DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF ZADAR

24th PSYCHOLOGY DAYS IN ZADAR

Book of Selected Proceedings

Publisher
University of Zadar
Department of Psychology

For publisher
Josip Faričić, Rector

Publishing Committee
Lena Mirošević, Chair

Edited by
Ivana Macuka

Reviewers
Irena Pavela Banai, J. J. Strossmayer University of Osijek, Croatia
Tamara Mohorić, Faculty of Humanities and Social Sciences, University of Rijeka, Croatia

Editorial Board
Irena Burić, University of Zadar, Croatia
Ana Slišković, University of Zadar, Croatia
Ana Šimunić, University of Zadar, Croatia
Pavle Valerjev, University of Zadar, Croatia
Theresa J. Brown, Georgian Court University, USA
Lisa E. Kim, University of York, United Kingdom
Ljiljana B. Lazarević, University of Belgrade, Serbia
Hui Wang, Education University, Hong Kong

Technical editor
Ivana Macuka

English language editing
Leonarda Lovrović

Graphic design
University of Zadar

ISBN 978-953-331-605-5 (online)

DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF ZADAR

INTERNATIONAL SCIENTIFIC CONFERENCE

24th PSYCHOLOGY DAYS IN ZADAR

Book of Selected Proceedings

morepress
morepress.unizd.hr

Zadar, Croatia, 2025.

Editor's Foreword

The *Psychology Days in Zadar* is a distinguished international conference organised biennially by the Department of Psychology at the University of Zadar. Since its beginnings, the conference has grown from a regional academic gathering into a highly regarded international event, attracting researchers and practitioners from across Europe and beyond. Its continued success reflects both a strong commitment to scientific excellence and the enthusiasm of its participants, whose engagement shapes and sustains its reputation.

The *24th Psychology Days in Zadar*, held from 26 to 28 May, once again brought together a dynamic and diverse academic community eager to exchange knowledge, present cutting-edge research, and engage in meaningful discussions on contemporary psychological issues. Organised by an international scientific and organisational committee, this year's conference upheld its tradition of fostering intellectual collaboration and innovation.

The *Book of Selected Proceedings* presents a carefully curated selection of empirical studies showcased at the conference, alongside invited lectures offering theoretical and review perspectives. Researchers from various countries contributed to this volume, presenting 12 papers across keynote lectures, symposia, and oral presentations. Each paper underwent a rigorous double-blind review process, ensuring the publication maintains the highest academic standards. The selected research spans a wide range of psychological disciplines, including personality psychology, educational psychology, developmental psychology, and clinical psychology - offering valuable insights for researchers and practitioners alike.

This publication is the product of a collaborative effort, and we are deeply grateful to all who have contributed to its realisation. We thank the authors for entrusting us with the dissemination of their work, and the reviewers for their invaluable insights and critical feedback. Our sincere appreciation also goes to the members of the Editorial Board for their commitment throughout the editorial process. Lastly, we acknowledge the continuous support of the University of Zadar, without which this volume would not have been possible.

On behalf of the Editorial Board and contributing authors, it is my great pleasure to present the Book of Selected Proceedings of the 24th Psychology Days in Zadar. We hope it will serve as a valuable resource and source of inspiration for researchers, educators, and practitioners in the field of psychology.

Ivana Macuka
Editor-in-Chief

Contents

1

Understanding Parental Burnout: Significance and Implications

Isabelle Roskam
Department of Psychology, UCLouvain, Belgium
isabelle.roskam@uclouvain.be

Moïra Mikolajczak
Department of Psychology, UCLouvain, Belgium

Abstract

Despite the joy and pride it can bring, parenting comes with its share of challenges and negative emotions. It can even sometimes lead to suffering in the parental role. Parental burnout (PB) encompasses a range of negative feelings including exhaustion, emotional distancing, and a loss of pleasure in parental roles. This article explores the complexities of parental burnout, distinguishing it from ordinary parenting stress, and examines its prevalence, antecedents, and consequences. Despite having been recognised since the 1980s, empirical research on PB has surged in the last fifteen years. A validated tool, the Parental Burnout Assessment (PBA), facilitates measurement and understanding of the disorder. It revealed that approximately 5% of parents globally are affected, with significant variations influenced by cultural contexts and individualistic values. Key risk factors include personal traits, parenting practices, and family dynamics, while specific mediators in the relation between culture and parental burnout, such as parental self-discrepancy, affect burnout levels. The article discusses the bidirectional effects of PB on children's behaviour and parental relationships, highlighting increased neglect and aggression. Prevention and treatment are considered essential, emphasising the importance of early identification and compassionate support. Interventions focused on rebalancing parental stressors and resources have shown promising results, reducing symptoms and associated risks of neglect and violence. Ultimately, the article underscores that while parental burnout is a serious concern, it is both a preventable and treatable condition, offering hope for affected families through targeted psychological support.

