Many of the women who take hormonal contraceptives discontinue because of unwanted side effects, including negative psychological effects. Yet scientific evidence of psychological effects is mixed, partly because causal claims are often based on correlational data. In correlational studies, possible causal effects can be difficult to separate from selection effects, attrition effects, and reverse causality. Contraceptive use and, according to the congruency hypothesis, congruent contraceptive use (whether a woman’s current use/non-use of a hormonal contraceptive is congruent with her use/non-use at the time of meeting her partner) have both been thought to influence relationship quality and sexual functioning. In order to address potential issues of observed and unobserved selection effects in correlational data, we studied a sample of up to 1,179 women to investigate potential effects of contraceptive use and congruent contraceptive use on several measures of relationship quality and sexual functioning: perceived partner attractiveness, relationship satisfaction, sexual satisfaction, and diary measurements including libido, frequency of vaginal intercourse, and frequency of masturbation. No evidence for substantial effects was found except for a positive effect of hormonal contraceptives on frequency of vaginal intercourse and a negative effect of hormonal contraceptives on frequency of masturbation. These effects were robust to the inclusion of observed confounders, and their sensitivity to unobserved confounders was estimated. No support for the congruency hypothesis was found. Our correlational study was able to disentangle, to some extent, causal effects of hormonal contraceptives from selection effects by estimating the sensitivity of reported effects. To reconcile experimental and observational evidence on hormonal contraceptives, future research should scrutinize the role of unobserved selection effects, attrition effects, and reverse causality.

Theory

An estimated 248 million women worldwide used some form of hormonal contraceptive in 2019—this represents 27% of all women aged 15–49 years who were using any form of contraception (United Nations, Department of Economic and Social Affairs, Population Division, 2019). The most common form of hormonal contraceptive, the birth control pill, also known as the pill, has had an enormous impact on women’s reproductive health and their role in modern society since its commercial release over six decades ago: It has reduced the number of unwanted pregnancies and helped change women’s economic status (Goldin & Katz, 2002). The pill has been called “the most important scientific advance of the 20th century” (G. Harris, 2010).

Despite the undoubted advantages of hormonal contraceptives, the pill can also cause medical side effects, including venous thromboembolism, headaches, and breast pain. Package inserts and some studies suggest that psychological side effects such as mood changes and decreased libido may also arise (e.g., Lee et al., 2017; Lindh et al., 2009; Sanders et al., 2001; Westhoff et al., 2007). However, as Graham (2019) noted, little research has investigated the link between the pill and sexual functioning in general. According to Graham, one of the most consistent findings in the research that does exist is the variability in women’s experience with the pill: While some women showed improved sexual functioning related to pill use, some showed adverse effects, and some showed no changes at all. Overall, treatment heterogeneity (i.e., interindividual differences in psychological responses to hormonal contraceptives) is not well understood. A few studies have investigated the effects...
of hormonal contraceptives on psychological outcomes using randomized controlled trials (RCTs) with a placebo control group (Graham et al., 1995; Zethraeus et al., 2016, 2017) or head-to-head RCTs comparing different forms of hormonal contraceptives or different forms of the pill (Ornanratanaphan & Taneapanichkul, 2006; Sabatini & Cagiano, 2006; Strufaldi et al., 2010). The benefit of control groups is that they allow comparisons between hormonal contraceptives and nonhormonal contraceptives in general while head-to-head RCTs have little to say about general effects of hormonal contraceptives.

Though RCTs with a placebo control can provide gold standard evidence of average causal effects, existing studies do not address the role of individual experience in women’s choices to go on and off hormonal contraceptives. Correlational studies in the psychological literature are more concerned with understanding individual differences through the analysis of moderators and mechanisms, but causal conclusions from these mainly cross-sectional studies must be treated with care because of the potential for selection effects, attrition effects, and reverse causality. Selection effects describe the possibility that interindividual differences in women affect both the choice of contraceptive method and psychological outcomes of interest here. Attrition effects occur when women who experience negative effects of hormonal contraceptives discontinue using them, so that the remaining hormonal contraceptive users are more likely to experience no effects or positive effects (Vitzthum & Ringheim, 2005). Selection and attrition could both exaggerate or mask possible causal effects of hormonal contraception. The possibility that relationships between psychological outcomes and hormonal contraceptive use in correlational studies might exist because the outcome influences the contraceptive choice (e.g., higher frequency of vaginal intercourse might lead to the decision to start using hormonal contraceptives) is called reverse causality.

In addition, most studies of sexual functioning have relied on retrospective reports (McAuliffe et al., 2007) which are less reliable due to retrospective biases or on reactions to artificial stimuli with limited external validity rather than experience sampling. In all, there is no complete picture of how women differ in their psychological reactions to hormonal contraceptives and how this shapes their choices about contraception. Our study aimed to help fill in this picture by disentangling selection effects from potential causal effects of hormonal contraceptive use. First, we investigated selection effects on choice of contraceptive method (no/nonhormonal vs. hormonal) and congruent contraceptive use (incongruent vs. congruent with use at the time of meeting the current partner). Second, we estimated the robustness of potential effects of choice of contraceptive method and congruent contraceptive use on relationship quality and sexual functioning after taking into account observed confounders. Finally, we estimated the sensitivity of the effects of hormonal contraceptives and congruent contraceptive use in the light of potential unobserved confounders.

Effects of Hormonal Contraceptives in Normally Cycling Women and Women Using Hormonal Contraceptives

Hormonal contraceptives contain synthetic progesterone (i.e., progestin), which suppresses the natural production of estrogen and progesterone. Hormonal contraceptives reduce variation in estrogen and progesterone across the menstrual cycle, flattening spikes in estrogen before ovulation and during the secretory phase as well as spikes in progesterone following ovulation (Fleischman et al., 2010). Altering the endocrine system could lead to unexpected changes to body and mind. In a study by Wiegratz et al. (2005), hormonal contraceptive use was found to lead to a decrease of free testosterone and an increase in the levels of serum-binding globulins, including sex hormone-, thyroxine-, and corticosteroid-binding globulins. A systematic review and meta-analysis by Zimmerman et al. (2014) concluded that oral hormonal contraceptives reduce levels of total as well as free testosterone and increase levels of sex hormone-binding globulin. In addition, some studies suggest that hormonal contraceptives affect brain structure and activity, although these studies were generally small (sample sizes between 28–56), not preregistered, and showed conflicting findings. For instance, Lisofsky et al. (2016) reported decreased gray matter in the anterior parahippocampal gyrus, while Pletzer et al. (2010) found larger anterior parahippocampal regions in women using the pill. The evidence base for effects of hormonal contraceptives on brain structure and activity remains uncertain. Because of the changes to the endocrine system, hormonal contraceptives have been predicted to affect sexuality and even partner preferences (Alvergne & Lummaa, 2010).

Hormonal Contraceptive Use and Partner Preferences

The idea that hormonal contraceptives could influence

---

1 While Zethraeus et al. (2016, 2017) included all women, Graham et al. (1995) specifically recruited women who had been sterilized or whose partners had been sterilized.

2 Although our study had a correlative nature and was therefore not able to directly determine causal effects, we controlled for potential selection variables in order to estimate potential causal effects. To make the proposed causal structure as transparent and comprehensible as possible, a directed acyclic graph was drawn, incorporating all included selection variables, predictors, and outcomes (see Figure 1). In addition, the sensitivity analyses quantitatively formulated the assumptions about unobserved confounders that must be taken into account when considering potential causal effects of hormonal contraceptives and congruent contraceptive use. As the aim of the study was to infer causal effects based on correlational data as closely as possible and to avoid hiding this causal aim behind correlational language (Grosz et al., 2020), the terms effects of hormonal contraceptives and effects of congruent contraceptive use are used throughout this manuscript.
partner preferences is based on the premise that women's hormone levels affect their partner preferences (Gildersleeve et al., 2014). Alvergne and Lummaa (2010) argued that hormonal contraceptives alter mate choice by removing the preference shift in the fertile phase of the menstrual cycle. Summarizing the evidence at that time, they concluded that while normally cycling women preferred more masculine, symmetrical, and genetically unrelated men during ovulation compared to the luteal phase of their cycle, women using the pill showed no variation in partner preferences across their cycle. Current studies focusing on the association between hormonal contraceptive use and preference for masculine faces are mixed: Some studies found evidence for a negative relationship (Feinberg et al., 2008; Little et al., 2002, 2015) and others reported positive associations (Cobey et al., 2015; Jones, Hahn, Fisher, et al., 2018) or no significant link at all (Marcinkowska et al., 2019).

Based on Alvergne and Lummaa's (2010) conclusions, Roberts et al. (2015) formulated the congruency hypothesis, which states that it is not the direct effect of hormonal contraceptives but whether a woman has changed contraceptive methods since meeting her current partner that influences her partner preferences and ultimately affects the relationship quality. If a woman's hormonal contraceptive use is congruent (i.e., is the same as when she met her partner), her partner preference is also likely to be unchanged, leading to higher sexual satisfaction and relationship satisfaction. Changing contraceptive methods (from hormonal contraceptives to no or nonhormonal contraceptives, or vice versa) could lead to changes in partner preferences owing to the absence of menstrual cycle changes in women using hormonal contraceptives. Changing contraceptive methods could result in a decrease in attraction to one's partner and relationship satisfaction.

However, nonreplications and methodological criticisms (Arslan et al., 2018; C. R. Harris et al., 2013; Jones, Hahn, & DeBruine, 2018; Jünger, Kordsmeyer, et al., 2018; Jünger, Motta-Mena, et al., 2018; Stern et al., 2020, 2021; Stern & Penke, in press; Wood et al., 2014) have cast doubt on reports that hormonal changes across the cycle cause changes in partner preferences, as concluded by Gildersleeve et al. (2014). Nevertheless, there is robust evidence that hormonal changes across the cycle lead to changes in sexual desire that do not occur in women using hormonal contraceptives (Arslan et al., 2018). In addition, normally cycling women evaluate male bodies, voices, and behavior as generally more attractive in their fertile phase (Jünger, Kordsmeyer, et al., 2018; Jünger, Motta-Mena, et al., 2018; Stern et al., 2020, 2021).

**Hormonal Contraceptive Use, Sexuality, and Satisfaction**

Beyond partner preferences, there is a growing body of research investigating effects of hormonal contraceptives on sexuality and satisfaction. Several mechanisms explaining how hormonal contraceptives might influence sexuality have been suggested. For example, the estrogen-induced increase in the production of sex hormone binding globulins based on hormonal contraceptives might lead to a decrease in libido by lowering the amount of free, biologically active testosterone (Pastor et al., 2013; Zimmerman et al., 2014). At the same time, other mechanisms like overcoming the fear of unwanted pregnancies, the resolution of gynecologic disorders (e.g., endometriosis, dysmenorrhea), and the reduction of body image concerns with an increase in self-esteem for women with clinical hyperandrogenism might be mechanisms behind potential positive effects of hormonal contraceptives on sexuality (Both et al., 2019). The recent review by Both et al. (2019) found that only a minority of women experienced changes in sexual functioning and concluded that the effects of hormonal contraceptives on sexual functioning— and sexual desire in particular—are understudied and poorly understood.

