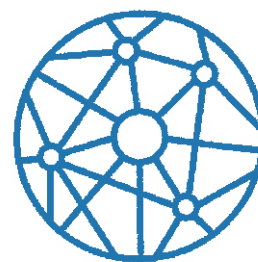


Kraków, 7<sup>th</sup> December 2025

**Evaluation of the Doctoral Dissertation**

**authored by Melanie Kowalczyk**

**The relationship between the menstrual cycle, anxiety and cognitive functioning – the moderating role of oral contraceptives**



**JAGIELLONIAN  
UNIVERSITY**

**FACULTY OF PHILOSOPHY  
INSTITUTE OF PSYCHOLOGY**

The material submitted for evaluation comprises a 27-page introduction providing a general overview of the conducted research, as well as a collection of co-authored publications, consisting of four English-language articles published in international peer-reviewed journals. Each article is accompanied by statements specifying the doctoral candidate's individual contribution according to the CRediT classification, which confirm the candidate's leading role in the conduct of all the research.

The empirical studies reported by the candidate focus on the relationship between the menstrual cycle, anxiety, and executive functioning in women who take oral contraceptives and those who cycle naturally. Two main hypotheses were tested: (1) there are differences in anxiety levels and executive functioning between women taking OC and naturally cycling women; (2) women's anxiety levels and executive functioning change through the menstrual cycle. The candidate and her co-authors address the hypotheses using three distinct yet complementary methodological approaches: a meta-analysis, a cross-sectional study, and a daily diary study. Undoubtedly, all studies presented in this thesis are novel and original. They address the scientific problem by demonstrating the differential effects of androgenic and anti-androgenic oral contraceptives on psychological functioning, thereby helping to explain inconsistencies in earlier research and offering an innovative contribution to the contemporary psychological literature.

I should preface my comments by saying that it is inherently challenging to assess papers which have already undergone a rigorous peer review process, as this requires re-evaluating decisions previously made by journal editors and their appointed reviewers. My remarks are not intended to call into question the outcomes of this process. Given the high quality of the studies, as evidenced by their publications in international peer-reviewed journals, my comments primarily concern methodological issues that may help further develop the findings presented in the thesis. In the following sections, I provide specific comments for each study, although certain remarks apply across all four papers.

**Paper-by-paper appraisal of the thesis**

**Kowalczyk, M., Kornacka, M., Wisiecka, K., Młyniec, A., Redeł, A., Szwykowska-Ziemniak, M., & Krejtz, I. (2023). The relationship between the menstrual cycle, oral contraceptives, and executive function–inhibition, updating, and shifting. *European Psychologist*, 28(4), 288-304.**

The article presents the first meta-analysis to integrate findings from studies on OCs and the three key executive functions in the Miyake & Friedman model, and it introduces a systematic assessment of study quality. Overall, the article offers a meaningful attempt to address an important scientific question: whether the menstrual cycle and oral contraceptive use are associated with differences in executive functions. The meta-analysis, which included 16 carefully selected studies, did not reveal a significant effect size, indicating no differences

**ul. R. Ingardena 6  
30-060 Kraków, Poland  
tel. +48 12 663 24 15**

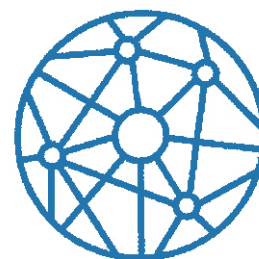
**[www.psychologia.uj.edu.pl](http://www.psychologia.uj.edu.pl)**

between OC users and naturally cycling women on the core executive functions. While the scientific value of the analysis conducted is unquestionable, several methodological issues require further clarification.

