The Congruency Hypothesis Revisited: A Reassessment of the Association between Hormonal Contraception, Sexual Satisfaction, and Relationship Satisfaction

Tatu Erlin

Master's thesis

University of Turku

Faculty of Social Sciences

Department of Psychology and Speech-Language Pathology

December 2016

Recent research has found empirical support for the hormonal contraceptive (HC) congruency hypothesis, which predicts that a change in women's HC consumption will result in lower intra-pair sexual satisfaction due to altering the preferences for partner's traits, although these studies have left room for speculation of the actual effect of HC congruency. To reassess the congruency hypothesis, a large sample of 948 Finnish women was gathered (M_{age} = 23.95 years, SD = 4.16) who completed an online survey, including items about their current and previous HC use, general relationship satisfaction, and sexual satisfaction. The present study attempted to replicate results from two recent studies that have found support for the congruency hypothesis. A direct replication of one study was performed, and in addition data regarding sexual satisfaction and relationship satisfaction was collected using measures that have been previously validated (*Perceived Relationship Quality Components* and *Female Sexual Function Index*). The present study had 98.7% power to detect an effect for the association between sexual satisfaction and HC congruency of the same size as reported in previous studies. However, no support was found for the congruency hypothesis in the present study despite excellent statistical power, and regardless of whether the same measures as used in previous studies of HC congruency or validated measures of sexual satisfaction and general relationship satisfaction were used. Instead, the present study found that the largest mean differences in terms of sexual satisfaction were between the two different congruent HC user groups (i.e., between women who had used HCs both when they first met their partner and at the time of study participation, and women who had not used HCs at either time point), contradicting the congruency hypothesis. These results cast serious doubt on the congruency hypothesis, but the findings suggest a pharmacological main effect of HC use on relationship-related outcomes.

TABLE OF CONTENTS

TITLE

ABSTRACT

1. INTRODUCTION

The consumption of hormonal contraceptives (HCs) is very common and widely spread, for example four out of five sexually experienced women have used the pill in the US (Daniels, Mosher & Jones, 2013) and it has been estimated in *Trends in Contraceptive Use Worldwide 2015*, a report by the UN, that 27.8 % of the women aged from 15 to 49 worldwide use some sort of hormonal contraception, with HC use frequencies being higher in Western countries. In the present study, the consumption of HCs was similarly high, with 77.1 % of the women having used HCs at some point of their current relationship. The study of how women's consumption of hormonal contraceptives influences women's sexual desire, partner preferences, and sexual satisfaction in romantic relationships has been a popular subject of study for the past decades.

However, different studies have yielded conflicting results, so that HC use appears to be associated with increased satisfaction (or sexual desire) in some studies, but decreased satisfaction (or desire) in others (see Roberts, Cobey, Klapilová, & Havlícek, 2014 for review). Recently, the HC *congruency hypothesis* (Roberts, Cobey, Klapilová, & Havlícek, 2013) was introduced in an effort to explain these contradictory findings.

The congruency hypothesis postulates that women with incongruent HC use – that is, women who have used (or not used) HCs when they first met their partner, but stopped (or begun) using them before the present – report more negative consequences for their romantic relationships compared to women with congruent HC use (i.e., women who have either used or not used HCs over the entire course of their current relationship). The congruency hypothesis builds on previous research showing that hormonal fluctuations (such as those naturally occurring in different phases of the menstrual cycle) predicts variation in preferences for different partner attributes, such as perceived health (Jones et al., 2005), facial masculinity (e.g., Penton-Voak et al., 1999), as well as relationship-related behaviors and emotions, such as mate guarding (Welling, Puts, Roberts, Little, & Burriss, 2012) and jealousy (Cobey, Roberts & Buunk, 2013). Consequently, the congruency hypothesis rests on the assumption that women prefer different attributes in a sexual or romantic partner while under the influence of the hormones that constitute the active ingredients of HCs, namely progestin, often but not necessarily in combination with estrogen (e.g., Penton-Voak et al., 1999, Jones et al., 2008, Jones et al., 2005). The rationale of the congruency hypothesis is illustrated in Figure 1. In other words, women

who have their hormone levels pharmacologically altered during the course of a relationship are assumed to be at increased risk of suffering negative relationship consequences as a result of the pharmacological modification of partner preference.
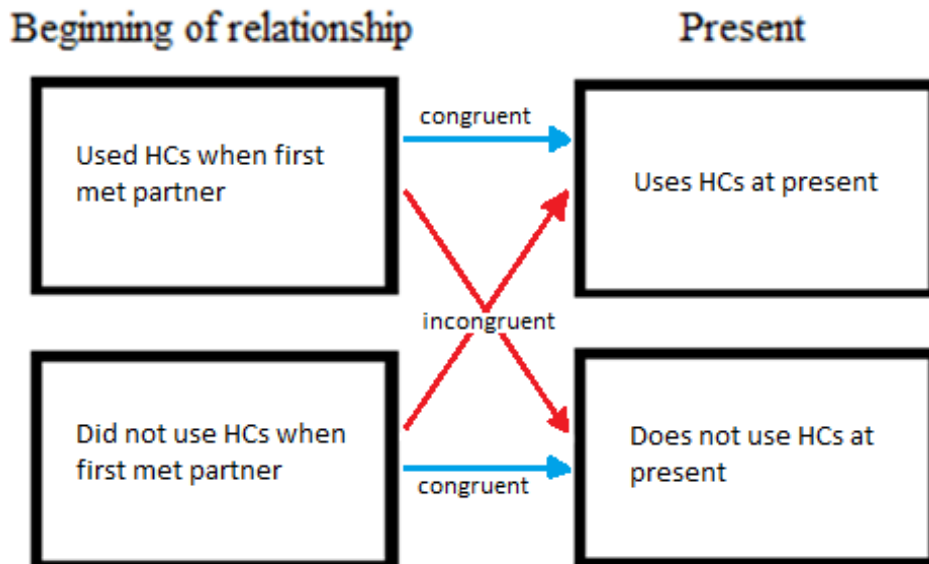

Figure 1. Blue arrows indicate routes to congruent use and red arrows to incongruent use.

