggest that this association differed between women and men. The latter finding suggests that estradiol is associated with narrative fluency to the same extent in both sexes and that women have higher narrative fluency than men because they have higher estradiol levels on average between menarche and menopause. Moreover, the association between narrative-writing fluency and hormone levels was specific to estradiol as we failed to find unique associations with salivary testosterone and progesterone, two hormones that also differ between the sexes during the reproductive life phase. However, the strength of the association between estradiol and narrative-writing fluency was only in the anecdotal range of Bayesian standards of evidence (Johnson, 2013), and estradiol did not fully account for the statistical effect of sex on narrative fluency.

One reason for the relatively weak association between estradiol and narrative-writing fluency may be that variations in circulating estradiol are only one factor that may affect cognitive functions such as narrative writing; variations in neuronal tissue sensitivity to estradiol (mediated, for instance, by estrogen receptor type or density) may be another critical contributor to the sex difference. To address this issue, we examined whether variations in hormonally organized prenatal brain development, as reflected in the digit ratio marker measure, moderate the association between estradiol and narrative-writing fluency and hence the mediating effect of estradiol for the sex/fluency link. Findings from a moderated mediation analysis modeling the interaction between digit ratio and estradiol suggested that this was the case. Follow-up analyses indicated that for individuals who had been exposed to high levels of estradiol and/or low levels of testosterone prenatally, as reflected in high digit ratio scores, and whose brains' sensitivity to estradiol therefore had presumably been retained (Zheng & Cohn, 2011), estradiol was a significant mediator of the sex effect on narrative-writing fluency. This was not the case for individuals prenatally exposed to the converse hormonal constellation—reflected in low digit ratio scores—whose brains were presumably less sensitive to circulating estradiol. However, like the simple mediation model described above, the moderated-mediation model failed to account for the full extent of the link between sex and narrative fluency.

Another likely reason for the incomplete (moderated) mediation effects may therefore reside in the considerable measurement error characterizing our mediators. Estradiol, which is present typically in the single-digit picogram per milliliter range in saliva, is notoriously difficult to quantify precisely (Rosner, Hankinson, Sluss, Vesper, & Wierman, 2013). Consistent with this, measurement error for salivary estradiol assays in our studies, estimated via the coefficient of variation between duplicate measurements, almost always exceeded thresholds that are deemed acceptable for hormones measured at higher concentrations. Nevertheless, despite these problems of measurement reliability, we consider the salivary estradiol assays used in the included studies valid because they typically yielded acceptable recovery of quality control checks (Oxford et al., 2017; Schultheiss et al., 2003; Schultheiss & Zimni, 2015; Stanton & Schultheiss, 2007), captured the expected difference between women and men, and also differentiated between women on and off oral contraceptives (see online supplemental materials). Unlike estradiol assays, which aim to quantify a biological effect directly, digit ratio measurements represent only indirect measures of prenatal hormone exposure. Here, soft tissues interfere with the exact measurement of bone length (Manning, 2002), the focal indicator of prenatal hormone effects on the body. Findings and their interpretations with this measure are further complicated by the fact that both estradiol and testosterone have counteracting effects on digit ratio (Zheng & Cohn, 2011).

Thus, although our hypothesis that estradiol mediates the effect of sex on narrative-writing fluency, particularly in individuals with brains that are highly sensitive to the effects of estradiol, may have merit, the lack of precision associated with both salivary estradiol and digit ratio measures and the ambiguity inherent in the latter measure may have considerably attenuated the observed strength of the hypothesized mediation effects. This conclusion is also consistent with

---

[2] We thank an anonymous reviewer for suggesting this alternative explanation.

the age-graded pattern of the sex difference in our overall meta-analysis, which would be difficult to explain without the consideration of hormonal factors (cf. Kolb & Whishaw, 2015). Replication studies are needed that employ more precise and/or more direct measures of circulating estradiol (e.g., via mass spectrometry; Schultheiss, Dlugash, & Mehta, 2019) and prenatal hormone exposure (e.g., through amniotic fluid assessment; Beking et al., 2018) to determine the validity of our hypothesis and the replicability of our findings.

