

Registration form

1a. Details of applicant

-Name, title(s): Dudley Díaz, Emma Klaasse

-University, Department (or Institute): Vrije Universiteit Amsterdam

-E-mail(s): klaasse.emma@gmail.com, deds15@gmail.com

1b. Title of research proposal

The effects of exercise on perceived stress among Canadian university students

1c. Abstract

This study investigates the impact of physical exercise on perceived stress levels among Canadian university students using national survey data and a Coarsened Exact Matching (CEM) technique. The participants were matched on covariates like age, sex, ethnicity, international student status, and study program. The treatment variable, defined as higher levels of exercise based on the Godin-Shephard Leisure-Time Physical Activity Questionnaire, was compared to lower levels of exercise. The Perceived Stress Scale measured perceived stress, which was compared between high and low exercise groups. Results from an Ordinary Least Squares (OLS) regression show that higher exercise levels correlate with a statistically significant decrease in perceived stress. Limitations include potential reversed causality, inability to control all confounding variables, and biases due to self-selection in sampling. These findings suggest that while exercise is associated with reduced stress, further research is needed to confirm causality and address these limitations.



Word Count: 145

Research proposal

2a. Description of the proposed research and societal results

Introduction and research question

University students often endure significant stress from academic demands and uncertainties about post-graduation plans, among other factors (Beiter, Nash et al., 2015). Although stress is an unavoidable part of life, managing this stress is crucial for their mental health and overall well-being. According to Aggarwal et al. (2014) stress can result in reduced learning and memory functions. One promising strategy is physical exercise, which has proved to provide numerous health benefits, including the reduction of stress (Salmon, 2001; Klaperski & Fuchs, 2021). Physical exercise is a well-documented method for improving mental health, offering benefits such as reduced symptoms of anxiety and depression, and enhanced mood (Salmon, 2001). Despite these known advantages, there is a need to better understand how exercise specifically affects stress levels in university students. Understanding this relationship could enable universities to better cater to the needs of their students. Despite these well-documented advantages, there remains a need to understand how exercise specifically impacts stress levels in university students.

Most existing studies have explored this relationship using Randomised Controlled Trials (RCTs) (Zieff et al., 2022). While valuable, RCTs can be expensive and are sometimes limited by ethical and logistical constraints. In contrast, using national survey data offers a cost-effective alternative that allows



researchers to analyse a larger and more diverse sample of university students. However, current studies that utilise survey data often fail to employ quasi-experimental statistical methods, which are essential for isolating causal effects. Such methods control for potential confounding factors like age and sex. This study aims to fill this gap by employing robust matching techniques to isolate the effect of exercise on perceived stress levels. Therefore, this study will answer the following research question: "What is the effect of physical exercise on perceived stress levels among university students?"

Theoretical framework and hypotheses

This research is grounded in the biopsychosocial model, which suggests that biological, psychological, and social factors interact to influence health and well-being. This model, proposed by Engel (1977), remains relevant today despite certain limitations (Lugg, 2022). Applying this framework to physical exercise, physical exercise is expected to affect psychological outcomes like stress levels through various mechanisms. Biologically, exercise triggers the release of endorphins, improves sleep quality, and enhances self-esteem, which in turn could reduce stress. Psychologically, it could provide a distraction from stressors. Socially, exercise often involves group activities, which could facilitate social interactions and foster a sense of community and belonging, which in turn would also lower stress levels.

Based on the theoretical framework and existing literature, we formulated the following hypotheses:

- *H0* (null hypothesis): There is no difference in perceived stress levels between students who engage in more exercise compared to students who engage in less exercise.
- H1 (alternative hypothesis): Students who engage in more exercise have different perceived stress levels compared to students who engage in less exercise.

Societal importance

Understanding the relationship between exercise and stress can have significant implications for student support services. If exercise is found to reduce stress effectively, universities could promote physical activity as a key component of student well-being programs. This could involve creating more accessible fitness facilities, organising group sports activities, and integrating physical activity into daily student life. By doing so, universities can help students manage stress better, thereby enhancing their academic performance and quality of life. Overall, the study aims to contribute valuable insights into how lifestyle interventions can mitigate stress, ultimately supporting students in achieving academic success and maintaining mental health.

2b. Approach: what is your methodological or experimental approach?

Data used

This study uses survey cross-sectional survey data collected from Canadian university students between September 22 and October 30, 2020 (Reeves & Paterson, 2022). Participants were recruited through social media, REACH BC, the Canadian Psychological Association's (CPA), and the Recruit Research Participant Portal (R2P2), resulting in 1,192 respondents. It is important to note that the survey period coincided with the COVID-19 pandemic, which may have influenced participants' stress levels. This context will be considered when interpreting the study's findings.