This is supported by a large amount of previous research on hormonal contraceptives producing mixed results for a wide range of psychological outcomes, including sexual functioning (Laessoe et al., 2014; Oranratanaphan & Taneepanichskul, 2006; Panzer et al., 2006; C. W. Wallwiener et al., 2015; M. Wallwiener et al., 2010; Zethraeus et al., 2016), libido (Caruso et al., 2005; Graham et al., 1995; Graham & Sherwin, 1993; Mark et al., 2016; McCoy & Matyas, 1996; Oranratanaphan & Taneepanichskul, 2006; Sabatini & Cagiano, 2006; Walker & Bancroft, 1990; Zethraeus et al., 2016; for reviews see Burrows et al., 2012; Davis & Castaño, 2004; Lee et al., 2017; Pastor et al., 2013; Schaffir, 2006), sexual and masturbation frequency (Alexander et al., 1990; Bancroft et al., 1991; Caruso et al., 2005; McCoy & Matyas, 1996), and sexual satisfaction (Alexander et al., 1990; Caruso et al., 2005; Jern et al., 2018; Oranratanaphan & Taneepanichskul, 2006), as well as relationship satisfaction (Jern et al., 2018; Taggart et al., 2018), jealousy (Cobey et al., 2011, 2012, 2015; Geary et al., 2001; Jern et al., 2018; Welling et al., 2012) and general well-being (Caruso et al., 2005; Egarter et al., 1999; Taggart et al., 2018; Zethraeus et al., 2017). There is therefore a need for methodologically sound and rigorous studies investigating effects of hormonal contraceptives on sexuality and satisfaction.

**Congruent Contraceptive Use, Sexuality, and Satisfaction**

Studies based on the congruency hypothesis reported associations between incongruent use of contraception and decreased perceived partner attractiveness (Roberts et al., 2012; Roberts, Cobey, et al., 2014), decreased relationship satisfaction (Roberts, Cobey, et al., 2014; but see French & Meltzer, 2020; Roberts, Little, et al., 2014), decreased sexual satisfaction (French & Meltzer, 2020; Roberts et al., 2012; Roberts, Little, et al., 2014; Russell et al., 2014), increased satisfaction with partner's paternal provision (Roberts et al., 2012), increased jealousy (Cobey et al., 2015), and increased appeal of alternative mates (Birnbaum et al., 2019). While many studies reported evidence for the congruency hypothesis, a recent large-scale replication study \(N = 948\) with balanced congruent and incongruent user groups found no effect of incongruent contraceptive use (Jern et al., 2018). This study, along with the many nonreplications and methodological criticisms concerning cycle preference shifts, undermine the foundations of the congruency hypothesis. Given this uncertainty, we aimed to
test the replicability of effects of congruent contraceptive use on perceived partner attractiveness, relationship satisfaction, and sexual satisfaction, as well as to expand the analyses to other outcomes including libido, frequency of vaginal intercourse, and frequency of masturbation.

**The Current Study**

The three main goals of this study were to (1) investigate selection effects on choice of contraceptive method (no/nonhormonal vs. hormonal) and congruent contraceptive use (incongruent vs. congruent); (2) estimate effects of choice of contraceptive method and congruent contraceptive use on perceived partner attractiveness, relationship satisfaction, sexual satisfaction, libido, frequency of vaginal intercourse, and frequency of masturbation after taking into account observed confounders; and (3) estimate the robustness of these effects in the light of potential unobserved confounders. We used data from 1,179 women from the Goettingen Ovulatory Cycle Diaries 2 and compared outcomes between groups of women using hormonal contraceptives (HC users) and women using no or nonhormonal contraceptives (non-HC users).

**Hypotheses**

The analyses in this study were not preregistered. The data used in this study had already been collected, accessed, and analyzed to address other research questions. The first author formulated the hypotheses without prior access to the data and before any analyses relevant to this study had been conducted.

**Selection Effects.** Women do not choose contraceptives at random. Any study investigating the causal effects of hormonal contraception must therefore take selection into account. Our first goal was to investigate the degree to which contraceptive use and congruent contraceptive use could be explained by selection effects of demographic variables (age, education, and income), personality traits (openness, conscientiousness, extraversion, agreeableness, neuroticism, religiosity), and relationship duration. **Figure 1** shows a directed acyclic graph (e.g., Pearl, 1995) of the suggested causal network. Directed acyclic graphs visually represent causal assumptions. They offer an intuitive approach for thinking about causal structures and help to answer questions about potential third variables (i.e., confounders). Even though they look a lot like structural equation models, directed acyclic graphs differ in two important points: While structural equations models allow for bidirectional relationships between variables, directed acyclic graphs include only one-headed arrows and whereas structural equation models assume linear, additive relationships (unless indicated otherwise) the arrows included in directed acyclic graphs might reflect any form of relationship. For a primer on directed acyclic graphs see Rohrer (2018).

Many previous studies did not control for any selection variables; others controlled for age, income, and relationship duration (e.g., Jern et al., 2018). Yet these three covariates seem insufficient given the presumably complex causal network underlying the relationships we aim to study. We therefore compared the simple model, adjusted only for age, income, and relationship duration, to a more complex model adjusted for further potential selection effects as well: Education, Big Five personality traits, and religiosity (Figure 1). Several other variables that might play a role as selection effects (information about sociosexuality, how happy women would be about an unplanned pregnancy, number of sexual partners, or number of days and nights spent with their partner per week) were considered but ultimately not adjusted in order to avoid controlling for potential colliders or mediators (Rohrer, 2018). A collider for a certain pair of variables is any variable that is causally influenced by both of them. For example, whether women would be happy if they found out that they are pregnant could be influenced by both of them. For example, whether women would be happy if they found out that they are pregnant could be influenced by the choice of contraception (e.g., through a hormonal effect). Even more likely, it could be influenced by the outcome relationship satisfaction. Controlling for a collider can potentially introduce a spurious (i.e., noncausal) association between its causes. An example of a mediator is

---

3 Most of the available literature focused only on the pill, and the following hypotheses were based on this literature. Some evidence suggests divergent effects of different forms of hormonal contraceptives on sexual desire (Boozalis et al., 2016; Sabatini & Cagiano, 2006), and some studies reported differences in outcomes depending on dosages of estrogen and progestin in the pill (Caruso et al., 2011; Graham et al., 1995; Greco et al., 2017; Kelly et al., 2010; Krzyzulec & Droszdol, 2008; Strufaldl et al., 2010; but see M. Wallwiener et al., 2010). However, the current study was interested in the effects all hormonal contraceptives had in common. In addition, the more general comparison of hormonal contraceptives with no/nonhormonal contraceptives enabled us to include more participants for analyses. Results from robustness analyses including only pill users compared to naturally cycling women are available on the supplementary website https://laurabotzet.github.io/effects_of_contraception/14_analyses_robust. The results did not differ substantially from the results based on our main analyses including all hormonal contraceptive users.
sociosexual desire, which could potentially mediate the effect of hormonal contraceptives on libido. Controlling for a mediator might lead to controlling for the very process of interest (Rohrer, 2018). Therefore, we excluded all possible selection effects that might be influenced by the outcomes or might mediate the effect of hormonal contraceptives on the proposed outcomes—with the exception of relationship duration. Although relationship duration might represent a potential collider, the literature suggests it is uniquely important as a selection variable (e.g., Jern et al., 2018).

**Hypothesis 1:** The complex model including all potential selection variables explains more variance than does the simple model including only age, income, and relationship duration in (1) the choice of contraception and (2) congruent contraceptive use.

### Effects of Hormonal Contraceptives.

Our second goal was divided into two parts. First, we aimed to estimate effects of choice of contraceptive method after adjusting for observed confounders. Because RCTs provide the most robust evidence regarding causal psychological effects of hormonal contraceptives, the following hypotheses on effects of current contraceptive use were based on findings from RCTs, which suggest negative psychological effects of hormonal contraceptives (Graham et al., 1995; Zethraeus et al., 2016, 2017).

**Hypothesis 2.1:** Hormonal contraceptives lead to decreased (1) perceived partner attractiveness, (2) relationship satisfaction, (3) sexual satisfaction, (4) libido, (5) frequency of vaginal intercourse, and (6) frequency of masturbation.

We also tested whether effects of hormonal contraceptives persisted after controlling for selection variables (age, education, income, openness, conscientiousness, extraversion, agreeableness, neuroticism, religiosity, and relationship duration).

**Hypothesis 2.2:** After controlling for all selection variables, hormonal contraceptives lead to decreased (1) perceived partner attractiveness, (2) relationship satisfaction, (3) sexual satisfaction, (4) libido, (5) frequency of vaginal intercourse, and (6) frequency of masturbation.

### Effects of Congruent Contraceptive Use.

The second part of this goal was to estimate effects of congruent contraceptive use after considering observed confounders. In line with the congruency hypothesis (Roberts et al., 2013), it was expected that congruent contraceptive use leads to positive effects on relationship quality and sexual functioning.

**Hypothesis 3.1:** Congruent contraceptive use leads to increased (1) perceived partner attractiveness, (2) relationship satisfaction, (3) sexual satisfaction, (4) libido, (5) frequency of vaginal intercourse, and (6) frequency of masturbation.

We also tested whether effects of congruent contraceptive use persisted after controlling for selection variables (age, education, income, openness, conscientiousness, extraversion, agreeableness, neuroticism, religiosity, and relationship duration).

**Hypothesis 3.2:** After controlling for all selection variables, congruent contraceptive use leads to increased (1) perceived partner attractiveness, (2) relationship satisfaction, (3) sexual satisfaction, (4) libido, (5) frequency of vaginal intercourse, and (6) frequency of masturbation.

### Sensitivity Analyses

Our third goal was to estimate the robustness of effects of both contraceptive use and congruent contraceptive use in the light of potential unobserved confounders. When making causal inferences based on observational data, most researchers adjust for observed covariates while implicitly assuming that there are no unobserved confounders. This assumption is untestable and, given the research question at hand, unlikely to hold. One way to solve this problem is to qualitatively debate potential unobserved confounders and their effects. Another possibility, which we employ here, is to apply sensitivity analyses to quantitatively examine the fragility of putative causal estimates when the underlying assumption of no unobserved confounding is challenged.

We estimated the sensitivity of effects of hormonal contraceptives and congruent contraceptive use on the outcomes in the light of potential unobserved confounders. Given the complexity of the relationships between contraceptive method, relationship quality, and sexual functioning, it is very likely that confounders exist that were not observed in the current study and therefore not included in the complex models. Unobserved confounding could explain or potentially even reverse the observed effects of hormonal contraceptives.

### Methods

#### Sample and Procedure

This study was based on the data collection *Goettingen Ovulatory Cycle Diaries 2*, conducted from May 2016 to January 2017. The data was primarily collected to investigate psychological cycle shifts in women (Arslan et al., 2016). Arslan, Reitz, et al. (2020) published a paper based on the data collection focusing on measurement arcana investigating the benefits of a planned missingness design in diaries and Schleifenbaum et al. (2021) studied fertile window effects on attractiveness on a within-subject level using hormonal contraceptive users as a quasi-control group who do not experience ovulation. A codebook for the full dataset including all measures is available at [https://ruber-narslan.github.io/gocd2](https://ruber-narslan.github.io/gocd2) (Arslan, Driebe, et al., 2020).

All data was collected online using the open source software [formr.org](http://formr.org) (Arslan, Walther, et al., 2020). In an initial survey, participants answered questions about contraceptive methods, demographics, sexuality, and personality. All variables except for libido, frequency of vaginal intercourse, and frequency of masturbation were derived from the initial survey. The survey was followed by a daily 5-minute diary filled out for 70 days. Libido, frequency of vaginal intercourse, and frequency of masturbation were based on the diary survey.