Notably, our meta-analytical findings also indicate that the sex difference is completely absent when narrative fluency is assessed orally but present when some type of writing (handwriting, typing on the keyboard) is involved. We tentatively rule out a mere motor ability difference because a detailed analysis (reported in the online supplemental materials) of the relative contributions of typing speed and sex to narrative-writing fluency effect in one of our studies indicated that the effect of the latter predictor is independent of the one of the former. However, the null effect for oral assessment is based on a comparatively small sample of studies ($k = 14$), which also come from the prepubertal and postmenopausal age brackets. So how generalizable are these findings? Research from other groups of researchers using similar picture-story measures as we do indicates that sex differences in narrative fluency are absent when stories are told orally by prepubertal children, pubertal adolescents (Costantino & Malgady, 1996), and adults alike (van den Daele, Yates, & Jenkins, 2018; Veroff, Atkinson, Feld, & Gurin, 1960). More broadly, our findings are consistent with previous meta-analyses that fail to find substantial sex differences in verbal fluency when direct oral measures are used (Leaper & Ayres, 2007; Leaper & Smith, 2004).

The difference in observed effects for oral and written storytelling could be due to differences in the neurological pathways and cognitive functions underlying spoken and written language (Basso, Taborelli, & Vignolo, 1978; Ruigrok et al., 2014). For instance, timed narrative-writing tasks like ours may depend more strongly on working memory than oral narrations (Azuma, 2004; McCutchen, 2011; Swanson & Berninger, 1996), and working memory capacity varies with fluctuations in estradiol (Hampson, 2018, 2019). These observations suggest that women show better narrative-writing fluency than men because (and when) they have higher estradiol, which in turn enhances their working memory (Hampson, 2018) and thereby helps them construct more elaborate written narratives under time restrictions. The same may not hold for spoken language. The proposed mechanism is sex-specific only to the extent that during a certain life phase, one sex tends to have higher estradiol than the other. However, it also explains within-sex variations in the strength of the estradiol/narrative-writing fluency association: As our results suggest, within each sex, individuals benefit from higher estradiol levels when they write narrative material, presumably because of their relatively better working memory capacity. It is also conceivable that in comparison to oral storytelling, narrative-writing tasks place greater demands on episodic memory (possibly in interaction with working memory; see McCutchen, 2011) and the retrieval of specific representative episodes. There is evidence that episodic memory may benefit from the high gonadal steroid levels in women during the gonadally active life phase (e.g., Rentz et al., 2017), an observation that is consistent with estradiol's well-documented beneficial effects on neuronal processes and synaptic connections, particularly in memory-relevant structures like the hippocampus (e.g., Luine, 2014).

We conclude that the sex difference in *written* narrative fluency is robust and highly replicable but absent in oral narrative fluency

or speech more generally. This also suggests that the wholesale dismissal (Wallentin, 2009) of sex differences in verbal ability may have been premature.

The fact that all studies included in our meta-analysis came from only two research teams is both a strength and a limitation. It is a strength because both teams use very similar protocols for collecting picture stories and have both done so extensively in different geographic regions. This resulted in a large data set of high methodological quality. It is a limitation because our findings may be specific to our laboratories, our data collection methods, and the specific task we used (i.e., the PSE). However, we do think that our findings may have some degree of generalizability when we look at others' research. For instance, in a large-scale Swiss cohort study with 736 participants aged 20 to 80 years, women consistently wrote longer PSE stories than men, while at the same time also showing evidence of a decrease of this sex difference over the life span ($ds = 0.78, 0.42$, and 0.27 in young, middle-aged, and aged adults, respectively; Denzinger, Backes, Job, & Brandstätter, 2016). At the other end of the age spectrum, a large-scale U.S. study of 2,495 children and adolescents 6 to 21 years old and requiring participants to compose a written newsletter (i.e., a different task) found an overall narrative-fluency female advantage of $d = 0.40$ (Scheiber, Reynolds, Hajovsky, & Kaufman, 2015). Similar to our findings, however, the effect depended on age: It was absent in children but present in adolescents and young adults. Last but not least, a recent meta-analysis of mandatory state-level tests of verbal ability in elementary, middle, and high school students in the United States (Petersen, 2018) documented not only an increase in overall sex differences in verbal abilities from third to ninth grade but also a specific sex difference in composition writing of $d = 0.45$ that was larger than sex differences on all other verbal ability tests. We conclude that our findings are consistent with the results obtained in not only other studies with high statistical power but also studies coming from other research groups and employing varied measures of narrative-writing fluency. They also feature the same age trends as our studies, with sex differences higher in adulthood but lower or absent in prepubertal childhood or after menopause.