To date, the HC congruency hypothesis has been subject to a few direct empirical tests, with most efforts having been undertaken to investigate the effects of HC congruency on sexual satisfaction and general relationship satisfaction. In the first of these, Roberts et al. (2014) showed, in a sample of 365 women, that incongruent HC use was associated with significantly lower values on women's sexual satisfaction scores compared to those women with congruent HC use. This association was robust also after controlling for a number of potential confounders, such as socioeconomic status, age, relationship duration, and general relationship satisfaction. However, Roberts et al. (2014) found no significant effects of HC congruency in the context of general relationship satisfaction. In a subsequent longitudinal study, Russell et al. (2014) replicated the negative effects of incongruent HC use on sexual satisfaction as previously demonstrated by Roberts et al. (2014). Russell et al. (2014) also reported that the association between incongruent HC use and general relationship satisfaction appeared to be moderated by the partner's facial attractiveness, so that women whose husbands had relatively less attractive face became less satisfied, but more satisfied if their husband had a relatively more attractive face. These results corroborate the congruency hypothesis, because women under the influence of HC can choose their partner differently (due to hormonally altered partner preferences)

than they would have without that influence, which in turn can result in less attractive partner as perceived by women after quitting the use of HC.

While the HC congruency hypothesis is an intriguing explanation for the ambiguous results concerning effects of HCs on relationship satisfaction-related factors, the studies that have investigated the HC congruency hypothesis have some potential limitations that contribute to the many reasons for conducting a replication study on the congruency hypothesis. Firstly, the questionnaire items used to measure sexual and general relationship satisfaction in the study by Roberts and others (2014) consisted of only two items each (*"How satisfied are you with your partner's ability to arouse you sexually?"* and *"How satisfied are you you're your husband's sexual adventurousness?"* in the case of sexual satisfaction; and *"How satisfied are you with your partner's intelligence / financial provision?"* in the case of general relationship satisfaction). Neither of the aforementioned measures had been validated. For example, at face value, the general relationship satisfaction domain used by Roberts et al. (2014) may not adequately capture factors that are important for one's satisfaction in a dyadic relationship (one's partner can, for instance, be intelligent and wealthy yet possess other traits that makes him difficult to be in a fulfilling relationship with, such as narcissistic or psychopathic traits). Russell and others (2014) did employ validated questionnaires, except in the second sample sexual satisfaction was measured with only one item: *"How satisfied are you with the quality of the sex you have had with your spouse over the past 4 months?"*

Secondly, in the study by Russell et al. (2014), the small sample size ($N = 48$ and 70 in two separate samples) suggests they were underpowered to detect an effect of the size (partial $\eta^2 = .018$) reported for the association between HC congruency and sexual satisfaction by Roberts et al. (2014). Due to these deficiencies in the measures used in previous studies, moderate correlations between these measures and the validated measures were expected, as it is not expected for the previous methods to measure accurately sexual satisfaction or general relationship satisfaction.

Thirdly, in psychology it is often the case that previous studies' results cannot be, for one or another reason, reproduced, which has recently led to repeated calls from researchers to place more emphasis on replication studies (e.g., Open Science Collaboration, 2015). This is especially important if the previous studies have had small sample sizes, or if other

3

factors known to possibly affect the outcomes of the study are identified (e.g., non-validated measures).

Fourthly, the aforementioned studies by Roberts et al. (2014) and Russell et al. (2014) have only taken into account two groups regarding the HC congruency; congruent and incongruent. Regarding the HC congruency, four groups can actually be distinguished; congruent users that always used HCs, congruent users who never used HCs, incongruent users who have quit using HCs and incongruent users who have started using HCs between forming the current relationship and the present (illustrated by the four arrows in Figure 1, respectively). This more precise categorization of congruency of HC use allows for observation of differences between the two congruent and the two incongruent groups. This is important, because in previous studies testing the HC congruency hypothesis, there have been discrepancies in group sizes. For instance, in a study testing implications of the HC congruency hypothesis on jealousy (Cobey et al., 2013), the incongruent group consisted mainly (87 %) of current HC users (only four individuals with incongruent use were current non-users of HCs), whereas the congruent group consisted mainly of current HC abstainers. Thus, if HCs exert a main effect on jealousy, this may turn up as an artefactual HC congruency effect if group sizes are unbalanced. Should the data support the congruency hypothesis, the two incongruent groups should differentiate from the two congruent groups in the analyses in which the four-group categorization is used.

The goal of the present study was to reassess the validity of the congruency hypothesis by using the same measures as used by Roberts et al. (2014), as well as previously validated questionnaires (*Female Sexual Function Index* and *Perceived Relationship Quality Components*). The present study aimed to gather a large sample of women in order to achieve good statistical power to detect the effect sizes reported by previous studies.

Based on the above, the following four hypotheses were formulated:

1) The results concerning the association between HC congruency and sexual satisfaction reported by Roberts et al. (2014) and Russell et al. (2014) would replicate, so that incongruent HC use was expected to predict lower sexual satisfaction;

2) A validity assessment of the measures of general relationship satisfaction and sexual satisfaction used by Roberts et al. (2014) was expected to show moderate validity for these measures (i.e., correlations of moderate size between these and well-validated measures of sexual satisfaction and general relationship satisfaction), and that statistical analyses using previously validated measures of general relationship satisfaction and sexual satisfaction would yield stronger associations (higher effect size, lower p values);

3) The present study set out to exploratively investigate whether the studies of Roberts et al. (2014) and Russell et al. (2014) could have been underpowered to detect a statistically significant main effect of incongruent HC use on general relationship satisfaction, expecting a statistically significant decrease (with small effect size) in general relationship satisfaction in the event that the data collection would yield a considerably larger sample (i.e., with better statistical power) than that used in previous studies; and

4) Women's reports of the partner's physical attractiveness were expected to moderate the association between HC congruency and general relationship satisfaction (Russell et al., 2014), so that women who perceive their partner to be relatively more attractive are more satisfied than women who perceive their partner to be relatively less attractive, in women who report incongruent HC use and more specifically in women who discontinued HC use during the relationship.