1. The heterogeneity of the included studies is very high ( $I^2 = 77.41\%$ ,  $Q(65) = 525.89$ ,  $p < .001$ ). Despite this, the conclusion that there are no differences in executive functions between OC users and naturally cycling women is stated in rather unambiguous terms. Given such substantial heterogeneity, a more appropriate interpretation would be that “the average effect is non-significant, but the results are highly inconsistent and do not permit straightforward conclusions regarding the absence of differences.” Furthermore, the authors interpret the absence of significant moderators (cycle phase, EF type) as evidence of “no moderating effect,” without explicitly acknowledging that the moderator analyses may be underpowered. It is well established that complex moderator tests in meta-analyses have limited statistical power when the number of studies is small and may therefore fail to detect true moderator effects.
2. It appears that the authors constructed their own quality assessment scale by combining elements from various sources and adding OC-specific items. Consequently, the tool is neither validated nor widely used (e.g., unlike GRADE or the Cochrane Risk of Bias instruments). They then use the mean score across the 16 items as a continuous moderator, which implicitly assumes that all items are equally important, although it is unclear whether all items should indeed carry the same weight.
3. Based on Cook’s distance, four outlying effects were identified and subsequently removed from the analysis. The authors did not report which studies or specific effects were excluded, nor did they present how the results changed following their removal. This omission reduces transparency and makes it difficult for the reader to assess the impact of these exclusions on the overall findings.
4. The authors use meta-CART to examine moderators and their interactions. However, this technique generally requires a substantial number of units, whereas only 16 studies were included in the present analysis. The absence of significant findings is interpreted as evidence of “no significant moderators,” although this may simply reflect insufficient statistical power due to the limited number of studies. Overall, the meta-analytic results suggest that the current body of research is too heterogeneous to yield a definitive conclusion.
5. Systematic review was not preregistered (however, preregistration is more common in the medical field than psychology, e.g. see: PROSPERO)

**Kowalczyk, M., Kornacka, M., Kostrzewa, Z., & Krejtz, I. (2024). Differences in anxiety, worry, and perceived stress among naturally cycling women and oral contraceptives users: a cross-sectional study investigating the role of contraceptive types. *Archives of Women’s Mental Health*, 27, 241–247.**

This study investigates the psychological correlates of different types of oral contraceptives in a large sample of women. Specifically, it examines whether the use of androgenic and anti-androgenic oral contraceptives is associated with distinct psychological profiles. This distinction between OC types enables a more nuanced understanding of the psychological effects of hormonal contraceptives and addresses an important gap that has often been overlooked in previous research. Nonetheless, several specific aspects of the study warrant further discussion.



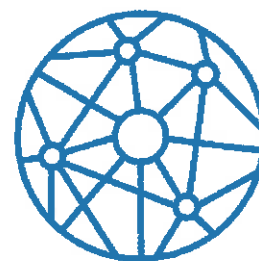
**JAGIELLONIAN  
UNIVERSITY**

**FACULTY OF PHILOSOPHY  
INSTITUTE OF PSYCHOLOGY**

ul. R. Ingardena 6  
30-060 Kraków, Poland  
tel. +48 12 663 24 15

[www.psychologia.uj.edu.pl](http://www.psychologia.uj.edu.pl)

1. The authors do not provide any clear justification for choosing the GAD-7 scale as a measure of anxiety. The GAD-7 was developed as a screening tool for detecting general anxiety disorder in primary care patients (Spitzer et al., 2006), rather than to measure anxiety. It is therefore unclear why this tool was selected instead of alternative instruments that directly assess anxiety, such as the State-Trait Anxiety Inventory (STAI).
2. Moreover, extended worry constitutes a core symptom of generalized anxiety disorder measured with the GAD-7, which suggests that these two variables may partially overlap.
3. The article lacks information on whether the authors verified the assumptions of the statistical analyses, such as distribution normality, homogeneity of variance, or covariate collinearity. The high correlation between the PSS, GAD-7, and PSWQ increases the risk of multicollinearity; without reporting VIF or tolerance values, it is unclear whether controlling for one of these variables in the ANCOVA analysis may have distorted the interpretation of the results. The analyses clearly showed that the choice of covariate changes the result: there are no differences when anxiety and age are controlled ( $p = .307$ ), but differences emerge when stress and age are controlled ( $p = .045$ ), and only then do the post-hoc tests indicate a difference between the anti-androgenic and NC groups ( $p = .023$ ). This suggests sensitivity to model specification and possible collinearity among PSS, GAD-7, and PSWQ.
4. The authors consistently report only  $F$  and  $p$  values and means; however, it is unclear how large the observed differences actually are. For example, the mean PSWQ difference between the NC group and individuals using anti-androgenic OCs is about 2 points (56.62 vs. 54.64) – likely a minimal effect. Moreover, the main conclusion (that anti-androgenic OC is associated with higher worry), is based on a single post-hoc comparison ( $p = .023$ ), with no consistency between controlling for “stress” versus “anxiety”.
5. The analyses lack tests of interaction and moderation, even though the hypotheses require them. The research question is: do the types of OCs affect the link between anxiety and its main maintenance factors: worry and perceived stress? Yet no interaction test was conducted in a regression model. Using a simple ANCOVA (with stress or anxiety as a covariate) does not actually test for interaction effects.
6. Furthermore, the study did not account for the reasons why participants chose a particular type of OC. Since anti-androgenic OCs are often used to improve appearance and reduce acne, it is worth considering whether individuals who opt for anti-androgenic rather than androgenic OCs may generally have a greater tendency to worry, which could manifest, for example, as greater concern about physical appearance. The observed difference in worry levels could therefore stem from a more fundamental personality trait rather than being a consequence of using a specific type of OC.
7. The data were collected between February and November 2022, which means that a large portion of the study took place during the COVID-19 pandemic. It may be worth considering whether this could have influenced the level of worry.
8. Interpretation of the findings may also be complicated by the different time horizons of the scales: GAD-7 (two weeks), PSS (the past month), and PSWQ (trait).



