PARTICIPANTS

The participants studied at three music academies in Poland: the Stanisław Moniuszko Academy of Music in Gdansk, the Feliks Nowowiejski Academy of Music in Bydgoszcz, and the Department of the Fryderyk Chopin University of Music in Białystok. A total of 85 participants were studying instrumental performance (39.9%), 58 music education (27.2%), 20 solo singing (9.4%), ten musical theatre (4.7%), nine studied conducting (4.2%), six church music (2.8%), six sound engineering (2.8%), six music theory (2.8%) and the remaining 13 were taking a variety of other courses (6.1%). In Poland, all music academy students take courses closely related to performance, such as classical piano, their own choice of one or more orchestral instruments, piano improvisation, accompaniment, or vocal studies. These courses usually end with a formal examination or public performance. As such, all participants can be considered music performance students regardless of their main specialization. They were taking different courses in different faculties and were in different year groups. There were very few missing data (less than 1% overall); these were imputed when necessary, using the expectation-maximization (EM) algorithm in SPSS 25.0, which provides unbiased estimates of parameters (Enders, 2001; Scheffer, 2002). The dataset can be made available from the corresponding author on request. This sample has already been used in a separate study published by Czerwiński et al., 2023.

INSTRUMENTS

Demographics. Respondents were asked about their age and gender as well as to provide estimates of the total number of hours they devote every week to studying at the university, both in and outside classes (e.g., at home or the library).

Study addiction. Study addiction was measured using the Bergen Study Addiction Scale (BStAS; Atroszko et al., 2015a). It consists of seven items pertaining to experiences during the past 12 months (e.g., "Studied in order to reduce feelings of guilt, anxiety, helplessness and depression?"), with a Likert-type response scale ranging from 1 (*never*) to 5 (*always*). It showed good psychometric qualities in previous research (Lawendowski et al., 2020). The Cronbach's α reliability coefficient in the current sample was .76.

Learning engagement. Learning engagement was measured using a single item, the question "How engaged in learning are you?" (Atroszko, 2014), with responses ranging from 1 (*I am not at all engaged*) to 7 (*I am completely engaged*). It showed good validity and test-retest reliability in previous research with an intraclass correlation coefficient of .77 for test-retest reliability (Atroszko, 2014; Łukowicz et al., 2017).

Big Five personality traits. The Mini-IPIP (Donnellan et al., 2006) was used to measure the Big Five personality traits. It consists of a 20-item inventory with four items measuring each of the Big Five personality factors: extraversion (e.g., "Talk to a lot of different people at parties"), agreeableness (e.g., "Sympathise with others' feelings"), conscientiousness (e.g., "Like order"), neuroticism (e.g., "Get upset easily") and intellect (e.g., "Have a vivid imagination"). Participants indicate how well each statement described them using a 5-point Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The scale showed good psychometric qualities in previous research (Czerwiński & Atroszko, 2020). The Cronbach's α reliability coefficients in the current study were the following: .74 for extraversion, .61 for agreeableness, .77 for conscientiousness, .71 for neuroticism, and .64 for intellect.

MPA. The Kenny Music Performance Anxiety Inventory – Revised (K-MPAI-R; Kenny, 2009) was used to measure MPA. The instrument consists of 40 items (e.g., "My worry and nervousness about my performance interferes with my focus and concentration") with a 7-point Likert-type response format, ranging from 0 (*strongly disagree*) to 6 (*strongly agree*). Although originally intended as a multidimensional tool, recent research indicates the unidimensional approach to be superior (Chang-Arana et al., 2018). The scale showed good psychometric qualities in previous research (Kantor-Martynuska & Kenny, 2018). The Cronbach's α reliability coefficient in the current sample was .93.

Perceived stress. The short version of the Perceived Stress Scale (PSS-4; Cohen et al., 1983) was used as a measure of perceived stress, with four items referring to the previous month (e.g., "In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?"). Response options are 0 (*never*), 1 (*almost never*), 2 (*sometimes*), 3 (*fairly often*), and 4 (*very often*). The scale showed good validity and reliability in previous research (Atroszko, 2015; Czerwiński et al., 2020). The Cronbach's α reliability coefficient in the current sample was .78.

General quality of life, general health and quality of sleep. Three single-item measures of different aspects of quality of life, developed on the basis of the WHOQOL-BREF (Skevington et al., 2004), were used. General quality of life was measured by the question "How would you rate your quality of life?" with a 9-point Likert-type scale ranging from *very poor* (1)

to *very good* (9). General health was measured by the question "How satisfied are you with your health?" with a 9-point Likert scale ranging from 1 (*very dissatisfied*) to 9 (*very satisfie*d). Sleep quality was measured by the question "How satisfied are you with your sleep?" with a 9-point Likert scale ranging from 1 (*very dissatisfied*) to 9 (*very satisfie*d). This instrument showed good validity and test-retest reliability in previous research with intraclass correlation coefficients of .86 for general quality of life, .72 for general health, and .81 for sleep quality (Atroszko et al., 2015b; Czerwiński et al., 2020).

GPA. The students were asked to provide information about their grade point average (GPA) from the semester prior to the study as accurately as possible. All universities used a scale ranging from 1 to 25. However, some courses at the Stanisław Moniuszko Academy of Music in Gdansk used a scale ranging from 2 to 6. In these cases, the GPA was recalculated using the conversion rate provided in the official rulebook of the academy.