2. MATERIALS AND METHODS

2.1. Sample description and data collection procedure

Statistical analyses were based on $N = 948$ Finnish women ($M_{age} = 23.95$ years, $SD = 4.16$), who answered in all questions relevant to the present study. In the spring of 2015, women aged 18 or above who were in steady relationships were targeted to participate in the present study. Recruitment was done through university email lists at five different universities or polytechnic schools in Turku and Helsinki, Finland, as well as through social media (Facebook, Twitter). The covering letter of the study can be viewed in appendix 1. In total, 1,590 individuals logged on to the survey, and of these, 468 did not answer any of the questions relevant to the present study. All data were collected through a secure, online questionnaire. The survey was kept open for a period of four weeks (the time frame was decided before commencing the data collection). In total, 1,590

5

individuals logged on to the survey, and of these, 468 did not answer any of the questions relevant to the present study. The online survey was constructed so that participation would terminate if the participant stated that 1) his/her gender was not female (30 individuals); 2) she was not in a relationship at present (33 individuals); 3) she was in a non-heterosexual relationship (29 participants); or 4) she was undergoing any sort of hormonal treatment, such as fertility treatment (10 individuals). Finally, following Roberts et al. (2014), women who were pregnant, trying to become pregnant, or uncertain if they were pregnant or not (66 individuals) were excluded. Six individuals were uncertain whether they had used HCs or not when they first met their partner (and could thus not be allocated to congruent or incongruent HC user groups for statistical analyses), leaving with a final study sample of 948 women. All participants provided written, informed consent. The research plan was approved by the Ethics Committee of the Åbo Akademi University (Turku, Finland) in accordance with the Helsinki Declaration.

2.2. Measures

Use of hormonal contraceptives. Participants responded to two dichotomous queries: *"Did you use any hormone-based contraceptives when you first began dating your current partner?"* [yes/no], and *"Do you use any hormone-based contraceptives at the moment?"* [yes/no]. In addition, inquiry was made about type and brand of HC, and for how long they had been using their current contraceptive method (if any).

General relationship satisfaction. In order to measure general relationship satisfaction, two different measures was used. The first was the same measure used by Roberts et al. (2014), a composite variable consisting of two questions (*"How satisfied are you with your partner's financial provision/intelligence?"*), with response alternatives on a 9-point Likert scale with the anchors "my partner does not satisfy me at all on this condition" and "my partner completely satisfies me on this condition". The internal consistency of this composite variable was modest (Cronbach's α = .419). The second was the *Perceived Relationship Quality Components* (PRQC), which has been shown to have good internal consistency and predictive validity in previous studies (Fletcher et al., 2000a, 2000b). The PRQC consists of six questions (e.g., *"How satisfied are you with your relationship?"*) with response alternatives on a 7-point Likert scale with the anchors "not at all" and "extremely". In the present study, the internal consistency of this measure was good

(Cronbach's α = .858). Roberts' and others' (2014) measure of general relationship satisfaction had not previously been validated. The correlation between this measure and the PRQC was modest ($r = .342$, $p < .001$).

Sexual satisfaction. In similar fashion, two different measures of sexual satisfaction was used: first, a composite variable consisting of two questions used by Roberts et al. (2014) (*"How satisfied are you with your partner's sexual adventurousness/ability to arouse you sexually?"*), with the same response alternatives as above. The internal consistency of this composite variable was fairly good (Cronbach's α = .770). Second, the sexual satisfaction subscale of the *Female Sexual Function Index* (FSFI) was used, which consists of three items (e.g., *"Over the past four weeks, how satisfied have you been with your sexual relationship with your partner?"*) responded to on a five-point Likert scale with the anchors "very dissatisfied" and "very satisfied". The FSFI has been found to possess good validity in previous studies (e.g., Rosen et al., 2000). The FSFI measures sexual function and satisfaction based on the past four weeks. Individuals who had not engaged in partnered sexual activity in the past four weeks were excluded from analyses using the FSFI sexual satisfaction domain (50 individuals; 5.3%). This composite variable had good internal consistency (Cronbach's α = .860). Again, Roberts et al's (2014) measure of sexual satisfaction had not previously been validated. The correlation between this measure and the sexual satisfaction subdomain of the FSFI was moderate ($r = .496$, $p < .001$).

Partner attractiveness. The following two questions were used to form a composite variable measuring how attractive the participants perceived their male partners to be: *"Compared to other men, how attractive do you think your partner's face and body is?"* and *"How satisfied are you with your partner's physical attractiveness?"* Both items were responded to on a 7-point scale, with higher scores indicating higher perceived attractiveness. The internal consistency of this composite variable was fairly good (Cronbach's α = .787).

Control measures. Based on the studies conducted by Roberts et al. (2014), data for a number of covariates was collected to be able to check that any detected associations would remain robust after controlling for these. These were: household income (9-point scale with the anchors "no income" and "more than €100,000 per year"), relationship

7

duration, age, and whether the participants had children together with their current partner.

2.3. Statistical analyses

All statistical analyses were conducted using SPSS 22.0. Analyses were began by computing descriptive statistics, and proceeded to calculate the main effects of HC use, then assess the validity of the sexual satisfaction and general relationship satisfaction measures used by Roberts et al. (2014). This was done by correlating these composite variables with the composite variables formed by summing the items of the FSFI sexual satisfaction subscale and the PRQC, respectively. Next, it was proceeded to assess the effect of HC congruency on sexual and general relationship satisfaction by fitting linear regression models (the Generalized Linear Model module of SPSS) to the data. The latter analyses were conducted first using the measures of sexual satisfaction and general relationship satisfaction used by Roberts et al. (2014) as dependent variables, and then again using the FSFI sexual satisfaction subscale and the PRQC as dependent variables. After this, all regression models were repeated (separate analyses for the two-group and four-group congruency variables) now controlling for household income, presence or absence of children, relationship duration, and age. Then, the regression models were repeated using the detailed (four-group) congruency variable as independent variable. Finally, an interaction model was fitted with HC congruency and partner's attractiveness as the predictors.