Another limitation of our study is that although we screened for and excluded women taking hormonal contraceptives for the primary data analyses involving salivary estradiol, we did not generally have this information available for the studies in our main meta-analysis. However, if our account about the role of estradiol in narrative fluency is correct, we expect women taking hormonal contraceptives combining estrogen and progesterone, which suppress endogenous gonadal steroid levels but substitute estrogen, to write stories of similar or perhaps greater length than naturally cycling women. For the same reason, we expect women using progesterone-only contraceptives, which suppress endogenous gonadal steroid levels without supplementing exogenous estrogen, to write shorter stories (Beltz et al., 2015; Griksiene & Ruksenas, 2011; Montoya & Bos, 2017). Future research needs to explore this issue as well as the more general question of how strong the association between estradiol and narrative-writing fluency can become when samples with more extreme estradiol levels (e.g., ovulating or pregnant women) are targeted and compared with low-estradiol controls.

A third limitation is the uneven spread of sampled effect sizes across the life course, which was due to the overrepresentation of samples of college students and young adults. Hence, our effect size estimates for younger and older samples do not have the same precision as our estimates for the early and middle adulthood samples.

A final limitation is that in our studies, we determined sex by self-report only, and response options did not allow indicating a nonbinary identity. We therefore may have misidentified some genetically female or male individuals who indicated the opposite sex deliberately or by mistake, who have a nonbinary gender identity, or who have a sex phenotype that diverges from their genotype due to various anomalies of sexual differentiation. Future research explicitly targeting the sex difference in narrative-writing fluency should therefore use more differentiated and multisource measures of sex and gender identity.

To conclude, our study provides evidence for a robust, medium-sized sex difference in narrative-writing fluency among pubertal adolescents and adults in the age bracket before female menopause that is absent in prepubertal children, postmenopausal adults, or samples that have been tested using oral speech instead of writing. Our study also shows that estradiol, which is higher in women than men during the gonadally active years, is positively associated with narrative-writing fluency in both sexes and partially mediates the association between sex and narrative-writing fluency. Moreover, this effect appears to depend on prenatal hormone exposure of the brain, being stronger in individuals with a high digit ratio, a morphological marker of relatively higher prenatal estradiol compared to testosterone, than in individuals with a low digit ratio.

References

Azuma, T. (2004). Working memory and perseveration in verbal fluency. *Neuropsychology, 18,* 69–77. http://dx.doi.org/10.1037/0894-4105.18.1.69

Basso, A., Taborelli, A., & Vignolo, L. A. (1978). Dissociated disorders of speaking and writing in aphasia. *Journal of Neurology, Neurosurgery & Psychiatry, 41,* 556–563. http://dx.doi.org/10.1136/jnnp.41.6.556

Beking, T., Geuze, R. H., van Faassen, M., Kema, I. P., Kreukels, B. P. C., & Groothuis, T. G. G. (2018). Prenatal and pubertal testosterone affect brain lateralization. *Psychoneuroendocrinology, 88,* 78–91. http://dx.doi.org/10.1016/j.psyneuen.2017.10.027

Beltz, A. M., Hampson, E., & Berenbaum, S. A. (2015). Oral contraceptives and cognition: A role for ethinyl estradiol. *Hormones and Behavior, 74,* 209–217. http://dx.doi.org/10.1016/j.yhbeh.2015.06.012

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2009). *Introduction to meta-analysis.* Hoboken, NJ: Wiley. http://dx.doi.org/10.1002/9780470743386

Brizendine, L. (2006). *The female brain.* New York, NY: Broadway Books.

Burton, L. A., Henninger, D., & Hafetz, J. (2005). Gender differences in relations of mental rotation, verbal fluency, and SAT scores to finger length ratios as hormonal indexes. *Developmental Neuropsychology, 28,* 493–505. http://dx.doi.org/10.1207/s15326942dn2801_3

Cherrier, M. M., Matsumoto, A. M., Amory, J. K., Ahmed, S., Bremner, W., Peskind, E. R., . . . Craft, S. (2005). The role of aromatization in testosterone supplementation: Effects on cognition in older men. *Neurology, 64,* 290–296. http://dx.doi.org/10.1212/01.WNL.0000149639.25136.CA

Collaborative Group on Hormonal Factors in Breast Cancer. (2012). Menarche, menopause, and breast cancer risk: Individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *The Lancet Oncology, 13,* 1141–1151. http://dx.doi.org/10.1016/S1470-2045(12)70425-4

Costantino, G., & Malgady, R. G. (1996). Development of TEMAS, a multicultural thematic apperception test: Psychometric properties and clinical utility. In G. R. Sadowsky & J. C. Impara (Eds.), *Multicultural assessment in counseling and clinical psychology* (pp. 85–136). Lincoln, NE: Buros Institute of Mental Measurements.