Keywords: *parental burnout, prevention, treatment, culture*

Introduction

In most cultures around the world, parenting is associated with positive emotions such as joy, pride, and love. The celebrations surrounding births reflect this widely held emotional context. However, being a parent also involves moments of doubt, despair, feelings of failure or helplessness, anger, and fear. These negative emotions are often less emphasised, less recognised, and less audible. Parents struggling with negative emotions in their role often feel disoriented, ashamed, and alone.

A wealth of literature has nevertheless focused on parental stress, considered an integral part of parenting (Deater-Deckard, 2014). Parental stress is a normal response to the challenges of parenting, which typically comes and goes. On the other hand, parental burnout, which emerged in the 1980s, is a more severe and chronic condition (Mikolajczak et al., 2019). It is marked by extreme exhaustion, emotional distancing from one's children, and a loss of satisfaction in the parenting role. The concepts of parental stress and parental burnout are related but distinct concepts with an association of around .50 (Mikolajczak et al., 2023). It is only in the last fifteen years that parental burnout has seen unprecedented growth in empirical research. The term burnout is well known in professional life. It was in the 1960s that the term "burnout" appeared in the professional field with Bradley (1969). Freudenberger (1974) observed that some of the volunteers who worked with him were suddenly exhausted, drained—literally "burnt out"—by contact with patients. He said of these volunteers that they resemble a candle that, after lighting for many hours, has burned out from within and offers only an outdated flame. For this author, burnout was a trouble of the fighter, of the hyper invested. It was Christina Maslach who, in the early 1970s, devoted her scientific work to professional burnout. She observed that some caregivers came to detach themselves from their patients, to consider them as "cases" and not as people. She noticed that these caregivers were exhausted, that they were no longer fulfilled, no longer did they work well, and sometimes even became cynical (Maslach, 1976). For this author, burnout was the trouble of those who take their job of caring and helping others too much to heart.

Maslach proposed a tridimensional conception of job burnout, which was considered as a syndrome resulting from lasting exposure to excessive work stress. People in burnout feel emotionally and physically exhausted by their work, they also tend to depersonalise the beneficiaries of their job (as the nurses who refer to their patients as "the stomach of room 21" rather than as Mrs Johnson, for example). And they feel as they are not as efficient as they used to be. Research has shown that this is not only a perception since burned out employees do make more errors (Garcia et al., 2019).

The parallel between job and parental burnout is particularly interesting because parents can be considered as persons who fight for their children, who are hyper involved in their role as parents. And parenting is a job of caring and helping someone else. The idea of parental burnout originally came from two men, Procaccini and Kiefaver (1983), who worked and consulted on job burnout. They did not do any research on parental burnout but wrote a book in which they drew a parallel between job burnout and parental burnout. Their idea was later tested by Pelsma (1989), who found evidence for the first and third dimensions of burnout but not the second. He then gave up the idea. Fortunately, the idea was later taken up by researchers in health psychology who showed that parents of ill or disabled children were more likely to suffer from burnout than control parents. Scientifically, the idea of parental burnout was born. But do parents really need to have an ill or disabled child to be at risk of parental burnout?

The first symptom of parental burnout is physical and emotional exhaustion related to the parental role, leaving the parent with zero energy to care for their children. The second symptom differs somewhat from the depersonalisation observed in job burnout. As Pelsma (1989) discovered, depersonalisation is not characteristic of parental burnout. Instead of perceiving their children as objects, burnt-out parents exhibit emotional distancing from their children. Parents continue to care for their basic needs. They feed them, drive them to school, but do little beyond that. They behave in an autopilot mode with very little emotional involve-

ment in parenting. The third factor is a loss of pleasure and feelings of being fed up with parenting. Parents do not feel happy when they spend time with the child whether during meals, play, or caregiving. The third factor was also somewhat different from the third symptom of job burnout since parents do not lose their sense of efficacy, they still know how to parent and care for their children. But what they lose is the pleasure to parent them. The fourth and last factor refers to the contrast perceived between the current and previous parental self: the parent no longer recognises themselves. They are not the parent they wanted to be, or they were before the episode of burnout. This contrast is often noticed by both the partner and the children who do not recognise their mother or father. Importantly, parents' testimonies showed that burnout was a context-related syndrome. These symptoms are experienced in the parental role. And for burnt-out parents, their job is like a "safe haven" where they feel good. The reverse is true for people experiencing job burnout. For them, the family is the safe heaven and in the family context, they do not feel exhausted, detached, and inefficacious.