Statistical power

In the study by Roberts et al. (2014) the reported effect size for the congruency effect was $\eta^2_p = .018$ and had the sample size of $N = 365$. Thus, it can be calculated that with the present study's sample size of 948 and number of predictors set to one (congruency) at the critical p-value of 0.05, the present study had the statistical power of 98.7 % of detecting the effect of a same size that reported by Roberts and others. Russell et al. (2014) report an effect size of $r = 0.14$ (equals $r^2 = 0.02$) in the analysis regarding HC congruency moderated by husband's facial attractiveness on marital (general relationship) satisfaction. The present study had the statistical power of 99.2 % to detect an effect of the same size.

8

3. RESULTS

3.1. Descriptive Statistics

Corroborating international reports, the consumption of HC was very common in the present study's data, with only 217 out of the 948 women (22.9 %) reporting not having used HCs at either time point. The distribution of participants over the four groups based on HC congruence and incongruence can be viewed in Table 1. Most women (58.1 %) reported congruent use. The incongruent group consisting of women who had used HCs when they first met their partner but stopped doing so before participating in the study was the smallest group (77 participants, 8.1 %).

Table 1.

Frequencies for the Different Groups based on Hormonal Congruency

	Frequency	Percent
Congruent - used at both time points	334	35.2
Incongruent - used when first met but not at present	77	8.1
Incongruent - did not use when first met but uses now	320	33.8
Congruent - did not use at either time point	217	22.9
Total	948	100

The two-item composite variable (as used by Roberts and others) of sexual satisfaction had a relatively high mean score of 7.25 and median of 7.5. The respective numbers for the two-item composite of general relationship satisfaction were very similar with the mean of 7.27 and median of 7.5. Composition score of PRQC had the mean score of 6.12 and median of 6.33. FSFI had the mean score of 4.06 and median of 4.21. These can be viewed among others in table 2.

Nearly half of the subjects (47.5 %) reported their household income being less than €25,000, and average household income was €25,000 – €37,499. Fifty women (5.2 %) had biological children with their current partner.

9

Table 2.

Descriptive Statistics for Outcome Variables and Covariates

	M	Md	SD	Min	Max
Age	23.96	23	4.17	18	58
Relationship length, months	41.59	31	38,46	1	398
Sexual satisfaction (Roberts et al., 2014)	7.25	7.5	1.68	1	9
Sexual satisfaction (FSFI, Rosen et al., 2000)	4.06	4.21	0.91	1.26	4.95
Relationship satisfaction (Roberts et al., 2014)	7.27	7.5	1.48	1	9
Relationship satisfaction (PRQC, Fletcher et al., 2000)	6.12	6.33	0.83	1	7

Note. M = mean ; *Md* = median ; *SD* = standard deviation ; *Min* = minimum value ; *Max* = Maximum value.

3.2. Direct impact of HC use on sexual and general satisfaction

The direct impact of HC use on sexual and general satisfaction were analyzed with regression models first to simply assess the pure impact of HC use before moving on to the congruency analyses. A direct impact of HC use was found (HC use predicting higher satisfaction scores) using Roberts' and other's sexual satisfaction measure with both present use (F = 4.993, p = .027, $\eta^2_{partial}$ = .005) and the use on the time of forming the relationship (F = 4.927, p = .026, $\eta^2_{partial}$ = .005). However, the model with the FSFI did not yield significant results from either time points (both ps > .074). Using the general relationship satisfaction measures, Robert's and others' measure did not yield significant results from either time points (both ps > .452), however the model with the PRQC was significant regarding the present use (F = 5.651, p = .018, $\eta^2_{partial}$ = .006) but not the past (p = .435, $\eta^2_{partial}$ = .001).

3.3. Validity assessment of measures of sexual and general satisfaction

All correlations between the four different variables measuring sexual and relationship satisfaction can be viewed in Table 3. Correlation of the two-item sexual satisfaction used

10

by Roberts et al. (2014) variable and the sexual satisfaction domain of the FSFI was moderate ($r = .496$, $p < .001$), and the respective correlation between two-item general relationship satisfaction used by Roberts et al. (2014) and the PRQC composite variable was also moderate ($r = .342$, $p < .001$). These correlations suggest that the two-item composite variables measure the same underlying constructs to some extent, but the relatively modest correlations between the validated and non-validated measures suggest that the measures used by Roberts et al. (2014) may not accurately represent sexual satisfaction or relationship satisfaction. These moderate correlations seem to contradict the second hypothesis of the present study, thus increasing expectations for the FSFI and PRQC to produce robust associations in line with the HC congruency hypothesis.

Table 3.

Correlations between Different Measures of Sexual Satisfaction and General Relationship Satisfaction

	Sexual satisfaction		Relationship satisfaction	
	Roberts et al.	FSFI	Roberts et al.	PRQC
Sexual satisfaction (Roberts et al.)	1			
Sexual satisfaction (FSFI)	.496 ***	1		
Relationship satisfaction (Roberts et al.)	.265 ***	.113 ***	1	
Relationship satisfaction (PRQC)	.529 ***	.353 ***	.342 ***	1

Note. FSFI = Female Sexual Function Index (Rosen et al., 2000), PRQC = Perceived Relationship Quality Components (Fletcher et al., 2000)