Denzinger, F., Backes, S., Job, V., & Brandstätter, V. (2016). Age and gender differences in implicit motives. *Journal of Research in Personality, 65,* 52–61. http://dx.doi.org/10.1016/j.jrp.2016.09.003

Donishi, T., Terada, M., & Kaneoke, Y. (2018). Effects of gender, digit ratio, and menstrual cycle on intrinsic brain functional connectivity: A whole-brain, voxel-wise exploratory study using simultaneous local and global functional connectivity mapping. *Brain and Behavior, 8*(1), e00890. http://dx.doi.org/10.1002/brb3.890

Gawda, B., & Szepietowska, E. M. (2013). Semantic and affective verbal fluency: Sex differences. *Psychological Reports, 113,* 246–256. http://dx.doi.org/10.2466/28.21.PR0.113x17z3

Gray, J. (1992). *Men are from Mars, women are from Venus.* New York, NY: HarperCollins.

Griksiene, R., & Ruksenas, O. (2011). Effects of hormonal contraceptives on mental rotation and verbal fluency. *Psychoneuroendocrinology, 36,* 1239–1248. http://dx.doi.org/10.1016/j.psyneuen.2011.03.001

Hampson, E. (1990a). Estrogen-related variations in human spatial and articulatory-motor skills. *Psychoneuroendocrinology, 15,* 97–111. http://dx.doi.org/10.1016/0306-4530(90)90018-5

Hampson, E. (1990b). Variations in sex-related cognitive abilities across the menstrual cycle. *Brain and Cognition, 14,* 26–43. http://dx.doi.org/10.1016/0278-2626(90)90058-V

Hampson, E. (2018). Estrogens, aging, and working memory. *Current Psychiatry Reports, 20,* 109. http://dx.doi.org/10.1007/s11920-018-0972-1

Hampson, E. (2019). Estrogens and androgens in the prefrontal cortex: Relevance for cognition and decision-making. In O. C. Schultheiss & P. H. Mehta (Eds.), *Routledge international handbook of social neuroendocrinology* (pp. 371–390). New York, NY: Routledge.

Hampson, E., & Kimura, D. (1992). Sex differences and hormonal influences on cognitive function in humans. In J. B. Becker, S. M. Breedlove, & D. Crews (Eds.), *Behavioral endocrinology* (pp. 357–398). Cambridge, MA: MIT Press.

Hausmann, M., Becker, C., Gather, U., & Güntürkün, O. (2002). Functional cerebral asymmetries during the menstrual cycle: A cross-sectional and longitudinal analysis. *Neuropsychologia, 40,* 808–816. http://dx.doi.org/10.1016/S0028-3932(01)00179-8

Hayes, A. F., & Preacher, K. J. (2013). Conditional process modeling: Using structural equation modeling to examine contingent causal processes. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 219–266). Charlotte, NC: Information Age Publishing.

Hirnstein, M., Coloma Andrews, L., & Hausmann, M. (2014). Gender-stereotyping and cognitive sex differences in mixed- and same-sex groups. *Archives of Sexual Behavior, 43,* 1663–1673. http://dx.doi.org/10.1007/s10508-014-0311-5

Hjelmervik, H., Westerhausen, R., Osnes, B., Endresen, C. B., Hugdahl, K., Hausmann, M., & Specht, K. (2012). Language lateralization and cognitive control across the menstrual cycle assessed with a dichotic-listening paradigm. *Psychoneuroendocrinology, 37,* 1866–1875. http://dx.doi.org/10.1016/j.psyneuen.2012.03.021

Hofer, J. (2010). Research on implicit motives across cultures. In O. C. Schultheiss & J. C. Brunstein (Eds.), *Implicit motives* (pp. 433–467). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195335156.003.0015

Hofer, J., & Busch, H. (2017). Why Citizen Kane was unhappy: Motive-goal incongruence. *Social and Personality Psychology Compass, 11*(8), e12330. http://dx.doi.org/10.1111/spc3.12330