Treatment and control variables

The research employs a matching technique to investigate the effect of exercise on perceived stress levels. Participants are matched based on age, sex, ethnicity, study program, and international student status to form two groups:

- 1. Treatment Group: Students engaging in higher levels of physical exercise, as determined by the Godin-Shephard Leisure-Time Physical Activity Questionnaire.
- 2. Control Group: Students engaging in lower levels of physical exercise.

The definition of the treatment variable, higher levels of exercise, is defined based on whether a person is considered 'active' on the Godin-Shephard Leisure-Time Physical Activity Questionnaire. In his paper, Godin (2011) argues that, based on the outcomes of the Godin-Shephard Leisure-Time Physical Activity Questionnaire, individuals can be classified as 'active' or as 'insufficiently active'. This distinction is based



on the amount of moderate and strenuous activities or 'units' per week of an individual. If a person engages in 24 or more moderate to strenuous units of exercise per week that person is classified as 'active', and if a person does less, they are classified as 'insufficiently active'. Amireault and Godin (2015) show that this distinction is valid, as people that are classified as 'active' have a significantly better aerobic capacity and a lower body fat percentage.

The primary outcome variable, perceived stress levels, is measured using the Perceived Stress Scale (Cohen, Kamarch & Mermelstein, 1983). The Perceived Stress Scale is used to measure individuals perceived stress levels. In this questionnaire, consisting of 10 questions, individuals are asked to reflect on stressful situations in the past month. Based on these questions a final score is computed which indicates the stress levels of an individual. With the minimum score being 5 and the maximum score being 50.

Matching techniques

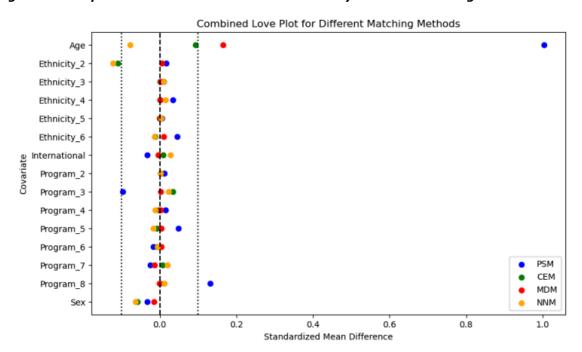
In order to improve causal inference, this research uses matching as a statistical technique to create a balanced sample of treated and control groups. There are different types of matching techniques. Some commonly used techniques include:

- (Coarsened) Exact Matching (CEM). This method ensures that the groups are comparable based on selected covariates, minimising bias and improving the reliability of causal estimates. CEM is different from exact matching, as it does not exactly match observations, but it temporarily coarsens the covariates on which the observations will be matched.
- Propensity Score Matching (PSM) calculates the probability (propensity score) that a participant receives a treatment based on observed characteristics, PSM matches individuals with similar scores from treatment and control groups. However, King and Nielsen (2019) find that using propensity scores for matching might actually increase imbalance instead of decreasing it.
- Mahalanobis Distance Matching: This is a statistical technique used to pair treated and control groups by measuring the distance between their covariates. This method accounts for the correlation between variables, ensuring that matches are based on overall similarity rather than individual covariate values.
- Nearest Neighbour Matching (NNM) pairs each observation in a treatment group with the most similar observation in a control group based on covariates.

Figure 1 shows a love plot that compares the balance of these four matching techniques. Overall, CEM achieves the best balance across all covariates compared to the other matching methods. It has the smallest absolute standardised mean difference.



Figure 1. Love plot for Standardised Mean Difference by different matching methods



Therefore, in this study, CEM is used to assess the effect of higher levels of exercise on perceived stress among university students. The treatment group consists of students with an activity score greater than 24, as measured by the Godin-Shephard Leisure-Time Physical Activity Questionnaire (Godin, 2011). The outcome of interest is the average perceived stress score, assessed using the Perceived Stress Scale (PSS; Cohen, Kamarch & Mermelstein, 1983). It is considered that SMD of ± 0.1 is indicative of good balance Austin (2009), Stuart (2010).