*** $p < .001$

11

3.4. HC congruency effect on sexual and relationship satisfaction

Two separate linear regression models were fitted using the two different measures of sexual satisfaction as dependent variables, and the two-group HC congruency variable as a factor to see if the present study's first hypothesis would gain support. Surprisingly, in the present study no association was found between HC congruency and sexual satisfaction. This was true both using sexual satisfaction measure used by Roberts and others (Wald $\chi^2[1] = .564$, $p = .453$) and when using the FSFI sexual satisfaction subdomain (Wald $\chi^2[1] = 1.346$, $p = .246$). The same was true for general relationship satisfaction: Corroborating Roberts and others' (2014) results, but in contradiction to the present study's third hypothesis, no significant main effect of HC congruency was found on general relationship satisfaction either with Roberts and other's measure, (Wald $\chi^2[1] = .262$, $p = .609$) or with the PRQC (Wald $\chi^2[1] = 2.059$, $p = .151$). The differences of the two groups are depicted in Figure 2.
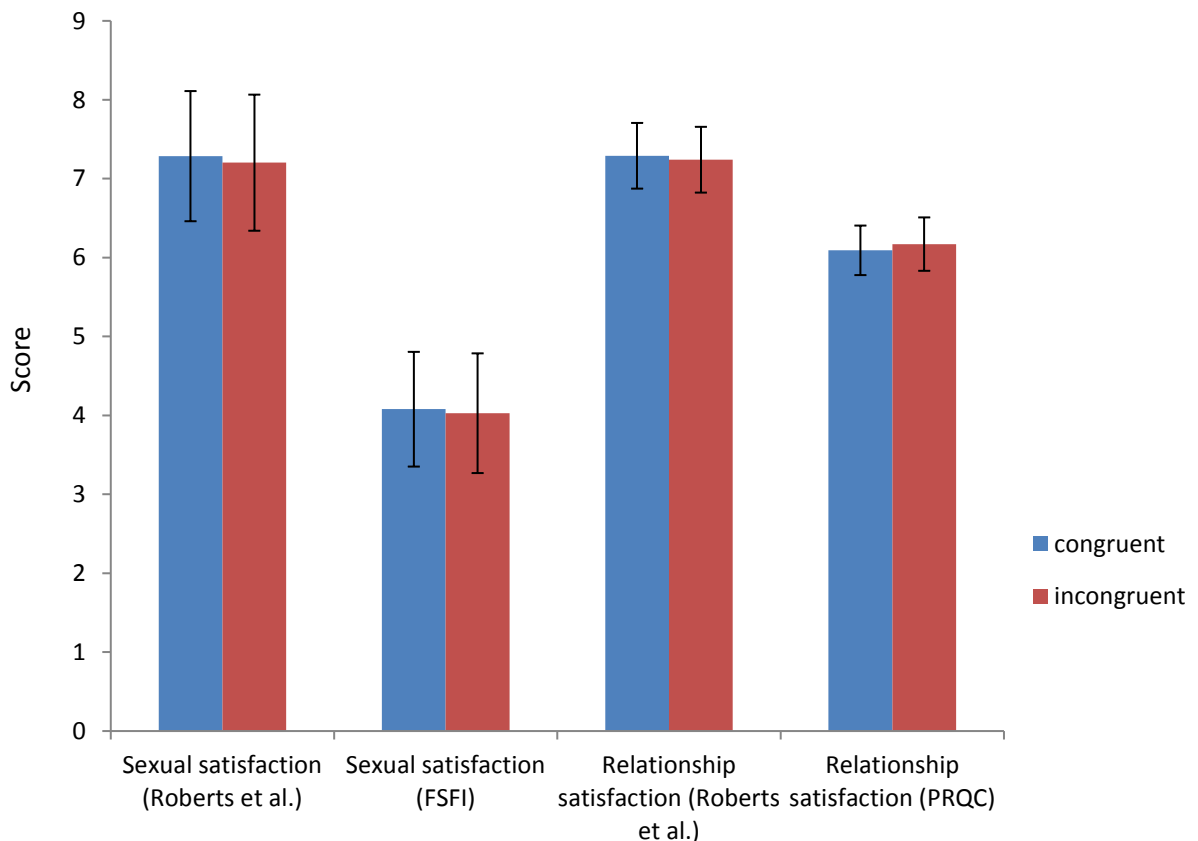

Figure 2. Mean scores of congruent and incongruent hormonal contraceptive user groups for four measures of sexual satisfaction and general relationship satisfaction. Error bars represent standard deviations. Note that the scale is not same for all variables.

12

Next, the regression models described above were rerunned while controlling for effects of potential confounders as described by Roberts et al. (2014). As inclusion of these covariates caused the effect size of the association between HC congruency and sexual satisfaction to become smaller in Roberts et al's (2014) study, this operation was not expected to improve the likelihood in detecting a significant effect of HC congruency. Indeed, none of the four analyses (i.e., two different measures of sexual satisfaction, and two different measures of general relationship satisfaction as dependent variables) generated a significant effect of HC congruency status on sexual or general relationship satisfaction after controlling for confounders (all ps > .146).