A total of 1,660 people initially enrolled in the study. Only women were allowed to participate. The proportion of
Table 1. Exclusion Criteria, Reasons for Exclusion, and Number of Excluded Participants

<table>
<thead>
<tr>
<th>Exclusion criteria</th>
<th>Reasons for exclusion</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not finishing the initial survey</td>
<td>Missing data</td>
<td>249</td>
</tr>
<tr>
<td>Not biologically female</td>
<td>Potential hormonal influences; no need to use contraceptives to prevent pregnancy</td>
<td>3</td>
</tr>
<tr>
<td>Not predominantly heterosexual</td>
<td>No need to use contraceptives to prevent pregnancy</td>
<td>26</td>
</tr>
<tr>
<td>Currently in a homosexual romantic relationship</td>
<td>No need to use contraceptives to prevent pregnancy</td>
<td>9</td>
</tr>
<tr>
<td>Older than 50</td>
<td>Potential hormonal influences</td>
<td>35</td>
</tr>
<tr>
<td>(Post-)menopausal</td>
<td>Potential hormonal influences</td>
<td>41</td>
</tr>
<tr>
<td>Pregnant</td>
<td>Potential hormonal influences</td>
<td>23</td>
</tr>
<tr>
<td>Breastfeeding</td>
<td>Potential hormonal influences</td>
<td>28</td>
</tr>
<tr>
<td>Trying to become pregnant</td>
<td>No need to use contraceptives to prevent pregnancy</td>
<td>61</td>
</tr>
<tr>
<td>“Taking a chance” of becoming pregnant</td>
<td>Need to use contraceptives to prevent a pregnancy is low</td>
<td>41</td>
</tr>
<tr>
<td>Using no contraceptive methods for other reasons</td>
<td>No information about contraceptive method</td>
<td>55</td>
</tr>
<tr>
<td>Choice of contraceptive method(^a)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morning-after pill</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breastfeeding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am infertile</td>
<td>Potential hormonal influences and/or no need for hormonal contraception</td>
<td>2</td>
</tr>
<tr>
<td>My partner is infertile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am sterilized</td>
<td></td>
<td></td>
</tr>
<tr>
<td>My partner is sterilized</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other contraceptive method</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incongruent information about contraceptive method(^b)</td>
<td>Potential hormonal influences</td>
<td>39</td>
</tr>
<tr>
<td>Medication including sex hormones</td>
<td>Potential hormonal influences</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>481</td>
</tr>
</tbody>
</table>

Note. \(^a\) Numbers of contraceptive methods add up to 35 (rather than 34) because women could report several hormonal contraceptives that led to exclusion. \(^b\) Participants, who reported recently changing their contraceptive method (in the last three months), were excluded because this might influence their hormonal status when participating in our study.

Hormonal Contraception and Sexuality: Causal Effects, Unobserved Selection, or Reverse Causality?

Exclusion Criteria

The exclusion criteria are summarized in Table 1. Note that some women were excluded for multiple reasons (e.g., not finishing the initial survey and being pregnant). Figure 2 shows a flowchart of applied exclusion criteria. In total, 481 participants were excluded, leaving 1,179 participants who were included in analyses.

For analyses that were based on diary information (libido, frequency of vaginal intercourse, and frequency of masturbation) 53,332 diary days for 1,179 participants were available. A total of 745 diary days were excluded because participants skipped these days and no information was available. An additional 142 diary days were excluded because participants indicated giving dishonest answers (e.g., randomly giving answers to speed through the survey). This resulted in 52,445 diary days for 1,138 participants. In addition, all participants who filled out fewer than 14 diary days were excluded in order to calculate reliable estimates for libido, frequency of vaginal intercourse, and frequency of masturbation \((n_{\text{excluded participants}} = 170, n_{\text{excluded days}} = 1,014)\). In total, 968 participants with 51,431 diary days and a mean number of 53.13 days per participant were included for the analyses with libido as an outcome.

For analyses with frequency of vaginal intercourse and frequency of masturbation as outcomes, we omitted women who said they did not need contraception because they currently had no sexual intercourse \((n = 84\) of all 1,179 included participants). Including these women among non-HC users would clearly introduce reverse causality, as their frequency of vaginal intercourse determined their need for contraception. This exclusion affected 71 women in the analyses for effects of hormonal contraceptives on frequency of vaginal intercourse and frequency of masturbation and 10 women in the analyses for effects of congruent contraceptive use and its interaction with current contraceptive use on frequency of vaginal intercourse and frequency of masturbation.
Figure 2. Flowchart of Applied Exclusion Criteria

Note. Numbers of choice of contraceptive methods add up to 30 (rather than 29) because women could report several hormonal contraceptives that led to exclusion.

Participants

The 1,179 eligible participants were on average 25.0 years old (SD = 5.1, range: 18-49 years). Most of the participants were students (72%), 22% were working, 3% were in secondary or vocational school, and 3% were homemakers or not working. A majority reported their religious denomination as Christian (51%), 42% described themselves as nonreligious, and 7% reported other religious denominations (including Buddhism, Hinduism, Islam, and Judaism). Most (66%) of the participants were in an ongoing romantic relationship (6% of all participants were married and 2% were engaged), with an average relationship duration of 3.4 years (SD = 3.7, range: 0.0-29.4 years). The vast majority (94%) had no children. Geographically, only Göttingen (the university town where this study was conducted) was visibly overrepresented.

Included Compared to Excluded Participants. Table S1 in the Supplemental Material shows all comparisons between included and excluded participants who finished the initial survey. Unsurprisingly given our exclusion of older, perimenopausal, and postmenopausal women, included participants were younger ($d = -1.43$ 95% CI: [-1.62; -1.24]), earned less money (Cramér's $V = 0.30$ [0.25; 0.37]), were more often single (Cramér's $V = 0.07$ [0.02; 0.12]), reported shorter relationship durations ($d = -0.87$ [-1.06; -0.68]), and reported higher libido ($d = 0.25$ [0.09; 0.42]). Included and excluded participants did not differ significantly in education, extraversion, neuroticism, agreeableness, conscientiousness, openness, religiosity, perceived partner attractiveness, relationship satisfaction, sexual satisfaction, frequency of vaginal intercourse, or frequency of masturbation ($all ps > .05$).

Measures and Indices

All variables were based on self-report. Item wordings are listed in Table S2 in the Supplemental Material. The following sections summarize measures and indices.

Selection Variables

Participants were asked to report their age and education in years and to indicate their income, choosing between five income groups (<500€ per month; 500–1,000€ per month;
1,000–2,000€ per month; 2,000–3,000€ per month; >3,000€ per month) and a sixth group in case they did not want to answer this question (do not want to disclose). Big Five personality traits—including extraversion (8 items), neuroticism (8 items), agreeableness (9 items), conscientiousness (9 items), and openness (10 items)—were measured using the German version of the Big Five Inventory (44 items, Lang et al., 2001) on a scale from 1 (does not apply at all) to 5 (fully applies). Level of religiosity was measured on the same scale with one item.

Relationship duration was measured in years and months. In order to incorporate single women into the analyses including relationship duration and to be able to estimate nonlinear effects of relationship duration, the variable was split into five categories: 0–12 months (n = 198), 13–28 months (n = 198), 29–52 months (n = 188), and more than 52 months (n = 190), and not in a relationship (n = 405). Figure S1 in the Supplemental Material displays a histogram of relationship duration.

**Contraception**

Current contraception was measured with one item asking for current contraceptive methods, and answers were categorized (hormonal contraception, no/nonhormonal contraception). Congruent use of hormonal contraception was determined by comparing women’s current use of hormonal contraceptives and their use of hormonal contraceptives at the time of meeting their partner.

Participants were sorted into two large groups according to current contraception (non-HC users: n = 688, 58%; HC users: n = 491, 42%; see Figure S2 in the Supplemental Material for allocation process and number of women in each group). No/nonhormonal contraception included nonhormonal intrauterine devices (n = 85, 7%), fertility awareness methods (n = 120, 10%), condoms (n = 580, 32%), and no contraception at all (n = 89, 8%), as well as miscellaneous other barrier-based methods (n = 14, 1%). Hormonal contraception included only the pill (n = 251, 21%), only other hormonal contraceptives (n = 61, 5%), and any combination of the pill and nonhormonal contraceptives (n = 161, 14%), other hormonal contraceptives and nonhormonal contraceptives (n = 17, 1%) as well as one combination of the pill, other hormonal contraceptives and nonhormonal contraceptives (n = 1, 0%). From all women using oral contraceptives (n = 413, 55%) only 25 women (2%) used progesterin-only oral contraceptives (i.e., "minipill"). Information about estrogen dosage was available for 377 women using regular oral hormonal contraceptives (M = 27.54µg; SD = 7.24µg; min = 2µg; max = 100µg) and 352 of all women using oral hormonal contraceptives reported the type of gestagen (Dienogest: n = 109, 31%; Levonorgestrel: n = 105, 30%; Chlormadinonacetat: n = 62, 18%; Desogestrel: n = 38, 11%; Drospirenon: n = 17, 5%; Cyprogenacetat: n = 14, 4%; Nomegestrolacetat: n = 5, 1%; Norgestimat: n = 2, 1%).

Of the 774 participants in an ongoing romantic relationship, 491 (63%) used congruent contraceptive methods (congruent HC users: n = 240, 31%; congruent non-HC users: n = 251, 32%). The remaining 283 (37%) used incongruent contraceptive methods (HC users → non-HC users: n = 150, 19%; non-HC users → HC users: n = 135, 17%).

**Outcomes**

Perceived partner attractiveness (two items; one measuring facial attractiveness and one on body attractiveness), relationship satisfaction (five items measuring aspects including satisfaction, fulfilling of needs, and—reverse-scored—conflicts), and sexual satisfaction (one item) were reported by partnered participants in the initial survey (1 = does not apply at all, 5 = fully applies). All participants reported on libido and sexual frequency in the diary. Libido was measured with one item every day (0 = not at all, 4 = very much) and the mean libido was aggregated across all diary days. Sexual frequency was measured with one item every day—the proportion of sexually active days was calculated by summing up all sexually active days and dividing them by the number of days the diary was filled out. Two measures of sexual frequency were included: (1) Frequency of vaginal intercourse included only the proportion of days when participants indicated that the sexual activity involved penetrative intercourse (not including anal sex); (2) frequency of masturbation included only the proportion of days when participants indicated that the sexual activity involved masturbation. Diary outcomes were all aggregated to reduce complexity and because we were not interested in any predictors at the within-woman level.

**Statistical Analyses**

All analyses were conducted using the statistical software R (R Core Team, 2013). Bayesian analyses were performed using the R-package brms (Bürkner, 2017), which implements an R interface to the probabilistic programming language Stan (Carpenter et al., 2017). Due to a lack of research especially on some of the outcomes (e.g., masturbation) and the fact that we switched between Bayesian and frequentist approaches for sensitivity analyses improper flat priors (which ensure consistency with maximum likelihood) for all parameters were used.

The focus of the current study was not to decide on accepting or rejecting null hypotheses but rather on estimating the strength of associations between contraception and outcomes. Nevertheless, a decision rule about whether an effect had a substantial practical impact was implemented in order to give recommendations for future research (Makowski et al., 2019). For each linear estimated effect size, a region of practical equivalence (ROPE) around the null value was estimated and combined with the 90% highest density interval (HDI). Following Kruschke (2018), the ROPE for each normally distributed outcome was set as

\[
0 \pm 0.1 \times SD(\text{outcome})
\]

A 90% HDI was chosen because a large number of sampling iterations would have been necessary to estimate the regions outside of a 95% HDI. In addition, applying a ROPE criterion is a fairly conservative way to decide about effect sizes; we therefore used a lower level of
The decision to follow Kruschke's (2018) suggestion reflects a lack of strong prior research that could inform our notion of practical equivalence, especially given widespread heterogeneity in outcome measures.

For non-normally distributed outcomes (frequency of vaginal intercourse and frequency of masturbation) the ROPE was set at $0 \pm 0.05$ because analyses assumed Poisson distributions, and the effects were estimated on a logarithmic scale. A difference of 0.05 on the log scale approximately translates to a difference of 1 percentage point in frequency of vaginal intercourse for the rates found in this sample.