Coarsened Exact Matching (CEM)

To ensure the validity of the CEM, two key assumptions must hold. First, there must be selection on observables. This means that the treatment assignment (exercise) is independent of the potential outcome (stress levels). If this is not the case, there might be confounding variables influencing the analysis, resulting in biased estimates. Second, the common support assumption must hold, which means that for every value of the covariate, there is at least one unit in both the treatment and control groups. If this assumption is violated, certain groups of the population might be excluded from the analysis, which limits the generalizability of the research. The following matching covariates are chosen to control for potential confounding variables that might influence both the treatment (exercise level) and the outcome (perceived stress):

- Age: reflects different life stages in which exercise and stress perceptions might differ;
- Sex: accounts for biological differences between men and women in stress response and exercise habits;
- Ethnicity: considers cultural variations in stress experiences and exercise participation;
- International student status: differentiates being an domestic or international student, as international students may face unique stressors compared to domestic students;
- Study program: different programs may have varying levels of stress.

These covariates are independent of the treatment. This means that an individual's level of exercise won't change their age, sex, ethnicity, international student status, or study program. These characteristics are fixed and unaffected by the treatment.

CEM requires the covariates to be coarsened. This means that the variables have to be changed into

meaningful categories. For the covariates used in this research, it means:

- Age: grouped into categories
 - Bin Ranges:
 - First Bin: Ages from 16 to 19.
 - Second Bin: Ages from 20 to 23.
 - Third Bin: Ages from 24 to 27.
 - Fourth Bin: Ages from 28 to 31.
 - Fifth Bin: Ages from 32 to 35.
 - Sixth Bin: Ages from 36 to 39.
 - Seventh Bin: Ages from 40 to 43.
 - Eighth Bin: Ages from 44 to 47.
 - Ninth Bin: Ages from 48 to 50.
 - Tenth Bin: Ages from 51 to 54.
 - Eleventh Bin: Ages from 55 to 58.
- Sex: binary classification (male/female)
- Ethnicity: broad categories of ethnicity (Caucasian, Asian, Indigenous, Hispanic/Latino, Black, other). Each ethnicity category was recoded to a dummy variable that equals 1 if an individual has a certain ethnicity, and 0 otherwise. The original dataset also included the category 'Metis'. This category was added to the 'Indigenous' category. A point of concern is that we do not know which ethnicities are included in the category 'other'.
- International student status: binary dummy variable (yes/no)
- Study program: broad study programs (business, education, engineering, fine arts, human and social development, humanities, sciences, social sciences). Each of these categories was transformed to a dummy variable that equals 1 when an individual follows a specific program and 0 otherwise.

After the variables have been coarsened, the observations are exactly matched based on this coarsened data. After matching, the balance of the covariates is checked to ensure that the treatment group and the control group are comparable using a love plot (as presented in figure 1).

A priori power analysis

CEM can cause certain observations to drop after matching. This happens when they do not have an exact (coarsened) match. In order to estimate the minimum sample size needed for the analysis, we conducted an a priori power analysis. We set the effect size to 0.2, and the power to 0.8. We assumed an allocation rate of 1 between the treatment and control group (N2/N1). Table 1 shows that for this analysis, we would need a total sample size of 788 individuals after matching to obtain a power of 0.8.

Table 1. A priori power analysis

Power analysis parameters	
Input	
Tails	two
Effect size d	0.225
α-error probability	0.05
Sample size group 1	531
Sample size group 2	517
Allocation ratio N2/N1	1.03
Output	
Noncentrality parameter	3.644
Critical t	1.962
Df	1046
Power (1-β error probability)	0.954

Robustness checks



Lastly, we calculate the Rosenbaum Bounds to evaluate the potential impact of unmeasured variables on the observed outcomes. Rosenbaum Bounds are a method used to assess the sensitivity of causal inference results to hidden biases in observational studies. They estimate how much an unobserved variable would need to influence the treatment assignment to invalidate the study's conclusions. By calculating the extent to which hidden confounders could alter the treatment effect, Rosenbaum Bounds help determine the robustness of the findings.

2c. Provisional results: What are your findings?

Descriptive statistics

This section provides an overview of the sample characteristics after applying matching techniques. We focus on our matching covariates sex, age, study program, ethnicity, and international student status. Table 1.1 and 1.2 show an overview of the covariates that were used in the model, and their corresponding mean and standard deviation before and after matching. The tables show that CEM generally improved the balance between the treated and the control groups for most variables. Some covariates still show slight differences that remain after matching, such as the variable sex and the dummy indicating students of Asian ethnicity. Another noticeable difference is that after matching, there are no more students present in the sample who are of Indigenous ethnicity.