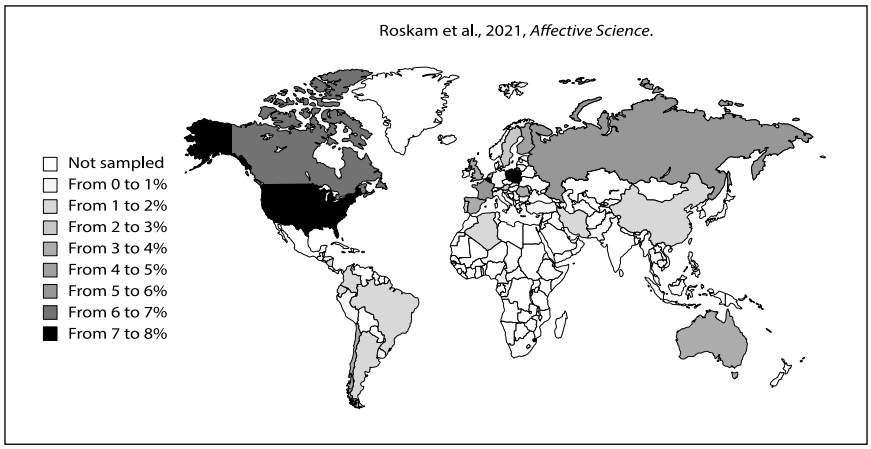
Based on these four dimensions, the Parental Burnout Assessment (PBA, Roskam et al., 2018) was developed to evaluate the presence and severity of this disorder. The PBA is currently the gold standard measure of parental burnout. It has been validated in the 15 most widely spoken languages, and externally validated clinical cut-offs are now available to discriminate between parents who are not in burnout, those who are at-risk, and parents who suffer from extreme fatigue (Roskam et al., 2022). Having a validated measurement tool has facilitated the rapid expansion of the research field in many regions around the world. Studies have particularly focused on the prevalence of parental burnout, its specificities regarding close concepts as parental stress, job burnout, depression or parenthood regret, its antecedents, its consequences, and ways to prevent and treat it. Hereafter, we will summarise the main findings in each of these issues.

Prevalence of Parental Burnout: A Global Perspective

What is the prevalence of parental burnout? How common is this phenomenon among parents, and does its prevalence vary across cultural contexts? To address these questions, we conducted a comparative analysis of the point prevalence of parental burnout across 42 countries (Roskam et al., 2021). Point prevalence refers to the proportion of a population experiencing a specific disorder within a defined timeframe. It is typically lower than both 12-month prevalence and lifetime prevalence, which indicates the likelihood of experiencing a burnout episode at some point during one's parenting role. As shown in Figure 1, our findings reveal that the worldwide point prevalence of parental burnout stands at approximately 5%, but that large variations exist across countries.

Figure 1

Point prevalence of parental burnout worldwide



To put this into context, around 200,000 parents in Belgium, 900,000 in France, and over 5 million in the United States are currently experiencing severe symptoms of burnout in their parental roles.

In most of the countries studied, we also observed a gender disparity, with a ratio of two burnt-out mothers for every father. This discrepancy likely arises from the societal perception of mothers as primary caregivers, rendering them more vulnerable to parenting stress than fathers. The prevalence of parental burnout also varies significantly across countries and cultures, with higher rates generally found in Western nations compared to non-Western countries. Culture, as a broad concept encompassing sociodemographic, economic, and sociological factors, plays a crucial role in understanding these differences. To identify which factors account for the variations in parental burnout across countries, we employed a multilevel model, treating parental burnout as the dependent variable. Our analysis indicated that individualistic values accounted for the greatest variance in parental burnout globally. Parents raising their children in cultures characterised by high individualism are at greater risk of experiencing burnout (Roskam et al., 2021).

What factors contribute to the heightened risk of parental burnout in individualistic cultures? In other words, how does raising children in a society shaped by individualistic values affect parental behaviours or attitudes, which in turn increase the risk of burnout in the parenting role? We addressed this question in a recent study (Roskam et al., 2023). The primary mediator identified was parental self-discrepancy. Parents from individualistic societies are more likely to perceive a gap between their socially prescribed parental identity and their actual parenting experience. This perception of inadequacy increases the likelihood of burnout. Western societal standards for parenting may be so demanding that parents feel perpetually insufficient in their roles, which leads to exhaustion and dissatisfaction.