Using the test for practical equivalence (Kruschke, 2018) makes it possible to distinguish between three eventualities: (a) rejecting the null hypothesis: The estimated effect sizes are interpreted as substantial because the HDI is outside the ROPE, (b) accepting the null hypothesis: The estimated HDI is completely within the ROPE, so it is unknown whether the association is outside the ROPE—in other words, the estimates are insufficiently precise and future research with larger samples is needed.

**Selection Effects on Current Contraceptive Use and Congruent Contraceptive Use**

In our investigation of the degree to which contraceptive use and congruent contraceptive use can be explained by selection effects, the simple models were based on probit regressions using the predictors age, income, and relationship duration. The complex models were based on probit regressions that also included education, openness, conscientiousness, extraversion, agreeableness, neuroticism, and religiosity as predictors. In order to investigate potential selection effects in light of the direction of incongruent contraceptive use (switching from hormonal contraceptives to no/nonhormonal contraceptives or vice versa) a third model with congruent contraceptive use as an outcome and additional predictors including contraceptive method when meeting one’s partner and its interaction with all other predictors was analyzed.

The models were compared by using approximate leave-one-out cross-validation (LOO-IC; Vehtari et al., 2016) to investigate whether the complex model explained current contraceptive use/congruent contraceptive use more precisely than the simple model did. Substantially better model performance was indicated if the absolute difference in expected log pointwise predictive density ($\text{diff}_{\text{ELPD}}$) was higher than twice the standard error of expected log pointwise predictive density ($\text{SE}_{\text{diff}(\text{ELPD})}$).

**Effects of Hormonal Contraceptives**

To estimate effects of choice of contraceptive method (no/nonhormonal vs. hormonal contraceptives and congruency) on relationship quality and sexual functioning, we used linear regressions for the outcomes perceived partner attractiveness, relationship satisfaction, sexual satisfaction, and libido. For the outcomes frequency of vaginal intercourse and frequency of masturbation we used Poisson regression and included offsets for the number of days the diary was filled out. We always estimated an unadjusted model including only current use of contraception (no/non-hormonal vs. hormonal contraceptives). In a second model, we then adjusted for potential selection variables: age, income, relationship duration, education, openness, conscientiousness, extraversion, agreeableness, neuroticism, and religiosity. To study congruent contraceptive use, models were specified as above but additionally included congruency and its interaction with current contraceptive use.

For effects of hormonal contraceptives our theoretical estimand of interest (Lundberg et al., 2021) was the average treatment effect of hormonal contraception on sexual satisfaction, frequency, and so on. We strove to identify this causal effect by adjusting for confounding variables. The so-estimated effect size was our empirical estimand. For linear regression models, the effect size of interest was the estimated coefficient. For Poisson regression models, we used the difference in percentage frequencies across the diary. Therefore, we based our ROPE for these outcomes on percentage frequencies and reported differences in percentage frequencies in addition to estimated effect sizes.

As robustness analyses, we performed all analyses of effects of hormonal contraceptives including only pill users compared to naturally cycling women. The results are reported on the supportive website [https://lau-rabotzet.github.io/effects_of_contraception/14 analyses robust](https://laurabotzet.github.io/effects_of_contraception/14 analyses robust). The patterns between oral contraceptive users and naturally cycling women did not differ substantially from the patterns reported in our main analysis including all forms of hormonal contraceptives.

**Sensitivity Analyses**

To estimate the robustness of effects of contraceptive use and congruent contraceptive use in the light of potential unobserved confounders, we used a sensitivity analysis that is an extended version of the omitted variable bias framework developed by Cinelli & Hazlett (2020). It estimates how robust results are to all potential unobserved confounders and how strong unobserved confounding would need to be in relation to the strength of observed confounders to change conclusions in a substantive manner. While sensitivity analysis cannot compensate unobserved confounders, it quantifies what one needs to believe in order to sustain that a given conclusion is not due to confounding.

The R-package [sensemakr v0.1.3](https://CRAN.R-project.org/package=sensemakr) was used for analysis. As this package does not implement sensitivity analyses for Bayesian analyses and only targets linear regressions, we performed frequentist linear regressions for all outcomes. For our scenario, linear regression is robust to this violation of assumptions and the effects closely approximated the conditional effects in Bayesian Poisson
Hormonal Contraception and Sexuality: Causal Effects, Unobserved Selection, or Reverse Causality?

Table 2. Counts, Means, Standard Deviations, Ranges, and Reliability

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>min</th>
<th>max</th>
<th>α</th>
<th>ω</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Age (in years)</td>
<td>1,179</td>
<td>25.03</td>
<td>5.09</td>
<td>18</td>
<td>49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Education (in years)</td>
<td>1,179</td>
<td>15.07</td>
<td>4.73</td>
<td>0</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Income (monthly)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- &lt;500€</td>
<td>287</td>
<td>4.27</td>
<td>0.74</td>
<td>1.2</td>
<td>5.0</td>
<td>.87</td>
<td>.77</td>
</tr>
<tr>
<td>- 500€–1,000€</td>
<td>565</td>
<td>3.83</td>
<td>0.79</td>
<td>1.0</td>
<td>5.0</td>
<td>.85</td>
<td>.71</td>
</tr>
<tr>
<td>- 1,000€–2,000€</td>
<td>215</td>
<td>3.67</td>
<td>0.62</td>
<td>1.4</td>
<td>5.0</td>
<td>.76</td>
<td>.68</td>
</tr>
<tr>
<td>- 2,000€–3,000€</td>
<td>63</td>
<td>3.62</td>
<td>0.89</td>
<td>1.2</td>
<td>5.0</td>
<td>.82</td>
<td>.70</td>
</tr>
<tr>
<td>- &gt;3,000€</td>
<td>16</td>
<td>3.35</td>
<td>0.61</td>
<td>1.5</td>
<td>5.0</td>
<td>.81</td>
<td>.63</td>
</tr>
<tr>
<td>- do not want to disclose</td>
<td>33</td>
<td>3.33</td>
<td>0.63</td>
<td>1.5</td>
<td>5.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Extraversion</td>
<td>1,179</td>
<td>3.46</td>
<td>0.78</td>
<td>1.12</td>
<td>5.0</td>
<td>.87</td>
<td>.77</td>
</tr>
<tr>
<td>(5) Neurotician</td>
<td>1,179</td>
<td>3.00</td>
<td>0.78</td>
<td>1.00</td>
<td>5.0</td>
<td>.85</td>
<td>.71</td>
</tr>
<tr>
<td>(6) Agreeableness</td>
<td>1,179</td>
<td>3.68</td>
<td>0.62</td>
<td>1.44</td>
<td>5.0</td>
<td>.76</td>
<td>.67</td>
</tr>
<tr>
<td>(7) Conscientiousness</td>
<td>1,179</td>
<td>3.53</td>
<td>0.66</td>
<td>1.56</td>
<td>5.0</td>
<td>.82</td>
<td>.70</td>
</tr>
<tr>
<td>(8) Openness</td>
<td>1,179</td>
<td>3.78</td>
<td>0.61</td>
<td>1.50</td>
<td>5.0</td>
<td>.81</td>
<td>.63</td>
</tr>
<tr>
<td>(9) Religiosity</td>
<td>1,179</td>
<td>2.20</td>
<td>1.34</td>
<td>1.00</td>
<td>6.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10) Relationship duration</td>
<td>405</td>
<td>6.96</td>
<td>7.21</td>
<td>4.73</td>
<td>5.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Single</td>
<td>198</td>
<td>6.96</td>
<td>7.21</td>
<td>4.73</td>
<td>5.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- 0–12 months</td>
<td>198</td>
<td>6.96</td>
<td>7.21</td>
<td>4.73</td>
<td>5.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- 13–28 months</td>
<td>188</td>
<td>6.96</td>
<td>7.21</td>
<td>4.73</td>
<td>5.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- 29–52 months</td>
<td>190</td>
<td>6.96</td>
<td>7.21</td>
<td>4.73</td>
<td>5.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11) Perceived partner attractiveness</td>
<td>774</td>
<td>4.25</td>
<td>0.74</td>
<td>1.00</td>
<td>5.0</td>
<td>.68</td>
<td>.68</td>
</tr>
<tr>
<td>(12) Relationship satisfaction</td>
<td>774</td>
<td>3.39</td>
<td>0.43</td>
<td>1.40</td>
<td>4.6</td>
<td>.88</td>
<td>.83</td>
</tr>
<tr>
<td>(13) Sexual satisfaction</td>
<td>774</td>
<td>4.00</td>
<td>1.05</td>
<td>1.00</td>
<td>5.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(14) Libido</td>
<td>968</td>
<td>1.19</td>
<td>0.59</td>
<td>0.00</td>
<td>3.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(15) Frequency of vaginal intercourse</td>
<td>897</td>
<td>7.27</td>
<td>7.19</td>
<td>0.00</td>
<td>42.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(16) Frequency of masturbation</td>
<td>897</td>
<td>6.96</td>
<td>7.21</td>
<td>0.00</td>
<td>50.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Means, standard deviations, ranges, and (if applicable) reliability measurements for variables and scales of all selection variables and outcomes are reported for numerical variables. For categorical variables (income and relationship duration), only the count for each group is displayed.

α = Cronbach’s alpha (Cronbach, 1951); ω = McDonald’s omega hierarchical (McDonald, 1999). * McDonald’s omega for perceived partner attractiveness could not be computed.

regressions.

Availability of Data, Code, and Analyses

All code (for both data wrangling and analysis), materials, and full statistical results pertaining to this study are openly available on the supportive website (https://laurabotzet.github.io/effects_of_contraception; Botzet, 2020) and uploaded as part of the accompanying project on the Open Science Framework (https://osf.io/rqxsa/). A codebook for the full dataset for the Goettingen Ovulatory Cycle Diaries 2 is available at https://rubenarslan.github.io/goecd2 (Arslan, Driebe, et al., 2020). Because we cannot share the data publicly due to the sensitive nature of sexual diary studies, we uploaded a synthetic dataset to the Open Science Framework following Quintana’s (2020) primer on synthetic datasets using the R-package synthpop (Nowok et al., 2016). The synthetic dataset mimics many of the central features of the real data, including means and bivariate associations (see https://laurabotzet.github.io/effects_of_contraception/11_check_synthetic_data.html for comparisons between the real and the synthetic dataset), and can thus be used by others to write code to test and build models using realistic fake data. Upon request we can share the partially anonymized data with anyone who has a valid reason and agrees not to attempt to reidentify the data.

Results

Selection Variables and Outcomes

Counts, means, standard deviations, ranges, and reliability measurements of selection variables and outcomes are reported in Table 2. Reliability measurements included Cronbach’s alpha and McDonald’s omega (hierarchical) and both indicated sufficient reliability for the used measurements. Table S3 in the Supplemental Material summarizes means and standard deviations for all selection variables and outcomes separately for singles and partnered women divided by current contraceptive method (no/nonhormonal vs. hormonal) and congruency of contraceptive method (incongruent vs. congruent).

Table S4 in the Supplemental Material shows zero-order correlations of all numerical selection variables and out-
comes. The outcomes perceived partner attractiveness, relationship satisfaction, sexual satisfaction, libido, and frequency of vaginal intercourse all correlated positively with each other (r between .09 and .41) except for the relationship between satisfaction and libido, which showed no significant correlation. Frequency of masturbation correlated positively with libido (r = .22) and negatively with perceived partner attractiveness, relationship satisfaction, and sexual satisfaction (r between -.11 and -.10). We found no significant correlation between frequency of vaginal intercourse and frequency of masturbation (r = -.05).