**JAGIELLONIAN  
UNIVERSITY**

FACULTY OF PHILOSOPHY  
INSTITUTE OF PSYCHOLOGY

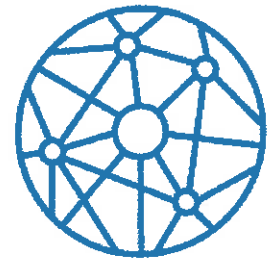
ul. R. Ingardena 6  
30-060 Kraków, Poland  
tel. +48 12 663 24 15

[www.psychologia.uj.edu.pl](http://www.psychologia.uj.edu.pl)

**Kowalczyk, M., Kornacka, M., & Krejtz, I. (2025). Anxiety and inhibitory control in women cycling naturally and women taking oral contraceptives. *Journal of Cognitive Psychology*, 1-15.**

The study compared anxiety and inhibitory control in naturally cycling women and those using oral contraceptives. The research gap is well identified: the authors point out the lack of studies comparing different types of oral contraceptives (androgenic vs. antiandrogenic) in the context of executive functions. By looking at different types of OCs, the researchers aimed to obtain a clearer understanding of how hormones might affect inhibitory control. The study used both a one-time survey and a daily diary approach, which helped capture short-term and long-term patterns. Below, I outline several comments regarding the study that the candidate may wish to consider in future research:

1. Similarly to the previous article, the rationale of the study would be strengthened by clarifying why the GAD-7 was chosen instead of instruments measuring state anxiety, such as the State-Trait Anxiety Inventory (STAI). Moreover, in both studies (cross-sectional and daily), no differences in anxiety levels were found between the NC and OC groups. It is possible that the GAD-7, developed as a screening tool for detecting general anxiety disorder, rather than to measure minor anxiety fluctuations, is not sensitive enough to detect subclinical anxiety changes in healthy subjects. The GAD-7 items refer to symptoms, such as “constant worrying” which tend to be relatively stable, therefore the tool may not capture short-term anxiety fluctuations associated with different phases of the hormonal cycle.
2. The study lacks a more detailed description of group characteristics; factors such as education level may influence both the decision to use oral contraceptives and cognitive performance. It is also not clear whether the authors controlled for other medications taken by participants, such as anxiolytics or antidepressants, which could influence both anxiety symptoms and reaction times.
3. The Emotional Stroop task was administered online, at home, and without supervision, meaning that there was no control over screen size, input device, internet latency, environmental distractions, time of day, or noise levels. All of these factors may have influenced participants' performance. In my opinion, performing the same Stroop task every day may lead to a practice effect, which was neither discussed nor controlled for (although I acknowledge that such effects are more difficult to elicit in the Stroop task than in many other cognitive tasks). Repeatedly completing the Stroop task with colors and words generally results in faster reaction times and a reduced interference effect (the difference between congruent and incongruent trials). Participants may also develop compensatory strategies, such as slightly blurring the image (for example by squinting) making the words less readable and thereby making it easier to ignore their semantic content and focus on the color instead. Moreover, repeated exposure to the same emotional words is very likely to reduce their emotional salience over time. The analyses do not model time trends across days (even though it is possible in the type of analysis used). Consequently, it remains unclear whether the observed reaction-time differences between groups and phases are stable or whether they may be partly driven by learning or habituation effects over the course of the study.
4. In study 1, it is not clear whether the ANOVA assumptions were evaluated and met. Moreover, pairwise comparisons involving the smallest group (N = 16) are