Hyde, J. S. (1981). How large are cognitive gender differences? A meta-analysis using ω^2 and *d*. *American Psychologist, 36,* 892–901. http://dx.doi.org/10.1037/0003-066X.36.8.892

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60,* 581–592. http://dx.doi.org/10.1037/0003-066X.60.6.581

JASP Team. (2019). *JASP (Version 0.9.2 and 0.10.2)* [Computer software]. Retrieved from https://jasp-stats.org/

Johnson, V. E. (2013). Revised standards for statistical evidence. *Proceedings of the National Academy of Sciences of the United States of America, 110,* 19313–19317. http://dx.doi.org/10.1073/pnas.1313476110

Kimura, D., & Hampson, E. (1994). Cognitive pattern in men and women is influenced by fluctuations in sex hormones. *Current Directions in Psychological Science, 3,* 57–61. http://dx.doi.org/10.1111/1467-8721.ep10769964

Kolb, B., & Whishaw, I. Q. (2015). *Fundamentals of human neuropsychology* (7th ed.). New York, NY: Worth.

Leaper, C., & Ayres, M. M. (2007). A meta-analytic review of gender variations in adults' language use: Talkativeness, affiliative speech, and assertive speech. *Personality and Social Psychology Review, 11,* 328–363. http://dx.doi.org/10.1177/1088868307302221

Leaper, C., & Smith, T. E. (2004). A meta-analytic review of gender variations in children's language use: Talkativeness, affiliative speech, and assertive speech. *Developmental Psychology, 40,* 993–1027. http://dx.doi.org/10.1037/0012-1649.40.6.993

Lee, P. A., Guo, S. S., & Kulin, H. E. (2001). Age of puberty: Data from the United States of America. *APMIS, 109*(S103), 81–88. http://dx.doi.org/10.1034/j.1600-0463.2001.d01-107.x

Lipsey, M. W., & Wilson, D. (2001). *Practical meta-analysis*. Atlanta, GA: Sage.

Luine, V. N. (2014). Estradiol and cognitive function: Past, present and future. *Hormones and Behavior, 66,* 602–618. http://dx.doi.org/10.1016/j.yhbeh.2014.08.011

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Palo Alto, CA: Stanford University Press.

Maki, P. M., Rich, J. B., & Rosenbaum, R. S. (2002). Implicit memory varies across the menstrual cycle: Estrogen effects in young women. *Neuropsychologia, 40,* 518–529. http://dx.doi.org/10.1016/S0028-3932(01)00126-9

Manning, J. T. (2002). *Digit ratio: A pointer to fertility, behavior, and health*. New Brunswick, NJ: Rutgers University Press.

Manning, J., Kilduff, L., Cook, C., Crewther, B., & Fink, B. (2014). Digit ratio (2D:4D): A biomarker for prenatal sex steroids and adult sex steroids in challenge situations. *Frontiers in Endocrinology, 5,* 9. http://dx.doi.org/10.3389/fendo.2014.00009

Mathuranath, P. S., George, A., Cherian, P. J., Alexander, A., Sarma, S. G., & Sarma, P. S. (2003). Effects of age, education and gender on verbal fluency. *Journal of Clinical and Experimental Neuropsychology, 25,* 1057–1064. http://dx.doi.org/10.1076/jcen.25.8.1057.16736

Mazer, N. A. (2009). A novel spreadsheet method for calculating the free serum concentrations of testosterone, dihydrotestosterone, estradiol, estrone and cortisol: With illustrative examples from male and female populations. *Steroids, 74,* 512–519. http://dx.doi.org/10.1016/j.steroids.2009.01.008

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96,* 690–702. http://dx.doi.org/10.1037/0033-295X.96.4.690

McCutchen, D. (2011). From novice to expert: Language and memory processes in the development of writing skill. *Journal of Writing Research, 3,* 51–68. http://dx.doi.org/10.17239/jowr-2011.03.01.3

McShane, B. B., & Böckenholt, U. (2017). Single-Paper Meta-Analysis: Benefits for Study Summary, Theory Testing, and Replicability. *Journal of Consumer Research, 43,* 1048–1063. http://dx.doi.org/10.1093/jcr/ucw085