For noncontinuous variables (income and relationship duration), we do not report correlations; rather, we analyzed linear regressions with the linear variable as an outcome and the noncontinuous variable as a predictor. For the outcomes frequency of vaginal intercourse and frequency of masturbation as outcomes, generalized linear models based on Poisson distributions were analyzed. Results from these analyses are summarized in the Supplemental Material.

Selection Effects

According to LOO-IC, the more complex model did not substantively improve upon the simple model when predicting current contraceptive method, changes in contraceptive method, or changes in contraceptive method separately for non-HC/HC users. When predicting current contraceptive method, the complex model showed only slightly improved LOO-IC performance (within one standard error of the difference). Unstandardized effect size estimates for the different predictors for models with current contraceptive method as an outcome are displayed in Figure 3. Descriptively, the simple model performed even better according to LOO-IC for the other two outcomes (see Table S5 in the Supplemental Material). Unstandardized effect size estimates for the different predictors for models with congruent contraceptive method as an outcome—separately for women who were using hormonal contraceptives when they met their partner (left) and women who were not (right)—are displayed in Figure 4.

Effects of Hormonal Contraceptives

Table 3 summarizes unstandardized effect size estimates and 90% HDIs for hormonal contraceptives on perceived partner attractiveness, relationship satisfaction, sexual satisfaction, libido, frequency of vaginal intercourse, and frequency of masturbation. Figure 5 displays all unstandardized effect size estimates and 90% HDIs for hormonal contraceptives compared to the ROPE criterion. Unstandardized effect size estimates for hormonal contraceptives in the uncontrolled models as well as the controlled models overlapped with the ROPE for perceived partner attractiveness, relationship satisfaction, sexual satisfaction, and libido. The unstandardized effect size estimates for hormonal contraceptives were outside of the ROPE for frequency of vaginal intercourse (uncontrolled: 0.25 [90% HDI: 0.21; 0.29]; controlled: 0.16 [0.12; 0.21]) and frequency of masturbation (uncontrolled: -0.40 [-0.44; -0.55], controlled: -0.30 [-0.34; -0.25]).

Based on the uncontrolled models this corresponded to a difference of 3.5 [3.0; 4.1] percentage points in estimated probabilities of penetrative intercourse per day (HC users: 15.9%; non-HC users: 12.4%) and a difference of 5.0 [4.5; 5.5] percentage points in estimated probabilities of masturbation per day (HC users: 10.5%; non-HC users: 15.3%). For the controlled models we computed average marginal effects (documented at https://laurabotzet.github.io/effects_of_contraception/18_marginal_effects.html) by setting number of diary days filled out as one, using all levels of categorical variables (income, relationship duration), using five categorical levels for the continuous variables age and years of education and assuming empirical mean values for all other continuous variables (Big Five personality traits, religiosity). Differences in average marginal effects were 2.8 [2.0; 3.6] percentage points in estimated probabilities of penetrative intercourse per day (HC users: 18.0%; non-HC users: 15.2%) and 3.1 [2.7; 3.6] percentage points in estimated probabilities of masturbation per day (HC users: 9.0%; non-HC users: 12.2%). Therefore, the observed differences were all larger than the difference in 1 percentage point that our predefined ROPE for analyses assuming Poisson distributions was based on.

Effects of Congruent Contraceptive Use

Table 4 summarizes unstandardized effect size estimates for hormonal contraceptives, congruent contraceptive use, and their interaction on perceived partner attractiveness, relationship satisfaction, sexual satisfaction, libido, frequency of vaginal intercourse, and frequency of masturbation. To illustrate effects, Figures 6 and 7 show the predicted means for current contraceptive method and congruent contraceptive use based on the uncontrolled and the controlled model, respectively. Unstandardized effect size estimates for hormonal contraceptives, congruent contraceptive use, and the interaction of the two in the uncontrolled models and the controlled models overlapped with the ROPE for perceived partner attractiveness, relationship satisfaction, sexual satisfaction, and libido. The unstandardized effect size estimates for congruent contraceptive use and its interaction with current use of HCs overlapped with the ROPE for frequency of vaginal intercourse and frequency of masturbation. The unstandardized estimates for the main effects of HCs on frequency of vaginal intercourse and masturbation were outside of the ROPE, as in the preceding analysis.

Sensitivity Analyses

Results for sensitivity analyses are summarized in Table 5 for models including current use of hormonal contraceptives (without and with control for observed confounders) and in Table 6 for models additionally including congruent contraceptive use and its interaction with current use of hormonal contraceptives (without and with control for observed confounders). Results are based on frequentist linear models for all outcomes.

Perceived Partner Attractiveness, Relationship Satisfaction, Sexual Satisfaction, and Libido

Hormonal contraceptives, congruent contraceptive use, and their interaction showed no significant effect on per-
Figure 3. Unstandardized Effect Size Estimates and 90% HDI for Models with Current Contraceptive Use (0 = No/Nonhormonal Contraceptives; 1 = Hormonal Contraceptives) as Outcome

Note. Simple model (top) includes age, income, and relationship duration as predictors. Complex model (bottom) also includes education, extraversion, neuroticism, agreeableness, conscientiousness, openness, and religiosity. Error bars show 90% HDIs.

HDI = highest density interval.

Received partner attractiveness, relationship satisfaction, sexual satisfaction, and libido in the uncontrolled or the controlled model. Unobserved confounders would have to explain between $RV_{\theta} = 0.2\%$ and $RV_{\theta} = 6.3\%$ of the residual variance of both the treatment and the outcome to bring effects of hormonal contraceptives, congruent contraceptive use, and their interaction on these outcomes to zero (see Tables 5 and 6 for further information).

**Frequency of Vaginal Intercourse and Masturbation**

Table 5 provides information about the sensitivity analysis for the effect of hormonal contraceptives on the frequency of vaginal intercourse and masturbation. Without control for observed confounders the effect of hormonal contraceptives on frequency of vaginal intercourse was statistically significant with an unstandardized effect size estimate of 0.04 [95% CI: 0.02; 0.05] (Bayesian analysis: 0.25 [90% HDI: 0.21; 0.29]). Adjusting for observed confounders reduced this effect by 50% (30% based on Bayesian analysis), resulting in an unstandardized effect size estimate of 0.02 [95% CI: 0.01; 0.04] (Bayesian analysis: 0.16 [90% HDI: 0.12; 0.21]). To bring the point estimate for hormonal contraceptives on frequency of vaginal intercourse to zero, unobserved confounders would have to explain $RV_{\theta} = 9.4\%$ of the residual variance of both the predictor and the outcome. This means that compared to observed confounders which explained $\Delta R^2_{Y-DX} = 0.7\%$ of the residual variance of the outcome unobserved confounders would need to explain $R^2_{Y-DX} = 1\%$ of the residual variance of frequency of vaginal intercourse to fully account for the effect of hormonal contraceptives.

Without control for observed confounders the effect of hormonal controls on frequency of masturbation was statistically significant with an unstandardized effect size estimate of -0.04 [95% CI: -0.06; -0.03] (Bayesian analysis: -0.06 [90% HDI: -0.08; -0.04]). Adjusting for observed confounders reduced this effect by 50% (30% based on Bayesian analysis), resulting in an unstandardized effect size estimate of -0.02 [95% CI: -0.03; -0.01] (Bayesian analysis: -0.03 [90% HDI: -0.04; -0.02]).

Even though the effect of hormonal contraceptives on relationship satisfaction and the effect of congruent contraceptive use on frequency of masturbation were significant in the uncontrolled model based on frequentist analyses (see Tables 5 and 6), Bayesian analyses showed that the 90% HDI overlapped with our predefined ROPE (see Tables 3 and 4). These effects were therefore treated as insignificant here.
Figure 4. Unstandardized Effect Size Estimates and 90% HDI for Models With Congruent Contraceptive Use as an Outcome (0 = Incongruent; 1 = Congruent) Separately for HC Users When They Met Their Partner (Left) and Non-HC Users When They Met Their Partner (Right)

Note. Simple model (top) includes age, income, and relationship duration as predictors. Complex model (bottom) also includes education, extraversion, neuroticism, agreeableness, conscientiousness, openness, and religiosity. Error bars display 90% HDI.

HC = hormonal contraceptive; HDI = highest density interval.

Discussion

Our study aimed to disentangle selection effects from causal effects of contraceptive use. It showed that additional selection effects (including information about demography and personality) did not describe the choice of contraceptive method and congruent contraceptive use substantially better than did selection effects of age, income, and relationship duration. Furthermore, there was no evidence for substantial effects of contraceptive method, congruent contraceptive use, and their interaction on perceived partner attractiveness, relationship satisfaction, sexual satisfaction, and libido. While congruent contraceptive use and its interaction with contraceptive use had no substantial effects on frequency of vaginal intercourse and frequency of masturbation, we found a positive effect of current use of hormonal contraceptives on frequency of vaginal intercourse and a negative effect of current use of hormonal contraceptives on frequency of masturbation. These links were robust to the inclusion of observed confounders and sensitivity analyses suggested that unobserved confounders would need to strongly influence outcomes (about as strong as all observed confounders taken together) in order to substantially alter conclusions.

Selection Effects

Including additional selection variables pertaining to demography (education) and personality (openness, conscientiousness, extraversion, agreeableness, neuroticism, and religiosity) did not substantially improve models predicting contraceptive method or congruent contraceptive use compared to models based only on age, relationship duration, and income. Therefore, Hypothesis 1—that the complex model explains more variance compared to the simple model in (1) the choice of contraception and (2) congruent contraceptive use—was rejected.

Of the predictors included in the simpler models, age and relationship duration had a significant effect on choice of contraceptive method and congruent contraceptive use; income was no significant predictor. Overall, age had a negative effect on the use of hormonal contraceptives, i.e. the percentage of women using hormonal contraceptives decreased with increasing age. In addition, age had a negative effect on congruency in women who were using hormonal contraceptives when they met their partner (i.e., older women were more likely to switch to no/nonhormonal contraceptives) but a positive effect on congruency in women who were using no/nonhormonal contraceptives when they met their partner (i.e., younger women were more likely to switch to hormonal contraceptives). Overall, women in a romantic relationship were more likely to use hormonal contraceptives. Relationship length played no significant role...
Table 3. Unstandardized Effect Size Estimates of Hormonal Contraceptives on Outcomes

<table>
<thead>
<tr>
<th>Outcome [ROPE]</th>
<th>n</th>
<th>Uncontrolled model Effect size estimate [90% HDI]</th>
<th>Controlled model Effect size estimate [90% HDI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived partner attractiveness [-0.07; 0.07]</td>
<td>774</td>
<td>0.08 [-0.00; 0.17]</td>
<td>0.07 [-0.02; 0.14]</td>
</tr>
<tr>
<td>Relationship satisfaction [-0.04; 0.04]</td>
<td>774</td>
<td>0.08 [0.03; 0.14]</td>
<td>0.06 [0.00; 0.11]</td>
</tr>
<tr>
<td>Sexual satisfaction [-0.11; 0.11]</td>
<td>774</td>
<td>0.14 [0.01; 0.26]</td>
<td>0.11 [-0.02; 0.24]</td>
</tr>
<tr>
<td>Libido [-0.06; 0.06]</td>
<td>968</td>
<td>0.02 [-0.04; 0.08]</td>
<td>0.00 [-0.06; 0.07]</td>
</tr>
<tr>
<td>Frequency of vaginal intercourse [-0.05; 0.05]</td>
<td>897</td>
<td>0.25 [0.21; 0.29]</td>
<td>0.16 [0.12; 0.21]</td>
</tr>
<tr>
<td>Frequency of masturbation [-0.05; 0.05]</td>
<td>897</td>
<td>-0.40 [-0.44; -0.35]</td>
<td>-0.30 [-0.34; -0.25]</td>
</tr>
</tbody>
</table>

Note. Effect sizes and estimated HDIs in bold were outside of the predefined ROPE, and the null hypothesis was rejected. For all other effect sizes, the estimated HDI overlapped with the predefined ROPE.
ROPE = region of practical equivalence; HDI = highest density interval.