3.5. Routes to (in)congruency (four congruency groups analyses)

Next, regression models were fitted to the data in a similar fashion as above, but instead using the four-group HC congruency variable as the key predictor to further determine if there are any differences in sexual or general relationship satisfaction between any of the four groups that couldn't be found with the simple congruency categorization. Similarly to the analyses above, four different models were run; one for each measure of sexual and general relationship satisfaction. A significant main effect of the four-group congruency variable was detectable within the Robert's and others (2014) sexual satisfaction measure (Wald χ^2[3] = 13.095, p = .004) and within the PRQC (Wald χ^2[3] = 9.706, p = .021), whereas the models with the FSFI (Wald χ^2[3] = 3.692, p = .297) and Robert's and others relationship satisfaction (Wald χ^2[3] = .823 p = .844) were not statistically significant. To see what groups are statistically different from each other, i.e. if the differences between the groups support the congruency hypothesis or not, pairwise comparisons were made to compare each of the four congruency group to the other three. Surprisingly, and countering the congruency hypothesis, significant between-group differences were found between the two congruent groups in both Robert's and others sexual satisfaction and in the PRQC (both ps < .007), the HC users having higher mean score of satisfaction in average in both. Besides these differences, significant difference were found regarding Robert's sexual satisfaction measure between congruent (HC users) and incongruent (those who started HC use during relationship), and between congruent (non-users) and incongruent (those who started HC use during relationship) (both ps < .036). The results are depicted in Figure 3.
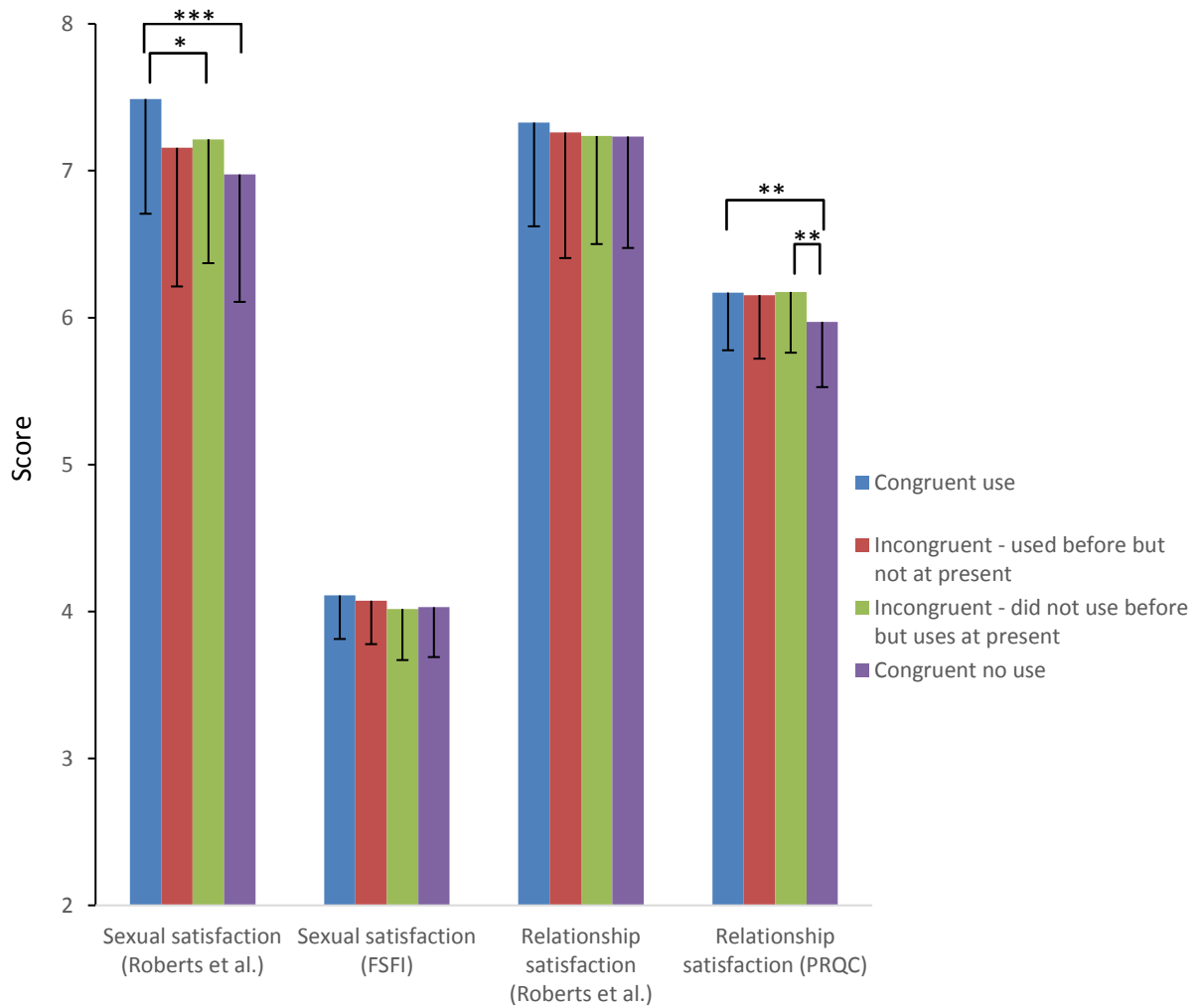
13

Figure 3. Mean satisfaction scores for each measure and congruency group. Error bars represent standard deviation. Significant differences are marked with lines above the respective bars. Note that the scale is not same for all variables.

* = p<.05, ** = p < .01, *** = p < .001.

3.6. Facial attractiveness as moderator of HC congruency's effect on relationship satisfaction

In the study by Russell and others (2014), a significant effect of HC incongruency on general relationship satisfaction mediated by the facial attractiveness of the women's spouse was found. In their study, women who discontinued their HC use became less satisfied if their spouses had relatively less attractive face, whereas those whose spouses had relatively more attractive face became more satisfied. An attractiveness assessment of the spouses' faces wasn't possible to execute within the present study, and therefore a two-item composite variable was used. As stated in the present study's fourth hypothesis, this moderation effect was expected to replicate with the present study's sample, which was not the case either with the PRQC ($F = 0.482$, $p = .488$) or with the Robert's and others' (2014) measure ($F = 0.033$, $p = .855$).

4. DISCUSSION

The present study did not yield any statistically significant results to corroborate the HC congruency hypothesis. On the contrary, the results concerning sexual satisfaction suggest that the greatest differences were between the two different congruent groups in relation to sexual satisfaction (using the measure of Roberts and others) and in relationship satisfaction (using the PRQC), which fundamentally speaks against the congruency hypothesis.

The first hypothesis of the present study, that the results concerning the association between HC congruency and sexual satisfaction reported by Roberts et al. (2014) and Russell et al. (2014) would replicate, so that incongruent HC use was expected to predict lower sexual satisfaction, was not supported by the results. HC congruency status did not predict sexual satisfaction scores with any measure in the present study, despite it having excellent statistical power to detect previously reported effect sizes.

The second hypothesis was partly confirmed: the validity of the measures of sexual and general relationship satisfaction used in previous studies (Roberts et al., 2014) was doubtful. As expected, the measures correlated only modestly with well-validated measures of sexual and general relationship satisfaction, keeping in mind that they should have measured the same constructs. However, the results reported by Roberts et al. (2014) and Russell et al. (2014) were still expected to replicate using the same measures, but

15

they did not. Furthermore, stronger statistical associations were expected to be detected in support of the HC congruency hypothesis when validated measures of sexual and general relationship satisfaction were used, but no support was found for the HC congruency effect with any measure. Thus, the latter part of the second hypothesis did not gain support from the results.

The measures used by Roberts et al. (2014) were not previously validated, and the present study cannot confirm them to reliably measure sexual satisfaction and relationship satisfaction. Therefore, in the light of the present study, even the significant results of the previous studies allegedly measuring sexual and general relationship satisfaction and correlating them to HC congruency cannot be said to predict sexual and relationship satisfaction reliably. In the present study, the significant difference between the two congruent groups in sexual satisfaction scores by the measure of Roberts and others could be explained with a confounding factor, for example the congruent users might value their sex life more and therefore put more effort in it (as could be reflected in the item about satisfaction of partner's sexual adventurousness). This, however, cannot be generalized as the entire concept of sexual satisfaction, as there were no significant differences between these groups with the FSFI.