Mehl, M. R., Vazire, S., Ramírez-Esparza, N., Slatcher, R. B., & Pennebaker, J. W. (2007). Are women really more talkative than men? *Science, 317,* 82. http://dx.doi.org/10.1126/science.1139940

Montoya, E. R., & Bos, P. A. (2017). How oral contraceptives impact social-emotional behavior and brain function. *Trends in Cognitive Sciences, 21,* 125–136. http://dx.doi.org/10.1016/j.tics.2016.11.005

Morgan, C., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology & Psychiatry, 34,* 289–306. http://dx.doi.org/10.1001/archneurpsyc.1935.02250200049005

Mulac, A., & Lundell, T. L. (1994). Effects of gender-linked language differences in adults' written discourse: Multivariate tests of language effects. *Language & Communication, 14,* 299–309. http://dx.doi.org/10.1016/0271-5309(94)90007-8

National Center for Health Statistics. (2000). *Percentile data files with LMS values: STATAGE data set*. Retrieved from https://www.cdc.gov/growthcharts/percentile_data_files.htm

O'Connor, D. B., Archer, J., Hair, W. M., & Wu, F. C. W. (2001). Activational effects of testosterone on cognitive function in men. *Neuropsychologia, 39,* 1385–1394. http://dx.doi.org/10.1016/S0028-3932(01)00067-7

Ojeda, S. R. (2012). Female reproductive function. In W. J. Kovacs & S. R. Ojeda (Eds.), *Textbook of endocrine physiology* (6th ed., pp. 194–238). New York, NY: Oxford University Press.

Oxford, J. K., Tiedtke, J. M., Ossmann, A., Özbe, D., & Schultheiss, O. C. (2017). Endocrine and aggressive responses to competition are moderated by contest outcome, gender, individual versus team competition, and implicit motives. *PLoS One, 12*(7), e0181610. http://dx.doi.org/10.1371/journal.pone.0181610

Paivio, A. (1986). *Mental representations. A dual coding approach*. Oxford, UK: Oxford University Press.

Pang, J. S., & Schultheiss, O. C. (2005). Assessing implicit motives in U.S. college students: Effects of picture type and position, gender and ethnicity, and cross-cultural comparisons. *Journal of Personality Assessment, 85,* 280–294. http://dx.doi.org/10.1207/s15327752jpa8503_04

Patterson, F. G. P., Holts, C. L., & Saphire, L. (1991). Cyclic changes in hormonal, physical, behavioral, and linguistic measures in a female lowland gorilla. *American Journal of Primatology, 24*(3–4), 181–194. http://dx.doi.org/10.1002/ajp.1350240305

Patterson, J. (2011). Verbal fluency. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of clinical neuropsychology* (pp. 2603–2606). New York, NY: Springer. http://dx.doi.org/10.1007/978-0-387-79948-3_1423

Pease, A., & Pease, B. (2000). *Why men don't listen and women can't read maps: How we're different and what to do about it*. New York, NY: Welcome Rain.

Petersen, J. (2018). Gender difference in verbal performance: A meta-analysis of united states state performance assessments. *Educational Psychology Review, 30,* 1269–1281. http://dx.doi.org/10.1007/s10648-018-9450-x

Quintana, D. S., & Williams, D. R. (2018). Bayesian alternatives for common null-hypothesis significance tests in psychiatry: A non-technical guide using JASP. *BMC Psychiatry, 18,* 178. http://dx.doi.org/10.1186/s12888-018-1761-4

Raihala, C., & Hansen, G. (2018). Power stress in primary school children. *Motivation and Emotion, 43,* 82–92. http://dx.doi.org/10.1007/s11031-018-9724-0

Raihala, C., & Kranz, D. (2019). Choose it or lose it: The implicit power motive in children and their resource control behavior. *Motivation Science, 5,* 86–91. http://dx.doi.org/10.1037/mot0000093

Rentz, D. M., Weiss, B. K., Jacobs, E. G., Cherkerzian, S., Klibanski, A., Remington, A., . . . Goldstein, J. M. (2017). Sex differences in episodic memory in early midlife: Impact of reproductive aging. *Menopause, 24,* 400–408. http://dx.doi.org/10.1097/GME.0000000000000771

Rosner, W., Hankinson, S. E., Sluss, P. M., Vesper, H. W., & Wierman, M. E. (2013). Challenges to the measurement of estradiol: An endocrine

society position statement. *The Journal of Clinical Endocrinology and Metabolism, 98,* 1376–1387. http://dx.doi.org/10.1210/jc.2012-3780