Table 2.1. Descriptive statistics of the matching covariates before matching (N=1192)

Matching covariates (N=1192)	Treated mean	Treated standard deviation	Control mean	Control standard deviation
Age	20.23	3.17	20.27	2.71
Sex (1=man, 2=woman)	1.78	0.41	1.84	0.37
International status	1.93	0.26	1.91	0.29
Ethnicity: Caucasian	0.68	0.39	0.57	0.46
Ethnicity: Asian	0.19	0.39	0.30	0.46
Ethnicity: Indigenous	0.01	0.11	0.00	0.06
Ethnicity: Hispanic/Latino	0.02	0.15	0.01	0.09
Ethnicity: Black	0.02	0.13	0.02	0.12
Ethnicity: Other	0.08	0.28	0.10	0.30
Study: business	0.09	0.17	0.10	0.16
Study: education	0.03	0.17	0.03	0.16
Study: engineering	0.11	0.31	0.08	0.28
Study: fine arts	0.02	0.16	0.04	0.19
Study: human and social development	0.02	0.14	0.03	0.17
Study: humanities	0.07	0.26	0.08	0.27
Study: sciences	0.41	0.49	0.40	0.49
Study: social sciences	0.25	0.43	0.24	0.43

Table 2.2. Descriptive statistics of the matching covariates before matching (N=1043)



Matching covariates	Treated mean	Treated standard	Control mean	Control standard
(N=1043)	_	deviation		deviation
Age	19.97	2.05	19.87	1.93
Sex (1=man, 2=woman)	1.81	0.39	1.87	0.34
International status	1.95	0.21	1.95	0.23
Ethnicity: Caucasian	0.73	0.39	0.61	0.46
Ethnicity: Asian	0.19	0.39	0.30	0.46
Ethnicity: Indigenous	0.00	0.00	0.00	0.00
Ethnicity: Hispanic/Latino	0.00 (0.002)	0.04	0.00 (0.002)	0.04
Ethnicity: Black	0.01	0.09	0.01	0.09
Ethnicity: Other	0.07	0.28	0.08	0.30
Study: business	0.09	0.16	0.11	0.15
Study: education	0.03	0.16	0.02	0.15
Study: engineering	0.11	0.31	0.07	0.28
Study: fine arts	0.02	0.14	0.02	0.19
Study: human and social development	0.02	0.08	0.01	0.17
Study: humanities	0.07	0.25	0.07	0.27
Study: sciences	0.41	0.50	0.44	0.49
Study: social sciences	0.25	0.44	0.26	0.43

Post-hoc power analysis

After matching, the sample size decreased from 1192 to 1043 individuals, 530 in the treatment group and 513 in the control group. We conducted a post-hoc power analysis to analyse the implications of this sample reduction for the power of the model. This analysis shows that the model has a power of 0.897. This means the model has high enough power to reject a false null hypothesis.

Table 3. Post hoc power analysis

Power analysis parameters	
<u>Input</u>	
Tails	two
Effect size d	0.2
α-error probability	0.05
Sample size group 1	530
Sample size group 2	513
Allocation ratio N2/N1	1.03
Output	
Noncentrality parameter	3.229
Critical t	1.962
Df	1041
Power (1-β error probability)	0.897

Results

To estimate the causal effects of physical exercise on stress levels we conducted an OLS regression on the matched sample. Table 2 presents the results of a simple regression without the control variables, to



show the effect of the individual predictor. The results show that the treatment, exercise, decreases the stress levels by 1.59 units on the Perceived Stress Scale. The effect is significant on the 5% significance level (p<0.05). However, the R^2 and the adjusted R^2 are 0.013 and 0.012 respectively, indicating that only 1% of the variance in the dependent variable is explained by the independent variables included in the model.

Table 4. OLS regression results without control variables (CEM)

OLS regression results	Coefficient	Standard error	P-value
Treatment: exercise	-1.5920*	0.302	0.000
Constant	30.2807*	0.423	0.000
\mathbb{R}^2	0.013		
Adjusted R ²	0.012		
F-statistic	14.16		
N	1043		

Note: * = significant at a 5% significance level

Table 3 presents the results of a multiple regression, including the control variables. It shows that higher exercise correlates with a -1.51 lower perceived stress score after controlling for age, sex, international status, ethnicity and study program. The result of the treatment remains significant on the 5% significance level (p<0.05). Based on this result, we can reject the null hypothesis that states that physical exercise has no effect on perceived stress levels. The results suggest that higher exercise levels are associated with lower perceived stress.