The second significant mediator is parental task-sharing. In individualistic cultures, the heavy responsibilities of parenting are often viewed as the sole burden of the individual parent, resulting in a reluctance to seek help. This self-reliance can hinder the sharing of parenting duties with other caregivers, supporting findings from other fields. Research in areas such as healthcare and corporate settings indicates that individualistic cultural norms, which favour self-sufficiency, are associated with increased burnout and related outcomes (Roskam et al., 2023).

The third mediator is the pursuit of agency and self-directed socialisation goals. Parents in individualistic contexts emphasise values of independence and self-reliance, instilling these ideals in their children from an early age. This focus on autonomy can diminish parental authority, requiring parents to navigate negotiations rather than enforce compliance. Consequently, parenting becomes more demanding and stressful, with no guarantee of a child's obedience (Roskam et al., 2023).

These three mediating mechanisms accounted for 21% of the relationship between country-level individualism and parental burnout. Although this figure is significant, it also suggests that 79% of the impact of individualism on parental burnout is mediated by other factors not addressed in the study (Roskam et al., 2023).

The work on cultural influences on parental burnout is still ongoing. However, we cannot yet dismiss the possibility that some observed differences might arise from methodological artifacts. Specifically, the PBA—the gold standard for measuring parental burnout—was developed by Western researchers based on the experiences of Western parents (Roskam et al., 2018). It remains feasible that the PBA may not adequately capture culturally specific symptoms of burnout, potentially leading to underreported prevalence in non-Western nations. Future research is needed to explore the existence of culturally specific signs of exhaustion, which may contribute to enriching our understanding of universal and culture-specific characteristics of parental burnout.

The Specificity of Parental Burnout

Is parental burnout really distinct from ordinary parenting stress? Is parental burnout merely a manifestation of lazy parenting? To explore whether the stress experienced by parents undergoing burnout is genuinely elevated, we investigated biological indicators, specifically hair cortisol levels, as one of the most reliable markers of chronic stress (Greff et al., 2019). Hair samples allow us to measure cortisol accumulation over varying timeframes: one month (1 cm), two months (2 cm), or three months (3 cm) from a strand collected at the back of the scalp. Beyond that length, the hair becomes too damaged for accurate assessment.

Our study compared hair cortisol levels between parents in burnout and demographically matched control parents—those of similar age, gender, number of children, income, education, and employment status who felt fulfilled in their parenting role. The findings revealed that parents experiencing burnout exhibit significantly increased levels of cortisol, approximately twice those of normal parents. Notably, these levels surpass those found in individuals suffering from chronic pain or victims of domestic violence. This underscores that parental burnout is not merely a psychological construct; it is also an objective reality measurable through biological markers (Brianda, Roskam, & Mikolajczak, 2020).

In distinguishing parental burnout from ordinary stress, we also considered its relationship with similar syndromes, particularly job burnout and depression. As both parental and job burnout are stress-related disorders, questions arise about their correlation. If burnout in one domain inevitably leads to burnout in another, we would expect a high correlation between the two among working parents. However, existing studies indicate a moderate correlation, suggesting that burnout is context-specific: one can experience burnout in one realm without succumbing to it in another. Our meta-analysis of seven studies revealed an average correlation of 0.36 between parental and job burnout. Similarly, the relationship between parental burnout and depression showed a moderate to large correlation, averaging 0.55 (Mikolajczak et al., 2023; Mikolajczak et al., 2020).

To further delineate the conceptual boundaries of parental burnout, we examined its proximity to the concept of parenthood regret (Roskam et al., in revision). While both constructs emerged and received scholarly attention around the same period, they have historically been studied within separate research streams. Parenthood regret is defined as a self-conscious emotion stemming from past decisions—specifically, the decision to have a child—which leads to current dissatisfaction in parental fulfilment. Regret is a universal emotion present from childhood but holds significant weight in the parental context due to its association with the child.

Using Latent Profile Analysis on two independent samples, we identified four distinct profiles: a group of parents with low levels of both parental burnout and parenthood regret; a second group exhibiting high levels of parental burnout but lower levels of regret; a third group with high levels of regret but lower burnout; and a final group with elevated levels of both burnout and regret. Notably, the largest cohort consisted of parents with low levels of both constructs, accounting for approximately 70% (69.1% in Study 1 and 78.9% in Study 2). Conversely, approximately 15% of participants demonstrated high parenthood regret with lower burnout; around 7% displayed high parental burnout with lower regret; and only about 4% exhibited elevated levels of both constructs.