The third hypothesis of the present study concerned the hypothesized impact of HC congruency on general relationship satisfaction. Despite the excellent statistical power of the present study and using the validated *Perceived Relationship Quality Components* – questionnaire, no association between HC use congruency and relationship satisfaction was found. Furthermore, in the congruency analysis using the PRQC (which is the most reliable measure for general relationship satisfaction in the present study) the incongruent group had higher, although not statistically significantly, satisfaction scores – contrary to what the congruency hypothesis would predict. This direction in the difference suggests that the statistically insignificant results are not caused by any statistical issues of the present study, for example insufficient statistical power, to detect the effect the congruency hypothesis predicts, and simply questions the congruency hypothesis itself.

As per the fourth hypothesis of the present study, women's reports of their partners' physical attractiveness were expected to moderate the association between HC congruency and relationship satisfaction in a similar fashion as reported by Russell et al.

16

(2014). In the present study, no such moderation effect could be found, and thus the fourth hypothesis did not gain support either.

The relevancy of the study by Russell et al. (2014) for the present study surfaced after the data collection for the present study had already begun, and therefore similar kind of measure for the spouses' facial attractiveness was not included in the present study. Secondly, due to the nature of the present study (online questionnaire aiming in large sample size), an expert review of the women's spouses' facial attractiveness was not done as in Russell and others' (2014) study. Nevertheless, data regarding women's subjective view of their partner's physical attractiveness was gathered, making it possible to replicate the analyses done in the aforementioned study. The measure is not as objective as in Russell and others' study; it is possible that the measure of the present study does not reflect genetic fitness as accurately as the measure of the study by Russell and others. However, similar results were expected presuming the facial moderating effect exists even with slightly different measures, as the present study had substantially greater statistical power compared to Russell et al. (2014), and as it was expected to the measure of physical attractiveness to correlate with a measure of facial attractiveness similar to that used by Russell et al. (2014). The aforementioned study had a small sample size with an especially small incongruent group of HC quitters (the only group in which this effect was significant) and the results, as it in the case of Roberts and others' (2014) study, could be a consequence of another factor altogether and the results may not be generalized.

All in all, contradicting the previous studies, the present study did not find any kind of evidence to support the congruency hypothesis. Possibly reasons for the phenomenon to not replicate in the present study are discussed next. As in the previous studies regarding the congruency hypothesis (Roberts et al., 2014; Russell et al., 2014; Cobey et al., 2013), the four congruency groups in the present study were clearly unbalanced (see Table 1); women who use HC during relationship formation and quit using during the relationship (without trying to become pregnant) was the smallest group at only 8.1 % representation of the sample with 77 individuals. However, in the study by Roberts et al. (2014) the sample was also unbalanced, but in a different way: the smallest group in their study was those who had started using HCs (19 individuals who started using HCs versus 111 of those who quit). This difference may have caused the congruency effect to seem significant. Considering the vast amount of reasons that could cause a woman to quit (or to start) the use of HCs, the congruency hypothesis could also be an effect of some other

17

factor of a couple's relationship (e.g., a couple discontinues to have intercourse for any reason and therefore the need to use HCs is discontinued) which correlates with sexual satisfaction scores and incongruent HC use, especially quitting HCs. This effect could be especially amplified in a relatively modest sample size and overrepresentation of the HC quitters, thus distorting the results. Furthermore, Roberts and other's study also has a small group of current HC users: 66 users versus 223 nonusers. To compare, in the present study nearly 70 % of the sample were current HC users, whereas the percentage of users in Roberts' and others' study is considerably lower at approximately 23 %. Given that Roberts and others did not find significant main effects of HC use (either previous or current), it might be possible for the skewed sample together with the strong representation of HC quitters to negate the main effect of HC use when comparing current users to nonusers, and thus cause the effect only to surface, as the congruency hypothesis predicts, when combining the incongruent groups. The present study found main effects of HC use, using the same measure used by Roberts and others, but the effect sizes were small ($\eta^2_{partial} = .005$ in both time points), which furthermore corroborates that the study by Roberts et al. (2014) might have been underpowered to detect this main effect, given their relatively small sample size.

The women who participated in the present study were relatively young on average (approximately 24 years) compared to the previous studies, but this shouldn't affect the presence of congruency hypothesis, since the congruency effect, if it exists, should be valid regardless of factors such as age or socioeconomic status by definition. In addition (and as expected), including covariates in the statistical analyses did not change the results of the present study. In the present study, no data on hysterectomy was gathered, however, given the rarity of the procedure and the relative young average age of the present study, it is very unlikely to effect the results. Another difference between the present study and the one of Roberts' and others (2014) regarding sample characteristics is that the average relationship duration was longer in Roberts' and others' study than in the present study. Roberts et al. (2014) did not report mean relationship duration, but a consultation with the lead author revealed that the median relationship duration was considerably longer, at 104 months ($M = 125$ months) compared to around 31 months in the present study (it should, however, be noted that relationship duration was controlled for in the analyses in the present study, as well as that of Roberts et al. [2014]). In addition, Roberts et al. (2012) found that women who were using HCs when they met their partner were less likely to

reject sex or undergo compliant sex in shorter relationships, but more likely to do so in longer relationships. Thus, it is conceivable that HC congruency effects on sexual satisfaction may take time before they manifest. However, the average relationship duration in the present study was still nearly three and a half years, which arguably is enough time for a relationship to settle: for example, a recent survey found that more than half of couples decide to get married within the first two years of dating (Francis-Tan & Mialon, 2015). Furthermore, HC congruency effects were detected in the studies of Cobey et al. (2013) and Russell et al. (2014) even though the average relationship durations were comparable to the present study, or shorter.

The present study had a large sample size and well-validated measures of sexual and relationship satisfaction. Statistical power was excellent to detect the same effect sizes as in previous studies on the topic. The present study leaves little to do better to quantitatively study the congruency hypothesis, although the association between HC use and sexual satisfaction is not particularly clear to date and should be studied more, given how large portion of women it concerns.