Additionally, we find that biological sex (woman/man) also significantly influences the perceived stress levels. In this model, being a woman increases perceived stress by 2.87 units on the Perceived Stress Scale. This effect is significant on the 5% significance level. Being a social science student also seems to increase stress levels by 1.78 units on the Perceived Stress Scale. Despite being matched for, the variable sex and the variable study program (social sciences) still show a significant effect on the dependent variable. This suggests that they independently influence the outcome beyond the matching process. Including these variables in the regression model helps account for any residual confounding, ensuring more precise estimates and reinforcing the robustness of the findings. However, the R² and the adjusted R² are still low, 0.056 and 0.042 respectively, indicating that only 4% to 6% of the variance in the dependent variable is explained by the independent variables included in the model.

Table 5. OLS regression results including control variables (CEM)



OLS regression results	Coefficient	Standard error	P-value
Treatment: exercise	-1.509*	0.423	0.000
Age	-0.016	0.108	0.886
Sex	2.866*	0.615	0.000
International status	-0.531	0.982	0.560
Ethnicity: Asian	-0.876	0.513	0.088
Ethnicity: Hispanic/Latino	1.585	4.819	0.742
Ethnicity: Black	-0.336	2.410	0.889
Ethnicity: Other	-0.666	0.823	0.419
Study: education	1.817	1.526	0.234
Study: engineering	0.020	1.062	0.985
Study: fine arts	2.259	1.595	0.157
Study: human and social development	1.070	2.264	0.636
Study: humanities	1.412	1.097	0.198
Study: sciences	0.533	0.808	0.509
Study: social sciences	1.781*	0.857	0.038
Constant	25.755*	3.219	0.000
R ²	0.056		
Adjusted R ²	0.042		
F-statistic	4.075		
N	1043		

Robustness checks

Table 6 shows the results of the Rosenbaum Bounds sensitivity analysis. The analysis shows to what extent we can violate the selection-on-observables assumption and still be confident about the estimate of the treatment effect. The analysis shows that the treatment effect remains consistent and statistically significant (p<0.05) even when the odds of treatment assignment due to unobserved factors increases by a factor of 6.

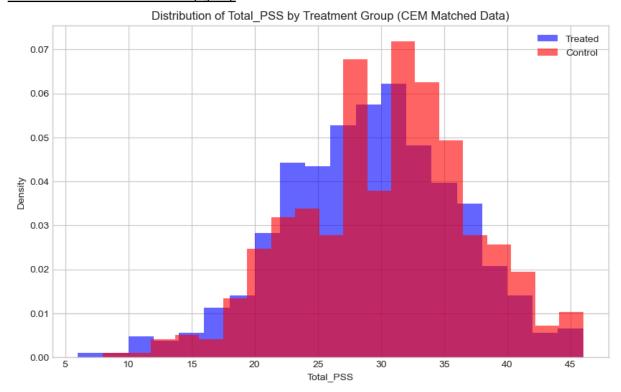
Table 6. Sensitivity analysis with Rosenbaum Bounds

Gamma	Lower bound	Upper bound
1	0.1587	0.1587
2	0.0786	0.2398
3	0.0416	0.2819
4	0.0228	0.3085
5	0.0127	0.3274
6	0.0072	0.3415

Note: Gamma is odds of differential assignment to treatment due to unobserved factors



Observation of the distribution by group



We also compared how the stress levels were distributed over both groups. The significant overlap and similar spread between the distributions of the outcome variable (PSS score) for the treated and control groups suggest that CEM has effectively balanced the groups. Also, the slight shift towards lower PSS score values for the treated group compared to the control group aligns with the negative treatment effect observed in the regression analysis. This shift indicates that the treatment is associated with a reduction in perceived stress scores. Finally, the similar densities and spread across a range of PSS score values indicate that the matching process has created comparable groups, which helps ensure that any observed treatment effect is due to the treatment itself rather than pre-existing differences between the groups.

<u>Limitations</u>

Even though the results seem to indicate a robust positive effect of exercise on perceived stress levels in Canadian university students, there are a couple of limitations to consider. Because of these limitations, our results should be interpreted with caution.

First, there is the issue of reversed causality. It is possible that stress levels could also influence people's willingness to exercise. For example, people that experience high stress might also exercise less. This means there is a possible risk of endogeneity in this causal relationship. For future research, this endogeneity could be removed by introducing an instrumental variable that is independent of the outcome (perceived stress levels), but does influence the treatment variable (physical exercise). A potential instrumental variable could be the amount of fitness facilities that are present at a university. Unfortunately, our dataset did not include any specific details about which university the students belonged to, which meant we could not create a variable that captured this data.