**JAGIELLONIAN  
UNIVERSITY**

**FACULTY OF PHILOSOPHY  
INSTITUTE OF PSYCHOLOGY**

**ul. R. Ingardena 6  
30-060 Kraków, Poland  
tel. +48 12 663 24 15**

**[www.psychologia.uj.edu.pl](http://www.psychologia.uj.edu.pl)**

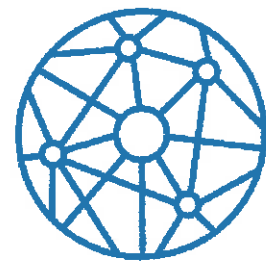
likely underpowered to detect anything other than moderate effects, which should be acknowledged when interpreting the results.

5. In Study 2, missing data are present in the EMA dataset. The paper does not examine whether this “missingness” is related to daily anxiety, group membership, or cycle phase. Because HLM assumes that data are Missing At Random, it is important, for example, to know whether participants may have been less likely to complete assessments on days characterized by higher anxiety, as such patterns could bias the results.
6. In the diary results, anti-androgenic OC users are slower than both androgenic OC users and NC women not only for fear-related words, but also for negative, positive, neutral, and color words. This pattern looks more like a general psychomotor or attentional slowing than a specific deficit in emotion-related inhibition.

**Kowalczyk, M., Kornacka, M., & Krejtz, I. (2025). Anxiety and depression in women cycling naturally and taking oral contraceptives—a daily diary study. *European Psychiatry*, 68(S1), S1204-S1205.**

The article examines daily fluctuations in anxiety, depression, perseverative cognition, and perceived stress in women who either cycle naturally or use different types of oral contraceptives. A particularly valuable part of this study is that it uses different ways of measuring the data, combining daily reports from each woman with comparisons between groups. This multi-level approach helps capture both day-to-day emotional changes and more stable differences between people, giving a clearer and more complete picture of how contraceptive use may relate to women’s everyday emotional experiences. Despite the study’s innovative approach and valuable findings, several issues still require clarification or further refinement in future research.

1. Key concepts in the manuscript would benefit from clearer conceptual delineation. The term *perseverative cognition* is used interchangeably with *worry* and *rumination*, despite these constructs representing distinct psychological processes with different theoretical backgrounds and empirical profiles.
2. As in the previous study, the assessment of daily anxiety is based on three GAD-7 items that do not specifically measure anxiety but rather reflect the symptom cluster associated with generalized anxiety disorder (e.g., feeling nervous, being on edge, worrying). Therefore, the score may not necessarily reflect the level of anxiety. Furthermore, the questionnaire, especially the three-item version, may not be sensitive enough to detect between-group differences.
3. Daily depression was assessed with three items that largely tap general self-evaluation and well-being. Daily self-esteem and daily satisfaction with life were then assessed as separate constructs using highly similar content, resulting in substantial conceptual and item overlap. Similarly, the daily anxiety items include “worrying too much about different things,” while perseverative cognition and worry are measured separately with overlapping themes. Such redundancy is likely to inflate the correlations among these variables, making the reported “significant links” between anxiety, depression, and perseverative cognition at least partly tautological.
4. A significant limitation is the absence of control for key confounding variables such as age at OC initiation, duration of OC use, ethinyl estradiol dosage, psychiatric diagnosis and treatment, substance use, psychiatric and medical



**JAGIELLONIAN  
UNIVERSITY**

FACULTY OF PHILOSOPHY  
INSTITUTE OF PSYCHOLOGY

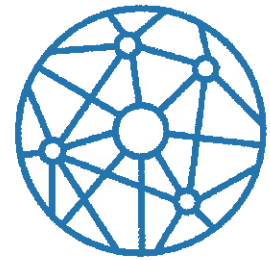
ul. R. Ingardena 6  
30-060 Kraków, Poland  
tel. +48 12 663 24 15

[www.psychologia.uj.edu.pl](http://www.psychologia.uj.edu.pl)



comorbidities. These factors strongly influence mood and could easily overshadow any cycle-related effects, making it unclear whether the reported differences are attributable to contraceptive use or to uncontrolled confounders. The study also lacks information on the reasons for OC use, such as PCOS, acne, or PMS, which are themselves associated with mood disturbances and emotional dysregulation.