Ruigrok, A. N. V., Salimi-Khorshidi, G., Lai, M.-C., Baron-Cohen, S., Lombardo, M. V., Tait, R. J., & Suckling, J. (2014). A meta-analysis of sex differences in human brain structure. *Neuroscience and Biobehavioral Reviews, 39,* 34–50. http://dx.doi.org/10.1016/j.neubiorev.2013.12.004

Scheiber, C., Reynolds, M. R., Hajovsky, D. B., & Kaufman, A. S. (2015). Gender differences in achievement in a large, nationally representative sample of children and adolescents. *Psychology in the Schools, 52,* 335–348. http://dx.doi.org/10.1002/pits.21827

Scheuringer, A., & Pletzer, B. (2017). Sex differences and menstrual cycle dependent changes in cognitive strategies during spatial navigation and verbal fluency. *Frontiers in Psychology, 8,* 381. http://dx.doi.org/10.3389/fpsyg.2017.00381

Scheuringer, A., Wittig, R., & Pletzer, B. (2017). Sex differences in verbal fluency: The role of strategies and instructions. *Cognitive Processing, 18,* 407–417. http://dx.doi.org/10.1007/s10339-017-0801-1

Schultheiss, O. C., & Brunstein, J. C. (2001). Assessment of implicit motives with a research version of the TAT: Picture profiles, gender differences, and relations to other personality measures. *Journal of Personality Assessment, 77,* 71–86. http://dx.doi.org/10.1207/S15327752JPA7701_05

Schultheiss, O. C., Dargel, A., & Rohde, W. (2003). Implicit motives and gonadal steroid hormones: Effects of menstrual cycle phase, oral contraceptive use, and relationship status. *Hormones and Behavior, 43,* 293–301. http://dx.doi.org/10.1016/S0018-506X(03)00003-5

Schultheiss, O. C., Dlugash, G., & Mehta, P. H. (2019). Hormone measurement in social neuroendocrinology: A comparison of immunoassay and mass spectrometry methods. In O. C. Schultheiss & P. H. Mehta (Eds.), *Routledge international handbook of social neuroendocrinology* (pp. 26–40). New York, NY: Routledge.

Schultheiss, O. C., Frisch, M., Özbe, D., Ossmann, A., Schultheiss, M., Lentz, S., . . . Rösch, A. G. (2019). Implicit motives show sex-dimorphic associations with digit ratio. *Motivation Science, 5,* 326–342. http://dx.doi.org/10.1037/mot0000122

Schultheiss, O. C., Liening, S., & Schad, D. (2008). The reliability of a picture story exercise measure of implicit motives: Estimates of internal consistency, retest reliability, and ipsative stability. *Journal of Research in Personality, 42,* 1560–1571. http://dx.doi.org/10.1016/j.jrp.2008.07.008

Schultheiss, O. C., & Pang, J. S. (2007). Measuring implicit motives. In R. W. Robins, R. C. Fraley, & R. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 322–344). New York, NY: Guilford Press.

Schultheiss, O. C., & Zimni, M. (2015). Associations between implicit motives and salivary steroids, 2D:4D digit ratio, mental rotation performance, and verbal fluency. *Adaptive Human Behavior and Physiology, 1,* 387–407. http://dx.doi.org/10.1007/s40750-014-0012-2

Stanton, S. J., & Schultheiss, O. C. (2007). Basal and dynamic relationships between implicit power motivation and estradiol in women. *Hormones and Behavior, 52,* 571–580. http://dx.doi.org/10.1016/j.yhbeh.2007.07.002

Swanson, H. L., & Berninger, V. W. (1996). Individual differences in children's working memory and writing skill. *Journal of Experimental Child Psychology, 63,* 358–385. http://dx.doi.org/10.1006/jecp.1996.0054

Taxel, P., Stevens, M. C., Trahiotis, M., Zimmerman, J., & Kaplan, R. F. (2004). The effect of short-term estradiol therapy on cognitive function in older men receiving hormonal suppression therapy for prostate cancer. *Journal of the American Geriatrics Society, 52,* 269–273. http://dx.doi.org/10.1111/j.1532-5415.2004.52067.x