Second, there were some potential confounding variables that were found in the literature that we could not control for. In some cases because we did not have the data. For example, Eisenberg et al. (2007) found that students with lower socioeconomic status are more likely to suffer from stress and mental



health problems. In other cases we could not control for the confounding variables because they were not independent from the treatment assignment. For example, poor sleeping habits can affect both stress levels and exercise levels, according to Orzech et al. (2011). When people start exercising more, this will affect their sleeping patterns. Therefore, using sleeping did not satisfy the 'selection on observables' assumption of matching.

Another limitation of this research is that the sampling method of the research was based on self-selection, and therefore not random. The students were recruited on multiple platforms, but they could themselves choose whether or not they wanted to participate. This meant they were able to opt-in based on their own interest or availability. This might have led to a sample that is biassed towards those who are more engaged or interested in the subject of mental health.

Lastly, the survey was conducted during the COVID-19 period, which might have affected the stress levels of the university students. This might affect the external validity of the research, as the outcomes have limited generalisability to students that were studying in other time periods.

Conclusion

The analysis supports the hypothesis that increased physical activity significantly reduces perceived stress scores. The results seem robust to both observed and unobserved confounding as evidenced by the CEM and the Rosenbaum sensitivity analysis. Coarsened Exact Matching has effectively balanced the treated and control groups, leading to reliable estimates of the treatment effect.

These findings suggest that interventions aimed at increasing physical activity can be effective in reducing perceived stress among university students. Therefore, programs and policies promoting physical activity could be beneficial for reducing their stress levels.

Future Work

- Further Research: Future studies could explore the long-term effects of physical activity on stress and other mental health outcomes.
- External Validity: The results on this research should be validated over other circumstances to ensure the hold on other environments
- Additional Covariates: Including more covariates or exploring different matching methods could further validate the robustness of the findings.

Word Count (Sections 2a - 2c): 3099

References

- Amireault, S. & Godin, G. (2015). The Godin-Shephard Leisure-Time Physical Activity Questionnaire: Validity Evidence Supporting its Use for Classifying Healthy Adults into Active and Insufficiently Active Categories. Perceptual and motor skills. 10.2466/03.27.PMS.120v19x7.
- Beiter, R., Nash, R., McCrady, M., Rhoades, D., Linscomb, M., Clarahan, M., & Sammut, S. (2015). The prevalence and correlates of depression, anxiety, and stress in a sample of college students. Journal Of Affective Disorders, 173, 90–96. https://doi.org/10.1016/j.jad.2014.10.054
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. Journal of Health and Social Behavior, 24(4), 385–396. https://doi.org/10.2307/2136404
- Engel G, (1977). The need for a new medical model: a challenge for biomedicine. Science, 196, 129–136.



- King, G. and Nielsen, R. 2019. "Why Propensity Scores Should Not Be Used for Matching." Political Analysis, 27, 4, Pp. 435-454. Publisher's Version Copy at https://tinyurl.com/v5b5vixo
- Godin, G. (2011). The Godin-Shephard Leisure-Time Physical Activity Questionnaire. The Health & Fitness Journal of Canada, 4(1), 18–22. https://doi.org/10.14288/hfjc.v4i1.82
- Klaperski, S., & Fuchs, R. (2021). Investigation of the stress-buffering effect of physical exercise and fitness on mental and physical health outcomes in insufficiently active men: A randomized controlled trial. Mental Health and Physical Activity, 21. https://doi.org/10.1016/j.mhpa.2021.100408
- Lugg, W. (2022). The biopsychosocial model history, controversy and Engel. Australasian Psychiatry, 30(1), 55-59. doi:10.1177/10398562211037333
- Reeves, J. T. H., Paterson, T. S. E. (2022). University Student Mental Health [Data set]. Scholars Portal Dataverse.
- Salmon, P. (2001). Effects of physical exercise on anxiety, depression, and sensitivity to stress: A unifying theory. Clinical Psychology Review, 21(1), 33–61. https://doi.org/10.1016/S0272-7358(99)00032-X
- Zieff, G. H., Stoner, L., Frank, B., Gaylord, S., Battle, S., & Hackney, A. C. (2022). Aerobic exercise, mindfulness meditation, and stress-reduction in high-stress, college-based young adults: A pilot study. Journal of American College Health, 72(5), 1331–1335. https://doi-org.vu-nl.idm.oclc.org/10.1080/07448481.2022.2076103