HC usage is an important topic for many women; for example, during the data collection phase of the present study many feedback emails were received, describing the qualitative effects of HC use on personal level. While not being any kind of scientific data, this fact still corroborates the thought that HCs have very different kinds of side effects on different people and also that many women feel the topic is important to study. On top of that, many women pointed out in their feedback that in their opinion the questionnaire was somewhat inconclusive; in these cases the women had started and stopped the use more than once during the relationship, and they had experienced drastic psychological changes in between. Although this information would not have been strictly important in the present study (as the intention was to conduct a replication as procedurally close as possible to previous studies), this highlights the importance of the qualitative aspects in studying the topic, primarily the reasons for quitting or starting the use.

To further investigate the effects of HCs on relationship-related factors and to explain the contradictional results of the previous studies, a proper longitudinal study to get information about when a woman quits HC use and when her sexual satisfaction decreases could be informative to give further understanding about the causation and possible confounding factors. Nevertheless, in the light of the present study, it seems that the

congruency hypothesis is not valid, and therefore resources should be directed in studies with a different viewpoint, rather than putting effort in finding supportive evidence to defend its existence.

REFERENCES

Cobey, K. D., Roberts, S. C., & Buunk, A. P. (2013). Hormonal contraceptive congruency: Implications for relationship jealousy. *Personality and Individual Differences*, *55*, 569-573.

Fletcher, G. J., Simpson, J. A., & Thomas, G. (2000). The measurement of perceived relationship quality components: A confirmatory factor analytic approach. Personality and Social Psychology Bulletin, 26, 340-354.

Francis-Tan, A., & Mialon, H. M. (2015). "A diamond is forever" and other fairy tales: The relationship between wedding expenses and marriage duration. *Economic Inquiry, 53,* 1919-1930.

Gangestad, S. W., & Thornhill, R. (1998). Menstrual cycle variation in women's preferences for the scent of symmetrical men. Proceedings of the Royal Society of London. Series B: Biological Sciences, 265, 927-933.

Jones, B. C., DeBruine, L. M., Perrett, D. I., Little, A. C., Feinberg, D. R., & Smith, M. J. L. (2008). Effects of menstrual cycle phase on face preferences. *Archives of sexual behavior*, *37*, 78-84.

Jones, B. C., Perrett, D. I., Little, A. C., Boothroyd, L., Cornwell, R. E., Feinberg, D. R., ... & Burt, D. M. (2005). Menstrual cycle, pregnancy and oral contraceptive use alter attraction to apparent health in faces. *Proceedings of the Royal Society of London B: Biological Sciences*, *272*, 347-354.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science, 6251, aac4716,* doi:10.1126/science.aac4716

Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K., & Minamisawa, R. (1999). Menstrual cycle alters face preference. *Nature, 399,* 741-2.

Roberts, S. C., Cobey, K. D., Klapilová, K., & Havlíček, J. (2013). An evolutionary approach offers a fresh perspective on the relationship between oral contraception and sexual desire. *Archives of sexual behavior*, *42*, 1369-1375.

Roberts, S. C., Little, A. C., Burriss, R. P., Cobey, K. D., Klapilová, K., Havlíček, J., ... & Petrie, M. (2014). Partner Choice, Relationship Satisfaction, and Oral Contraception The Congruency Hypothesis. Psychological science, 0956797614532295.

Roberts, S. C., Cobey, K. D., Klapilová, K., & Havlíček, J. (2014). Oral contraception and romantic relationships – From the lab to the real world. *Human Ethology Bulletin, 29,* 4-13.

Roberts, S. C., Klapilová, K., Little, A. C., Burriss, R. P., Jones, B. C., DeBruine, L. M., ... & Havlíček, J. (2011). Relationship satisfaction and outcome in women who meet

their partner while using oral contraception. Proceedings of the Royal Society B: Biological Sciences, rspb20111647.

Rosen, C. Brown, J. Heiman, S. Leiblum, C. Meston, R. Shabsigh, D. Ferguson, R. D'Agostino, R. (2000). The Female Sexual Function Index (FSFI): a multidimensional self-report instrument for the assessment of female sexual function. Journal of Sex &Marital Therapy, 26, 191-208.

Russell, V. M., McNulty, J. K., Baker, L. R., & Meltzer, A. L. (2014). The association between discontinuing hormonal contraceptives and wives' marital satisfaction depends on husbands' facial attractiveness. Proceedings of the National Academy of Sciences, 111, 17081-17086.

Trends in Contraceptive use worldwide 2015, a report by United Nations (p.50)

Welling, L. L., Puts, D. A., Roberts, S. C., Little, A. C., & Burriss, R. P. (2012). Hormonal contraceptive use and mate retention behavior in women and their male partners. Hormones and Behavior, 61, 114-120.

APPENDIX 1

Covering letter of the study

Hei!

Teemme tutkimusta hormonaalisten ehkäisymenetelmien vaikutuksista parisuhteeseen. Vastaajien tulisi olla parisuhteessa olevia ja vähintään 18-vuotiaita naisia. Vastauksesi ovat tärkeitä siitä riippumatta, käytätkö vai etkö käytä hormonaalista ehkäisyä. On mahdollista, että opit jotakin itsestäsi tai parisuhteestasi kyselyyn vastaamalla. Vastaamalla voit voittaa 50€ arvoisen lahjakortin Stockmannille (arvotaan 2kpl osallistujien kesken).

Noin 40% suomalaisista nuorista naisista käyttää jonkinlaista hormonaalista ehkäisyä, ja siksi on tärkeää saada tietoa ehkäisymenetelmien mahdollisista sivuvaikutuksista. Tutkimus toteutetaan Åbo Akademin ja Turun yliopiston psykologian laitosten yhteistyönä. Vastaaminen on täysin anonyymiä ja kestää noin 20 minuuttia.

Kysymyksiä ja palautetta kyselyä koskien voi lähettää sähköpostilla osoitteisiin
Tatu Erlin, PsK (tverli@utu.fi)
Janna Hujanen, psykologian ylioppilas (janhuj@utu.fi)

Tutkimusta ohjaa Patrick Jern (PsT, psykologian dosentti) ja Antti Kärnä (FT, psykologi).

Linkki kyselyyn:
http://surveyanalytics.com/t/AF6z1ZN6VA

Lämmin kiitos ajastasi!