Effects of Hormonal Contraceptives

The evidence for effects of hormonal contraceptives is inconclusive on perceived partner attractiveness, relationship satisfaction, sexual satisfaction, and libido, therefore neither accepting nor rejecting Hypotheses 2.1(1–4)—hormonal contraceptives lead to decreased (1) perceived partner attractiveness, (2) relationship satisfaction, (3) sexual satisfaction, and (4) libido—and Hypothesis 2.2(1–4)—after controlling for all selection variables, hormonal contraceptives lead to decreased (1) perceived partner attractiveness, (2) relationship satisfaction, (3) sexual satisfaction, and (4) libido. The estimates were insufficiently precise; future research with even larger samples is needed to reach a conclusion. Nevertheless, given the rather small effect sizes, it appears unlikely that use of hormonal contraceptives has a strong association with these outcomes.

Hormonal contraceptives had a positive effect on frequency of vaginal intercourse, even after controlling for observed confounders—thereby rejecting Hypotheses 2.1(5) and 2.2(5). Contrary to the RCTs by Graham et al. (1995) and Zethraeus et al. (2016) that provided evidence for negative effects of hormonal contraceptives on sexual desire, sexual arousal, and sexual pleasure, the results of our study are in line with studies based on correlational data that found a positive relationship between hormonal contraceptives and sexual frequency (Alexander et al., 1990; Caruso et al., 2005; McCoy & Matyas, 1996).

Hormonal contraceptives had a negative effect on frequency of masturbation, even after controlling for observed confounders—thereby accepting Hypotheses 2.1(6) and 2.2(6). The fact that income was not associated with contraceptive use might be because the current study was based on a German sample. Germany’s health system covers gynecological exams and consulting, and, for women younger than 22, the cost of hormonal contraceptives. Therefore, this result might not be generalizable to women in countries with no mandatory health insurance (e.g., the United States) or in countries where contraceptives tend not to be covered by health insurance (e.g., Canada).
2.2(6). Most studies show no difference in frequency of masturbation between HC users and non-HC users (Alexander et al., 1990; Bancroft et al., 1991), but a recent study by Mark et al. (2016) provided evidence of a positive association between hormonal contraceptives and women’s dyadic libido and a negative association between hormonal contraceptives and women’s solitary libido. The libido item in our study included dyadic and solitary libido (“I experienced increased libido [desire to have sexual intercourse/to masturbate/to be sexually active].”) and did not distinguish between them as proposed by Spector et al. (1996). Thus, it seems possible that the divergent relationships described by Mark et al. (2016) resulted in the overall null relationship between hormonal contraceptives and libido that we observed. Our study could therefore provide evidence for behavioral consequences (measured as frequency of vaginal intercourse and frequency of masturbation) of the divergent relationships between hormonal contraceptives and dyadic and solitary libido described by Mark et al. (2016).

**Effects of Congruent Contraceptive Use**

Evidence was inconclusive on effects of congruent contraceptive use on perceived partner attractiveness, relationship satisfaction, sexual satisfaction, libido, frequency of vaginal intercourse, and frequency of masturbation after considering observed confounders. We could therefore neither accept nor reject Hypotheses 3.1—congruent contraceptive use leads to increased (1) perceived partner attractiveness, (2) relationship satisfaction, (3) sexual satisfaction, (4) libido, (5) frequency of vaginal intercourse, and (6) frequency of masturbation—and 3.2—after controlling for all selection variables, congruent contraceptive use leads to increased (1) perceived partner attractiveness, (2) relationship satisfaction, (3) sexual satisfaction, (4) libido, (5) frequency of vaginal intercourse, and (6) frequency of masturbation.

While these findings do not support most of the literature based on the congruency hypothesis (Birnbaum et al., 2019; Cobey et al., 2013; French & Meltzer, 2020; Roberts, Cobey, et al., 2014; Roberts et al., 2012; Roberts, Little, et al., 2014; Russell et al., 2014), they are in line with a recent large-scale replication attempt by Jern et al. (2018). Marcinkowska et al. (2019) provided additional evidence that questions the congruency hypothesis: In a large-scale study (n = 6,482), they found no evidence that women using the pill had weaker preferences for male facial masculinity than did women not using the pill. Differences in partner preferences have been suggested to be the driving mechanism behind the congruency hypothesis: Incongruent contraceptive methods are thought to lead to a shift in partner preferences, resulting in less satisfaction with the current

---

**Figure 5. Unstandardized Effects Size Estimates and 90% HDIs of Hormonal Contraceptives on Outcomes Based on Uncontrolled and Controlled Models**

Note. Thick dotted lines indicate ROPEs for outcomes, thin dotted lines indicate zero. Blue indicates that the 90% HDI overlapped with the ROPE, red indicates that the 90% HDI was outside the ROPE.

HDIs = highest density intervals; ROPE = region of practical equivalence.
romantic partner.

Jern et al. (2018) pointed out an important difference between their study and earlier studies on the congruency hypothesis: Earlier studies often had unequal distributions of congruent and incongruent users such that one group was almost entirely based on HC users or non-HC users. For instance, in the study by Cobey et al. (2015) the group of incongruent current HC users consisted of only four participants (3% of the final sample), while the group of congruent HC users consisted of 71 participants (59% of the final sample). This is especially problematic because most studies featured relatively small incongruent HC user groups and relatively large congruent HC user groups. Considering the small expected effect sizes based on the congruency hypothesis, main effects of current contraceptive use could have led to a spurious effect of congruency based on unequal distributions. Although the sample sizes of congruent and incongruent use differed in our study, the subgroups were relatively balanced (congruent non-HC users: 32%; congruent HC users: 31%; incongruent non-HC users: 19%; incongruent HC users: 17%) and the models always accounted for current contraceptive use and its interaction with contraceptive congruency.

Another possibility is that the reported effects based on the congruency hypothesis were false positives (Simmons et al., 2011). None of the earlier studies used preregistered hypotheses, and sample sizes were relatively small (ranging between \( n = 48 \) and \( n = 365 \), apart from the study by Roberts et al. (2012), which found a positive effect of contraceptive congruency on perceived attractiveness and sexual satisfaction (\( n = 993 \)). Indeed, two failed large-scale replication attempts (Jern et al., 2018 and the current study) and a range of recent evidence question the theory underlying the congruency hypothesis (Arslan et al., 2018; C. R. Harris et al., 2013; Jones, Hahn, & DeBruine, 2018; Jünger, Kordsmeyer, et al., 2018; Jünger, Motta-Mena, et al., 2018; Stern et al., 2020; Stern & Penke, in press; Wood et al., 2014; for a current discussion on evidence for psychological cycle shifts see Gangestad et al., 2019a, 2019b; Higham, 2019; Jones, Hahn, & DeBruine, 2018; Roney, 2019; Stern et al., 2019).

Overall, recent work has cast doubt on the evidence for both the assumed mechanism and the interaction effect underlying the congruency hypothesis. Our study could not accept the null hypotheses that there are no effects of contraceptive congruency on perceived partner attractiveness and sexual satisfaction, relationship satisfaction, sexual satisfaction, libido, frequency of vaginal intercourse, and frequency of masturbation because the sample size was too small and because we applied rigorous decision criteria for accepting the null hypotheses. Future research on contraceptive congruence should be preregistered, be adequately powered to detect small effects, and appropriately account for current and past contraceptive use.

### Table 4. Unstandardized Effect Size Estimates of Hormonal Contraceptives, Congruent Contraceptive Use, and Their Interaction on Outcomes

<table>
<thead>
<tr>
<th>Outcome [ROPE]</th>
<th>n</th>
<th>Predictor</th>
<th>Uncontrolled model Effect size estimate [90% HDI]</th>
<th>Controlled model Effect size estimate [90% HDI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived partner attractiveness [-0.07, 0.07]</td>
<td>774</td>
<td>HCs</td>
<td>0.17 [0.02; 0.31]</td>
<td>0.14 [-0.01; 0.29]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congruency</td>
<td>0.14 [0.01; 0.27]</td>
<td>0.10 [-0.03; 0.23]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interaction</td>
<td>-0.13 [-0.31; 0.04]</td>
<td>-0.09 [-0.28; 0.09]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HCs</td>
<td>0.05 [-0.03; 0.13]</td>
<td>0.04 [-0.05; 0.12]</td>
</tr>
<tr>
<td>Relationship satisfaction [-0.04; 0.04]</td>
<td>774</td>
<td>HCs</td>
<td>-0.10 [-0.18; 0.03]</td>
<td>-0.07 [-0.14; 0.01]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congruency</td>
<td>0.06 [-0.05; 0.16]</td>
<td>0.03 [-0.07; 0.14]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interaction</td>
<td>0.18 [-0.02; 0.39]</td>
<td>0.12 [-0.10; 0.33]</td>
</tr>
<tr>
<td>Sexual satisfaction [-0.11; 0.11]</td>
<td>774</td>
<td>HCs</td>
<td>0.15 [-0.03; 0.33]</td>
<td>0.04 [-0.15; 0.23]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congruency</td>
<td>-0.08 [-0.33; 0.19]</td>
<td>-0.01 [-0.28; 0.26]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interaction</td>
<td>-0.00 [-0.12; 0.12]</td>
<td>-0.02 [-0.14; 0.11]</td>
</tr>
<tr>
<td>Libido [-0.06; 0.06]</td>
<td>632</td>
<td>HCs</td>
<td>0.08 [-0.03; 0.18]</td>
<td>0.02 [-0.09; 0.12]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congruency</td>
<td>-0.03 [-0.18; 0.12]</td>
<td>0.04 [-0.11; 0.19]</td>
</tr>
<tr>
<td>Frequency of vaginal intercourse [-0.05; 0.05]</td>
<td>622</td>
<td>HCs</td>
<td>0.20 [0.12; 0.27]</td>
<td>0.14 [0.06; 0.22]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congruency</td>
<td>0.09 [0.02; 0.15]</td>
<td>-0.05 [-0.11; 0.01]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interaction</td>
<td>-0.09 [-0.18; 0.01]</td>
<td>0.05 [-0.05; 0.13]</td>
</tr>
<tr>
<td>Frequency of masturbation [-0.05; 0.05]</td>
<td>622</td>
<td>HCs</td>
<td>-0.37 [-0.46; -0.27]</td>
<td>-0.38 [-0.47; -0.28]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congruency</td>
<td>0.11 [0.04; 0.18]</td>
<td>0.04 [-0.04; 0.13]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interaction</td>
<td>0.01 [-0.10; 0.13]</td>
<td>0.10 [-0.02; 0.22]</td>
</tr>
</tbody>
</table>

Note. Effect sizes and estimated HDIs in bold were outside of the predefined ROPE; the null hypothesis was thus rejected. For all other effect sizes, the estimated HDI overlapped with the predefined ROPE.