5. Although the authors describe their work as an “exploratory diary study” in the Statistical Analysis section, they formulate three explicit hypotheses concerning daily levels of the studied variables. Typically, exploratory studies do not advance directional hypotheses, particularly when prior findings are mixed or inconclusive, as is the case in this research area. Moreover, the hypotheses presented do not clearly emerge from the literature review, which further underscores the inconsistency between the exploratory framing of the study and the confirmatory nature of the stated predictions.
6. I also have some concerns regarding the formulation of Hypothesis 4. As currently phrased, the hypothesis essentially states that OC users “do not experience significant fluctuation” in anxiety, depression, and related factors across the menstrual cycle. This makes the research hypothesis equivalent to a “no effect” or “no difference” statement. Within the standard NHST framework, failing to obtain a statistically significant effect (e.g., a non-significant effect of cycle day in the hierarchical linear model) does not provide positive evidence for the absence of fluctuation; it may simply reflect limited statistical power, measurement error, or other sources of variability. If the authors’ theoretical position were genuinely that OC users do not exhibit meaningful cycle-related changes, this claim would be more appropriately evaluated with methods designed to test for equivalence or absence of an effect (e.g., equivalence testing, Bayesian approaches that permit assessment of evidence for the null, or specifying a smallest effect of interest and demonstrating that the confidence intervals fall within that range).
7. The authors acknowledge that no group differences emerged in daily worry, anxiety, or perceived stress, however, they do not offer any substantive explanation for these null findings. Their discussion merely notes inconsistencies in the existing literature but does not address methodological or theoretical factors that could account for the absence of effects in the present study.
8. The article does not report measures of model adequacy. Without these diagnostics, readers cannot determine whether the selected models were the best-fitting or even appropriate for the data structure.
9. The manuscript does not examine whether missingness in data is related to mood (e.g., participants skipping on worse days) or group status. HLM can handle missing data under “Missing at Random” assumptions, but those assumptions are not tested or discussed.
10. Although the theoretical framework suggests that emotional states may vary as a function of the interaction between menstrual-cycle phase and oral contraceptive type, none of the Hierarchical Linear Models included a phase  $\times$  OC-type interaction term. Instead, phases and groups are analyzed separately, which prevents a direct examination of whether the effect of phase differs across contraceptive types. As a result, the analyses do not actually evaluate the study’s core predictions, and conclusions regarding phase-dependent differences in emotional functioning among OC users versus naturally cycling women remain unsubstantiated.



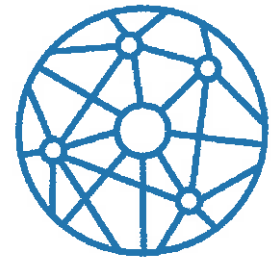
**JAGIELLONIAN  
UNIVERSITY**

FACULTY OF PHILOSOPHY  
INSTITUTE OF PSYCHOLOGY

ul. R. Ingardena 6  
30-060 Kraków, Poland  
tel. +48 12 663 24 15

[www.psychologia.uj.edu.pl](http://www.psychologia.uj.edu.pl)