STATISTICAL ANALYSES

Mediation and moderation analyses. Because the study was confirmatory in nature and the analyses were driven by clear hypotheses, no adjustments for multiple testing were applied. It should also be noted that correcting for multiple comparisons is controversial, with some researchers arguing that it is incorrect (see Gelman et al., 2012; Perneger, 1998; Rothman, 1990).

References

Atroszko, P. (2014). Developing brief scales for educational research: Reliability of single-item self-report measures of learning engagement and exam stress. In M. McGreevy & R. Rita (Eds.), *Proceedings of the 1st Biannual CER Comparative European Research Conference* (pp. 172–175). Sciemcee.

Atroszko, P. A. (2015). *The structure of study addiction: Selected risk factors and the relationship with stress, stress coping and psychosocial functioning* (Unpublished doctoral thesis). University of Gdansk, Poland.

Atroszko, P. A., Andreassen, C. S., Griffiths, M. D., & Pallesen, S. (2015a). Study addiction – a new area of psychological study: Conceptualization, assessment, and preliminary empirical findings. *Journal of Behavioral Addictions, 4*, 75–84. https://doi.org/10.1556/2006.4.2015.007

Atroszko P. A., Bagińska P., Mokosińska M., Sawicki A., & Atroszko B. (2015b). Validity and reliability of single item self-report measures of general quality of life, general health and sleep quality. In M. McGreevy & R. Rita (Eds.), *Proceedings of the 4th biannual CER Comparative European Research conference* (pp. 207–211). Sciemcee.

Chang-Arana, Á. M., Kenny, D. T., & Burga-León, A. A. (2018). Validation of the Kenny Music Performance Anxiety Inventory (K-MPAI): a cross-cultural confirmation of its factorial structure. *Psychology of Music, 46*, 551–567. https://doi.org/10.1177/0305735617717618

Cohen S., Kamarck T., & Mermelstein R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*, 385–396. https://doi.org/10.2307/2136404

Czerwiński, S., & Atroszko, P. (2020). Scores of short and free scale for Big Five explain perceived stress at different stages of life: Validity, reliability and measurement invariance of the Polish adaptation of Mini-IPIP. *Current Issues in Personality Psychology, 8*, 73–82. https://doi.org/10.5114/cipp.2020.95149

Czerwiński, S. K., Lawendowski, R., Kierzkowski, M., & Atroszko, P. A. (2023). Can perseverance of effort become maladaptive? Study addiction moderates the relationship between this component of grit and well-being among Music Academy students. *Musicae Scientiae, 27*, 568–595. https://doi.org/10.1177/1029864922109513

Czerwiński, S., Mackiewicz, J., Mytlewska, W., & Atroszko, P. (2020). Factorial validity, measurement invariance and concurrent validity of Hospital Anxiety and Depression Scale in a sample of Polish undergraduate students. *Psychiatria i Psychologia Kliniczna, 20*, 13–18. https://doi.org/10.15557/pipk.2020.0002

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The Mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18*, 192–203. https://doi.org/10.1037/1040-3590.18.2.192

Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling, 8*, 128–141. https://doi.org/10.1207/S15328007SEM0801_7

Gelman A., Hill J., & Yajima M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness, 5*, 189–211. https://doi.org/10.1080/19345747.2011.618213

Kantor-Martynuska J., & Kenny D. T. (2018). Psychometric properties of the Kenny-Music Performance Anxiety Inventory modified for general performance anxiety. *Polish Psychological Bulletin, 49*, 332–343. https://doi.org/10.24425/119500

Kenny, D. T. (2009). Negative emotions in music making: Performance anxiety. In P. Juslin & J. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 1–39). Oxford University Press.

Lawendowski, R., Bereznowski, P., Wróbel, W. K., Kierzkowski, M., & Atroszko, P. A. (2020). Study addic-

tion among musicians: Measurement, and relationship with personality, social anxiety, performance, and psychosocial functioning. *Musicae Scientiae, 24*, 449–474. https://doi.org/10.1177/1029864918822138

Łukowicz, P., Choynowska, A., Świątkowska, A. M., Bereznowski, P., Wróbel, W. K., & Atroszko, P. A. (2017). Validity of single-item self-report measure of learning engagement. In J. Nyćkowiak & J. Leśny (Eds.), *Badania i rozwój młodych naukowców w Polsce – nauki humanistyczne i społeczne. Część IV* [Research and development of young scientists in Poland – humanities and social sciences. Part IV] (pp. 41–49). Młodzi Naukowcy. Retrieved from https://experior.ug.edu.pl/wp-content/uploads/Validity-of-single-item-self-report-measure-of-examination-stress.pdf

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *BMJ, 316*, 1236–1238. https://doi.org/10.1136/bmj.316.7139.1236

Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology, 1*, 43–46. https://doi.org/10.1097/00001648-199001000-00010

Scheffer, J. (2002). Dealing with missing data. *Research Letters in the Information and Mathematical Sciences, 3*, 153–160.

Skevington, S. M., Lotfy, M., O'Connell, K. A., & WHOQOL Group (2004). The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A report from the WHOQOL group. *Quality of Life Research, 13*, 299–310. https://doi.org/10.1023/B:QURE.0000018486.91360.00