HCs = hormonal contraceptives; ROPE = region of practical equivalence; HDI = highest density interval.
Sensitivity to Unobserved Confounders

We estimated the robustness of the effects of hormonal contraceptives and congruent contraceptive use in light of potential unobserved confounders. Sensitivity analysis suggested that the influence of unobserved confounders would need to be nearly 1.5 times as strong as the influence of observed confounders to fully account for the effect of hormonal contraceptives on frequency of vaginal intercourse, and nearly as strong as the influence of observed confounders to fully account for the effect of hormonal contraceptives on frequency of masturbation. Even when taking into account the broad range of included observed confounders (demography, personality, and romantic relationship information) it seems plausible that unobserved confounders might exist that would fully explain the reported effects of hormonal contraceptives on frequency of vaginal intercourse and frequency of masturbation. Besides potential unobserved selection effects, we now consider three additional possible challenges: reverse causality, attrition effects, and further unobserved confounders.

Selection Effects and Reverse Causality

There was a positive effect of hormonal contraceptive use on frequency of vaginal intercourse. Although frequency of vaginal intercourse was measured after contraception in the diary, it is somewhat habitual and thus stable. Reverse causality might therefore be at play, even after excluding women who were not sexually active and therefore not using hormonal contraception. Women who have sex more frequently might place a larger premium on safety and convenience for contraception. Higher frequency of vaginal intercourse is associated with a higher risk of (unwanted) pregnancy, and therefore safe contraception is even more important, especially if a woman does not want to forego sexual intercourse or use additional contraceptive methods. In addition, higher frequency of vaginal intercourse affects economic considerations: At higher sexual frequencies, the pill can be cheaper than condoms. Reverse causation would explain why there are effects on behavior (frequency of vaginal intercourse) but not on the psychological outcomes that might be expected to precede the behavior in the causal chain (libido and sexual satisfaction). A similar, if slightly more speculative, explanation could be plausible for frequency of masturbation. If women who have sexual intercourse only infrequently eschew the pill and its cost and side effects, they might instead opt to use condoms. If these women have the same level of libido as women who have sex more frequently, they might masturbate more. Because the stable component of some of these outcomes could be quite large, these are plausible unobserved confounders, and repeated longitudinal data would be needed to adjust for them.
Attrition Effects

Unlike our study, RCTs reported negative effects of hormonal contraceptives on libido, sexual arousal, and sexual pleasure (Graham et al., 1995; Zethraeus et al., 2016) as well as on general well-being (Zethraeus et al., 2017). One potential reason for the positive effect of hormonal contraceptives on frequency of vaginal intercourse reported in our study and in earlier correlational studies (Alexander et al., 1990; Caruso et al., 2005; McCoy & Matyas, 1996) are attrition effects. As Graham (2019) noted, there is great variability in women’s experiences with hormonal contraceptives, with reports of negative, positive, and no effects. Women with negative experiences were more likely to stop using hormonal contraceptives (Bancroft & Sartorius, 1990; Sanders et al., 2001), and discontinuation rates are high. For instance, 11.6% of Swedish women who took hormonal contraceptives for the first time stopped using them within six months (Josefsson et al., 2013). Predictors of discontinuation include emotional side effects, worsening of the premenstrual syndrome, decreased frequency of sexual thoughts, and decreased psychosexual arousability (Sanders et al., 2001). Women with depressive and premenstrual complaints tend to discontinue hormonal contraceptive use, leaving the remaining users with greater reported well-being (Bancroft & Sartorius, 1990). Therefore, it is likely that empirical, correlational evidence suggesting positive effects of hormonal contraceptives on sexual functioning stems at least in part from the fact that women with negative experiences of hormonal contraceptives switch to other contraceptive methods. Much of the current evidence on positive effects of hormonal contraceptives might thus rest on samples skewed toward women who have already tailored their contraception regimen to their experiences with hormonal contraceptives; reported correlations could even be the reverse of the average causal effect. This may also be the case in our study: The estimated effect of hormonal contraceptives may not equal their average effect because women with negative experiences of hormonal contraceptives had already stopped taking them. Women who continued using hormonal contraceptives would be more likely to have had positive experiences with them, which would result in an overall positive relationship between hormonal contraceptive use and frequency of vaginal intercourse based on correlational data masking negative causal effects on average. Analyses on the congruency of contraceptive use at the time of meeting the current partner can only partly address this, as participants were generally old enough to have been able to try out different methods of contraception before meeting their current partner. This limitation implies that estimated effects should not be expected to generalize to the experiences of women using hormonal contraceptives for the first time. Attrition effects could be studied using longitudinal data. In addition, re-
search on women using hormonal contraceptives for the first time could provide more information on how preferences for contraceptives form.

**Further Unobserved Confounders**

Besides the already included selection and outcome variables, frequency of vaginal intercourse has been found to relate to less restricted sociosexuality (Grøntvedt et al., 2020), increased satisfaction with own body image (Ackard et al., 2000), and increased satisfaction with life (Muise et al., 2016) in women. Frequency of masturbation has been found to be positively associated with less restricted sociosexuality (Velten & Margraf, 2017), body acceptance and orgasm frequency (Burri & Carvalheira, 2019), and greater importance of sex and higher levels of general anxiety and depression (Rowland et al., 2020) in women. Regnerus et al. (2017) reported a negative relationship between frequency of masturbation and contentment with sexual frequency.

While some of these potential unobserved confounders were not measured in the available dataset (in particular orgasm frequency, importance of sex, anxiety, depression, and contentment with sexual frequency), others would have been available (in particular sociosexuality, general life satisfaction, and satisfaction with own body image) but we decided not to include them in the current study to prevent controlling for potential colliders or mediators (Rohrer, 2018). Nevertheless, they could be strong unobserved confounders that could explain the observed links between hormonal contraceptive use and frequency of vaginal intercourse and frequency of masturbation. For example, higher desire for penetrative intercourse could lead to higher frequency of vaginal intercourse. At the same time, it could lead to the decision to use hormonal contraceptives because they are among the safest contraceptive methods available. Body acceptance could lead to higher frequency of masturbation and, at the same time, to the decision to use no/nonhormonal contraceptives in order to avoid artificial hormones. Therefore, even though our study quantitatively estimated the needed strength of unobserved confounders, it is unable to definitively rule out the possibility that the observed relationships are due to the influence of potential unobserved confounders. In future research on larger samples, more pointed comparisons of contraceptives with similar Pearl indices indicating the effectiveness of this birth control (e.g., the pill and intrauterine devices) might answer some of these questions, and again, examining within-subject changes in sexuality in longitudinal data would reduce some of the concerns about potential unobserved confounders.

**Causal Effects of Hormonal Contraceptives on Frequency of Vaginal Intercourse and Frequency of Masturbation?**

Our study provides evidence for a positive effect of hormonal contraceptives on frequency of vaginal intercourse and a negative effect of hormonal contraceptives on frequency of masturbation. Both effects were somewhat attenuated when adjusting for observed confounders. Fairly strong unobserved confounders would be necessary to nullify or reverse the remaining effects but some plausible candidates exist. The questions of reverse causality, selection effects, and attrition effects regarding the reported effects persist.

**Limitations and Strengths**

Our study was not without limitations: First, while the total sample size of our study was relatively large, the sample sizes used for the analyses—especially those investigating effects of congruent contraceptive use—were too small to reach a definite conclusion about potential effects. Second, even though our study provides valuable insights into the links between contraceptive use, relationship quality, and sexual functioning, the conclusions based on correlational data remain inconclusive about any putative causal effect of hormonal contraceptives.

Nevertheless, our study had several key strengths: First, even though conclusions remain cautious because of the available sample size, the size of the sample still exceeds domain standards. Second, measures for libido, frequency of vaginal intercourse, and frequency of masturbation were based on diary reports, which have been shown to be more reliable than retrospective behavior measurements (McCuliffe et al., 2007) and described as the gold standard for measuring sexual frequency by Graham et al. (2005). Third, by providing a directed acyclic graph, controlling for observed confounding, and estimating the sensitivity to unobserved confounding, this study is better positioned than previous correlational work to disentangle selection effects from causal effects.

**Constraints on Generality**

Following the guidelines on constraints on generality (Simons et al., 2017) the following four factors reduce the broad generalizability of the current results: First, the sample studied consisted of heterosexual WEIRD (Henrich et al., 2010) women with a high proportion of undergraduate psychology students. Although our main explanation assumes a universal biological mechanism for the results, different absolute hormone levels in less prosperous and well-nourished populations (Vitzthum, 2009) might affect the frequency of ovulation and hence the observable effect size. Moreover, some of our alternative explanations (such as reverse causality) may be much more dependent on circumstances. Second, even though the gold standard for measuring sexual frequency and masturbation frequency by using diary reports was applied, all results are purely based on self-reports and generalizability to other measures might be limited. Third, from a temporal perspective the composition of hormonal contraceptives (e.g., the dosage of estrogen and progestin) has changed over the decades and will change further in the future. If the reported effects are due to certain estrogen or progestin dosages, results might not be replicable based on samples from different times.
Table 5. Sensitivity Analyses for Effects of Hormonal Contraceptives

| Uncontrolled model | Outcome                        | Unstandardized Effect size | SE  | df   | t    | R²_{Y-D|X} | RV_{q=1} | RV_{q=1, α=0.05} |
|--------------------|--------------------------------|-----------------------------|-----|------|------|-----------|-----------|------------------|
|                    | Perceived partner attractiveness| 0.08                        | 0.05| 772  | 1.56 | 0.3%      | 5.5%      | 0%               |
|                    | Relationship satisfaction      | 0.08                        | 0.03| 772  | 2.73 | 1.0%      | 9.4%      | 2.7%             |
|                    | Sexual satisfaction           | 0.11                        | 0.08| 772  | 1.39 | 0.3%      | 4.9%      | 0%               |
|                    | Libido                         | 0.02                        | 0.04| 966  | 0.53 | 0%        | 1.7%      | 0%               |
|                    | Frequency of vaginal intercourse| 0.04                        | 0.01| 895  | 3.96 | 1.7%      | 12.4%     | 6.4%             |
|                    | Frequency of masturbation      | -0.04                       | 0.01| 895  | -5.14| 2.9%      | 15.8%     | 10.1%            |

| Controlled model including observed confounders | Outcome                        | Unstandardized Effect size | SE  | df   | t    | R²_{Y-D|X} | RV_{q=1} | RV_{q=1, α=0.05} |
|-------------------------------------------------|--------------------------------|-----------------------------|-----|------|------|-----------|-----------|------------------|
|                                                 | Perceived partner attractiveness| 0.09                        | 0.06| 756  | 1.54 | 0.3%      | 5.4%      | 0%               |
|                                                 | Relationship satisfaction      | 0.06                        | 0.03| 756  | 1.80 | 0.4%      | 6.3%      | 0%               |
|                                                 | Sexual satisfaction           | 0.11                        | 0.08| 756  | 1.39 | 0.3%      | 4.9%      | 0%               |
|                                                 | Libido                         | 0.01                        | 0.04| 949  | 0.16 | 0%        | 0.5%      | 0%               |
|                                                 | Frequency of vaginal intercourse| 0.02                        | 0.01| 878  | 2.94 | 1.0%      | 9.4%      | 3.2%             |
|                                                 | Frequency of masturbation      | -0.03                       | 0.01| 878  | -3.49| 1.4%      | 11.1%     | 5.0%             |

Note. Results are based on frequentist analyses. Substantial significant predictors based on Bayesian analyses are in bold.  
R²_{Y-D|X} = partial R² of the predictor with the outcome; RV_{q=1} = robustness value for bringing the point estimate of the predictor exactly to zero (percentage of residual variance of both the predictor and the outcome that unobserved confounders would have to explain to bring the point estimate to zero); RV_{q=1, α=0.05} = robustness value for testing the null hypothesis that the coefficient of the predictor is zero (percentage of residual variance of both the predictor and the outcome that unobserved confounders would have to explain for bringing the point estimate to a range where it is no longer statistically different from 0, at the significance level of 0.05).
### Table 6. Sensitivity Analyses for Effects of Hormonal Contraceptives, Congruent Contraceptive Use, and Their Interaction