Turk, A. A., Brown, W. S., Symington, M., & Paul, L. K. (2010). Social narratives in agenesis of the corpus callosum: Linguistic analysis of the Thematic Apperception Test. *Neuropsychologia, 48,* 43–50. http://dx.doi.org/10.1016/j.neuropsychologia.2009.08.009

van den Daele, L., Yates, A., & Jenkins, S. R. (2018). Differential performance of professional dancers to the Music Apperception Test and the Thematic Apperception Test. *Rorschachiana, 39,* 20–49. http://dx.doi.org/10.1027/1192-5604/a000098

Van de Vijver, F. J. R., & Poortinga, Y. H. (1982). Cross-cultural generalization and universality. *Journal of Cross-Cultural Psychology, 13,* 387–408. http://dx.doi.org/10.1177/0022002182013004001

Van Goozen, S. H. M., Cohen-Kettenis, P. T., Gooren, L. J. G., Frijda, N. H., & Van de Poll, N. E. (1995). Gender differences in behaviour: Activating effects of cross-sex hormones. *Psychoneuroendocrinology, 20,* 343–363. http://dx.doi.org/10.1016/0306-4530(94)00076-X

Veroff, J., Atkinson, J. W., Feld, S. C., & Gurin, G. (1960). The use of thematic apperception to assess motivation in a nationwide interview study. *Psychological Monographs, 74,* 1–32. http://dx.doi.org/10.1037/h0093760

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., . . . Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review, 25,* 58–76. http://dx.doi.org/10.3758/s13423-017-1323-7

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., . . . Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review, 25,* 35–57. http://dx.doi.org/10.3758/s13423-017-1343-3

Wallentin, M. (2009). Putative sex differences in verbal abilities and language cortex: A critical review. *Brain and Language, 108,* 175–183. http://dx.doi.org/10.1016/j.bandl.2008.07.001

Weber, M. T., Maki, P. M., & McDermott, M. P. (2014). Cognition and mood in perimenopause: A systematic review and meta-analysis. *The Journal of Steroid Biochemistry and Molecular Biology, 142,* 90–98. http://dx.doi.org/10.1016/j.jsbmb.2013.06.001

Weis, S., Hausmann, M., Stoffers, B., Vohn, R., Kellermann, T., & Sturm, W. (2008). Estradiol modulates functional brain organization during the menstrual cycle: An analysis of interhemispheric inhibition. *The Journal of Neuroscience, 28,* 13401–13410. http://dx.doi.org/10.1523/JNEUROSCI.4392-08.2008

Weiss, E. M., Kemmler, G., Deisenhammer, E. A., Fleischhacker, W. W., & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and Individual Differences, 35,* 863–875. http://dx.doi.org/10.1016/S0191-8869(02)00288-X

Wolf, O. T., & Kirschbaum, C. (2002). Endogenous estradiol and testosterone levels are associated with cognitive performance in older women and men. *Hormones and Behavior, 41,* 259–266. http://dx.doi.org/10.1006/hbeh.2002.1770

Wolf, O. T., Preut, R., Hellhammer, D. H., Kudielka, B. M., Schürmeyer, T. H., & Kirschbaum, C. (2000). Testosterone and cognition in elderly men: A single testosterone injection blocks the practice effect in verbal fluency, but has no effect on spatial or verbal memory. *Biological Psychiatry, 47,* 650–654. http://dx.doi.org/10.1016/S0006-3223(99)00145-6

Zell, E., Krizan, Z., & Teeter, S. R. (2015). Evaluating gender similarities and differences using metasynthesis. *American Psychologist, 70,* 10–20. http://dx.doi.org/10.1037/a0038208

Zheng, Z., & Cohn, M. J. (2011). Developmental basis of sexually dimorphic digit ratios. *Proceedings of the National Academy of Sciences of the United States of America, 108,* 16289–16294. http://dx.doi.org/10.1073/pnas.1108312108

Correction to Schultheiss et al. (2021)

In the article "Evidence for a Robust, Estradiol-Associated Sex Difference in Narrative-Writing Fluency," by Oliver C. Schultheiss, Martin G. Köllner, Holger Busch, and Jan Hofer (*Neuropsychology*, 2021, Vol. 35, No. 3, pp. 323–333, https://doi.org/10.1037/neu0000706), there was an error in Table 1. The *df* for "18–50 years (adult men and premenopausal women)," originally read "1, 17," but should have read "1, 71." The online version of this article has been corrected.

https://doi.org/10.1037/neu0000762