11. Although daily stress was measured as part of the diary protocol, it was not included as a covariate or predictor in any of the hierarchical linear models. This omission is problematic, given that stress is theoretically and empirically linked to anxiety, depressive symptoms, and perseverative cognition, and the authors themselves highlight its centrality in the conceptual framework. Importantly, the descriptive statistics show that daily stress exhibits the largest within-person variance of all measured variables, indicating that day-to-day fluctuations in stress are substantial and could meaningfully affect momentary emotional states. Failing to statistically control for this highly variable construct means that observed between-group differences and within-person associations may be confounded by unaccounted fluctuations in daily stress. Consequently, the interpretability and internal validity of the reported effects are weakened, as it remains unclear whether the identified patterns reflect genuine influences of oral contraceptive use or simply variability in stress exposure across days.
12. Some previous studies indicate that premenstrual disorders, which include symptoms of anxiety and depression, became more prevalent during the pandemic, particularly among young women (Puthusserry & Delariarte, 2022; Shebl et al., 2022). Thus, as in the case of the second study, the levels of depression, stress, and anxiety reported by participants could have been influenced by the fact that a large portion of the research was conducted during the COVID-19 pandemic.
13. My general impression is that the empirical articles (with the exception of the systematic review) appear to be based on a single study: the 2024 article analyzes data from the screening phase, whereas the 2025 articles draw on data from the main study. This is not necessarily a methodological flaw; indeed, the project is extensive, involves numerous variables, and requires a large participant sample, making it difficult to conduct the studies independently. However, in the discussion section the authors write, for example: "As a result, there is a lack of studies on these topics, and the results are contradictory. A study conducted by Kowalczyk et al. [35] concluded that women taking anti-androgenic OC had significantly higher levels of worry than NC women," and "The lack of differences between OC users and NC women on daily levels of anxiety is in line with certain studies conducted previously [35]." It appears that the authors are using findings from the screening phase (reported in one article) to support conclusions presented in the later article, even though both sets of findings originate from the same overarching study. If so, this should be clearly acknowledged to avoid giving the impression of independent replication where none exists.



**JAGIELLONIAN  
UNIVERSITY**

**FACULTY OF PHILOSOPHY  
INSTITUTE OF PSYCHOLOGY**

### **Assessment of candidate's ability to conduct research independently**

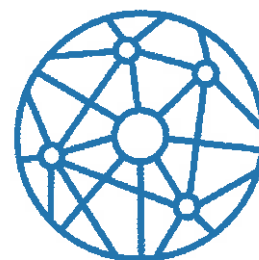
While some methodological limitations are to be expected given the scale and complexity of the project, the PhD candidate's research skills are nevertheless of a very high standard. The body of work presented across the four articles demonstrates that the candidate possesses a high level of competence in the design, execution, and interpretation of psychological research.

**ul. R. Ingardena 6  
30-060 Kraków, Poland  
tel. +48 12 663 24 15**

**[www.psychologia.uj.edu.pl](http://www.psychologia.uj.edu.pl)**

### **Assessment of candidate's general theoretical knowledge in psychology**

The candidate demonstrates comprehensive theoretical knowledge in the field of psychology, particularly within the domains of anxiety, executive functions, cognitive models, hormonal influences, women's mental health, and methodological approaches in psychological science. All empirical studies and the meta-analysis show that she can evaluate, critique, and synthesize diverse research findings, revealing an advanced understanding of theoretical debates, methodological limitations, and gaps in the field.



### **Assessment of novelty and contribution to the field**

Based on the four articles included in the PhD dissertation, I conclude that the candidate has clearly and convincingly addressed a significant scientific problem concerning the relationship between women's emotional and cognitive functioning and the use of hormonal contraception. The author not only identified key shortcomings in previous research but also systematically addressed them in her own studies: she employed large samples, differentiated between types of contraception, and applied both cross-sectional and diary methodologies. Consequently, the dissertation not only fills an important gap in the literature but also provides a structured and reliable foundation for future research on the influence of hormones on women's emotional and cognitive functioning.

**JAGIELLONIAN  
UNIVERSITY**

**FACULTY OF PHILOSOPHY  
INSTITUTE OF PSYCHOLOGY**

### **Conclusion**

To summarize, the dissertation demonstrates the candidate's broad theoretical knowledge and confirms her ability to conduct original research independently. The dissertation constitutes an original contribution to the scientific problem under investigation. The doctoral dissertation fulfills the requirements stipulated by the Act, in accordance with Article 187, points 1 and 2.

dr hab. Katarzyna Prochwicz, prof. UJ

ul. R. Ingardena 6  
30-060 Kraków, Poland  
tel. +48 12 663 24 15

[www.psychologia.uj.edu.pl](http://www.psychologia.uj.edu.pl)