<p>| Outcome                               | Predictor  | Unstandardized Effect size | SE  | Df   | t    | $R^2_{Y-D|X}$ | $RV_q = 1$ | $RV_q = 1, \alpha = 0.05$ |
|---------------------------------------|-----------|----------------------------|-----|------|------|---------------|-------------|---------------------|
|                                      |           |                            |     |      |      |               |             |                     |
| <strong>Uncontrolled model</strong>                |           |                            |     |      |      |               |             |                     |
| Perceived partner attractiveness      | HCs       | 0.17                       | 0.09|      | 1.90 | 0.5%          | 6.6%        | 0%                  |
|                                       | Congruency| 0.14                       | 0.08| 770  | 1.82 | 0.4%          | 6.4%        | 0%                  |
|                                       | Interaction| -0.13                      | 0.11|      | -1.22| 0.2%          | 4.3%        | 0%                  |
|                                       | HCs       | 0.05                       | 0.05|      | 0.97 | 0.1%          | 3.4%        | 0%                  |
| Relationship satisfaction             | Congruency| -0.10                      | 0.04| 770  | -2.37| 0.7%          | 8.2%        | 1.5%                |
|                                       | Interaction| 0.06                       | 0.06|      | 0.89 | 0.1%          | 3.2%        | 0%                  |
|                                       | HCs       | 0.18                       | 0.13|      | 1.46 | 0.3%          | 5.1%        | 0%                  |
| Sexual satisfaction                   | Congruency| 0.15                       | 0.11| 770  | 1.40 | 0.3%          | 4.9%        | 0%                  |
|                                       | Interaction| -0.08                      | 0.16|      | -0.48| 0%            | 1.7%        | 0%                  |
|                                       | HCs       | -0.00                      | 0.07|      | -0.05| 0%            | 0.2%        | 0%                  |
| Libido                                | Congruency| 0.08                       | 0.06| 628  | 1.26 | 0.3%          | 4.9%        | 0%                  |
|                                       | Interaction| -0.03                      | 0.09|      | -0.31| 0%            | 1.2%        | 0%                  |
| Frequency of vaginal intercourse      | Congruency| 0.03                       | 0.02|      | 1.59 | 0.4%          | 6.2%        | 0%                  |
|                                       | Interaction| -0.02                      | 0.02|      | -0.09| 0%            | 0.3%        | 0%                  |
| Frequency of masturbation             | HCs       | -0.04                      | 0.02|      | -2.75| 1.2%          | 10.5%       | 3.1%                |
|                                      |           |                            |     |      |      |               |             |                     |
| <strong>Controlled model including observed confounders</strong> |           |                            |     |      |      |               |             |                     |
| Perceived partner attractiveness      | HCs       | 0.14                       | 0.09|      | 1.54 | 0.3%          | 5.4%        | 0%                  |
|                                       | Congruency| 0.10                       | 0.08| 754  | 1.22 | 0.2%          | 4.3%        | 0%                  |
|                                       | Interaction| -0.09                      | 0.11|      | -0.78| 0.1%          | 2.8%        | 0%                  |
|                                       | HCs       | 0.04                       | 0.05|      | 0.72 | 0.1%          | 2.6%        | 0%                  |
| Relationship satisfaction             | Congruency| -0.07                      | 0.05| 754  | -1.49| 0.3%          | 5.3%        | 0%                  |
|                                       | Interaction| 0.04                       | 0.06|      | 0.56 | 0%            | 2.0%        | 0%                  |
| Sexual satisfaction                   | HCs       | 0.12                       | 0.13| 754  | 0.91 | 0.1%          | 3.3%        | 0%                  |
|                                      | Congruency| 0.04                       | 0.11|      | 0.33 | 0%            | 1.2%        | 0%                  |</p>
<table>
<thead>
<tr>
<th></th>
<th>Interaction</th>
<th>Libido</th>
<th>HCs</th>
<th>Congruency</th>
<th>Interaction</th>
<th>Libido</th>
<th>HCs</th>
<th>Congruency</th>
<th>Interaction</th>
<th>Libido</th>
<th>HCs</th>
<th>Congruency</th>
<th>Interaction</th>
<th>Libido</th>
<th>HCs</th>
<th>Congruency</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interaction</strong></td>
<td>-0.13</td>
<td>0.16</td>
<td>0.08</td>
<td>0%</td>
<td>0.3%</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Libido</strong></td>
<td>-0.02</td>
<td>0.07</td>
<td>-0.30</td>
<td>0%</td>
<td>1.2%</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HCs</strong></td>
<td>0.02</td>
<td>0.06</td>
<td>0.25</td>
<td>0%</td>
<td>1.0%</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Frequency of vaginal intercourse</strong></td>
<td>0.04</td>
<td>0.09</td>
<td>0.48</td>
<td>0%</td>
<td>1.9%</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HCs</strong></td>
<td>0.02</td>
<td>0.02</td>
<td>1.08</td>
<td>0.2%</td>
<td>4.3%</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Frequency of masturbation</strong></td>
<td>-0.01</td>
<td>0.02</td>
<td>0.62</td>
<td>-0.31</td>
<td>0%</td>
<td>1.2%</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HCs</strong></td>
<td>-0.04</td>
<td>0.02</td>
<td>-2.69</td>
<td>1.2%</td>
<td>10.4%</td>
<td>2.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Congruency</strong></td>
<td>0.00</td>
<td>0.01</td>
<td>0.08</td>
<td>0%</td>
<td>0.3%</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Interaction</strong></td>
<td>0.02</td>
<td>0.02</td>
<td>0.74</td>
<td>0.1%</td>
<td>3%</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Results are based on frequentist analyses. Substantial significant predictors based on Bayesian analyses are in bold.

**HCs** = hormonal contraceptives; $R^2_{Y|X,E}$ = partial $R^2$ of the predictor with the outcome; $RV_{q = 1}$ = robustness value for bringing the point estimate of the predictor exactly to zero (percentage of residual variance of both the predictor and the outcome that unobserved founders would have to explain to bring the point estimate to zero); $RV_{E \neq 0, a=0.05}$ = robustness value for testing the null hypothesis that the coefficient of the predictor is zero (percentage of residual variance of both the predictor and the outcome that unobserved founders would have to explain for bringing the point estimate to a range where it is no longer statistically different from 0, at the significance level of 0.05).
Future Research

Even after rigorous control for selection variables and considering the role of potential unobserved confounders, cross-sectional, correlational data seems insufficient to answer the question of whether hormonal contraceptive use causally affects women’s relationship quality and sexual functioning.

Conducting a RCT with a nonhormonal contraceptive placebo control group would be the most straightforward way to investigate causal effects of hormonal contraceptive use. Using the RCTs considering effects of the pill on libido, sexual arousal, and sexual pleasure (Graham et al., 1995; Zethraeus et al., 2016) as well as general well-being (Zethraeus et al., 2017) as role models, a RCT with a broader focus on relationship quality as an outcome could be conducted. In addition, instead of measuring overall sexual functioning retrospectively, a diary survey measuring libido, frequency of vaginal intercourse, and frequency of masturbation could be administered across at least three months, including a time window before and after the experimental treatment. The research focus could be broadened further by including not only the pill but other hormonal contraceptives (e.g., hormonal implants).

One problem of RCTs is that participants need to use nonhormonal contraceptive methods on top of hormonal contraceptives—otherwise placebo control groups would have no pregnancy protection. While Graham et al. (1995) only recruited women who had been sterilized or whose partners had been vasectomized, Zethraeus et al. (2016, 2017) included all women, and free condoms were distributed to prevent pregnancies (there were two pregnancies in the control group). The need to use additional contraceptive methods might mask some of the positive effects of hormonal contraceptives on relationship quality and sexual functioning (e.g., spontaneity).

Although RCTs are the superior approach for assessing the causal effects of the intervention, they have limited ecological validity when it comes to studying the processes of selection and attrition. These processes, by which women try out different contraceptive methods, are crucial for women’s satisfaction with their chosen contraception and how well they can tailor it to their individual needs. However, these processes presumably occur over much longer time frames.

Two issues we encountered could be solved by using longitudinal panel data. First, many hypotheses were undecidable because the sample size in the current study was too small. Second, attrition effects and reverse causality could not be eliminated as possible alternative explanations. Large longitudinal panel datasets would make it possible to investigate the effect of hormonal contraceptives use while controlling for potential confounders using propensity score matching. Attrition effects could be investigated by analyzing differences between women who continue and discontinue using hormonal contraceptives. To eliminate potential reverse causality, adjustment for the stable component of certain outcomes (e.g., sexual satisfaction) is necessary.

Conclusion

We found evidence that the use of hormonal contraceptives positively predicts frequency of vaginal intercourse and negatively predicts frequency of masturbation. Evidence for association of hormonal contraceptives with perceived partner attractiveness, relationship satisfaction, sexual satisfaction, and libido, as well as for the congruency hypothesis, was weak, but uncertain. These results were robust to the inclusion of observed confounders. Unobserved confounders would need to have a strong influence to nullify or reverse the observed relationships; however, some are plausible and could be assessed with longitudinal data. This study disentangled potential causal effects of hormonal contraceptives on frequency of vaginal intercourse and frequency of masturbation from selection effects to some extent but further research is needed to incorporate attrition effects and reverse causality.

Contributions

Laura J. Botzet (LJB), Tanja M. Gerlach (TMG), Lars Penke (LP), and Ruben C. Arslan (RCA) made substantial contributions to the conception and design of the current study. TMG, Julie C. Driebe (JCD), LP, and RCA were involved in the conception of the Goettingen Ovulatory Cycle Diaries 2 and data collection. LJB and RCA analyzed the current data and jointly interpreted the results. LJB drafted the article and TMG, LP, JCD, and RCA revised it critically. LJB, TMG, JCD, LP, and RCA approved the final version of this manuscript to be published.

Funding Information

Data collection of the Goettingen Ovulatory Cycle Diaries was partially funded by the Leibniz ScienceCampus Primate Cognition Göttingen. We acknowledge support by the Open Access Funds of the Göttingen University.

Acknowledgements

We would like to thank everybody involved in the Goettingen Ovulatory Cycle Diaries 2, especially Julia Ostner, and Julia Stern. In addition, we thank Dorle Schaper, who wrote her Bachelor’s thesis about selection effects of contraceptive methods based on data from the Goettingen Ovulatory Cycle Diaries 2. Finally, we would like to thank Deb Ain for her careful editing. All remaining errors are ours.

Competing Interests

No competing interests exist.

Supplemental Material

The Supplemental Material mentioned throughout this manuscript is available in “Supplemental Material: Psychological Effects of Hormonal Contraception” uploaded as a Word document with this submission. This manuscript also contains supporting information online at https://lau-
Data Accessibility Statement

We cannot share the data publicly due to the sensitive nature of sexual diary studies. Therefore, we uploaded a synthetic dataset to the Open Science Framework (https://osf.io/rqxsa/) that mimics many of the central features of the real data, including means and bivariate associations. It can therefore be used to write code to test and build models using realistic fake data. Upon request we can share the partially anonymised data with anyone who has a valid reason and agrees not to attempt to re-identify the data.

Submitted: March 06, 2021 PDT, Accepted: October 04, 2021 PDT
REFERENCES


Lang, F. R., Lüdtke, O., & Asendorpf, J. B. (2001). Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen [Test performance and psychometric equivalence of the German version of the Big Five Inventory (BFI) in young, middle-aged, and old adults]. *Diagnostica, 47*(3), 111–121. https://doi.org/10.1002/1924.47.111


SUPPLEMENTARY MATERIALS

Peer Review History

Supplementary Material