Our analysis suggests that a model in which parental burnout and parenthood regret manifest as separate latent factors fits the data better than a model unifying them under a single construct of parental distress. Consequently, we can conclude that parental burnout and parenthood regret are distinct phenomena: one may experience parental burnout without expressing regret, and vice versa. In some instances, both experiences may coexist (Roskam et al., in revision).

The Antecedents of Parental Burnout

Do certain parent populations face a higher risk of burnout than others? For instance, are parents of young children, adolescents, single parents, or parents of children with special needs at greater risk? Our studies have contested several prevailing myths. We found no increased risk for parents of young children or adolescents, nor for those with gifted children. The risk was slightly elevated for single parents and those in precarious situations, but it was particularly high among parents with children with special needs (Roskam & Mikolajczak, 2023).

If there is no profile of parents at risk of burnout, why do some parents experience burnout while others do not? This question invites an exploration of the aetiology of parental burnout. Much research has been conducted in the organisational sector, where the prevailing model suggests that burnout occurs when individuals face chronic high demands in a context lacking sufficient resources to meet those demands. Does this model hold true within the realm of parenting?

If this hypothesis is correct, we would anticipate that parental burnout arises when the factors contributing to parental stress exceed those that mitigate it. We examined five sets of risk factors. The first set encompasses sociodemographic factors. These include the number and ages of children, family structure (e.g., single-parent, same-sex, or blended families), net household income, and living space size. The second one comprises the situational factors. This category encompasses challenges such as having a child with chronic illnesses, learning difficulties, or special needs. The third one is about personal factors. Traits such as emotional competence, emotion regulation, neuroticism, and parental perfectionism fall under this category. Fourth, we studied the set of parenting factors including poor child-rearing practices, for example. Fifth, we considered a set of family functioning factors. A lack of support from a co-parent and overall family disorganisation are some examples. We finally considered corresponding sets of factors that contribute to parental resources (i.e., factors that alleviate parental stress) to provide a balanced perspective (Mikolajczak, Raes, et al., 2018).

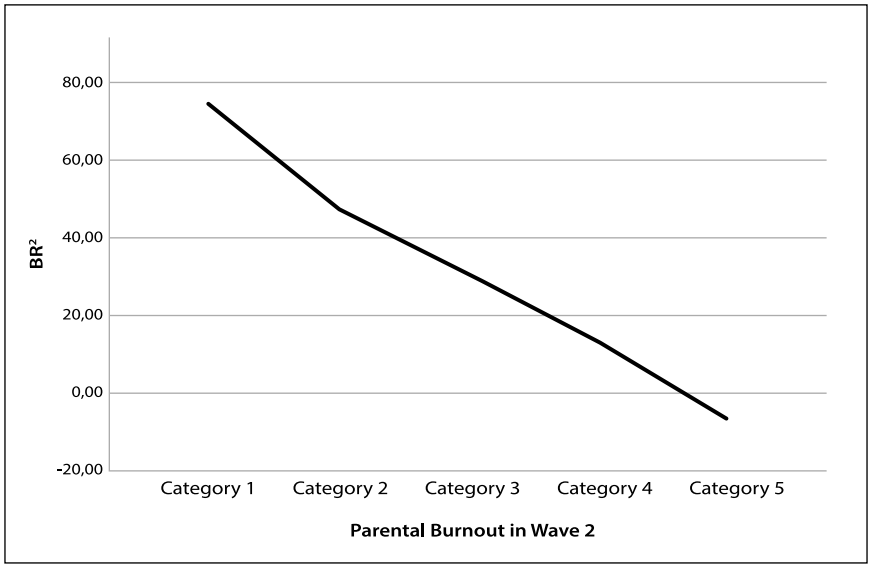
Our findings revealed that sociodemographic and situational factors account for less variance in parental burnout compared to personal factors, parenting practices, and family dynamics. This led us to develop an instrument designed to measure the balance (or imbalance) between significant risks and resources, referred to as the Balance of Risks and Resources (BR^2; Mikolajczak & Roskam, 2018).

To validate our model, we conducted a longitudinal study investigating both cross-sectional and prospective relationships between the balance score and instances of parental burnout. As illustrated in Figure 2, parents whose balance leans toward the negative—indicating that the cumulative weight of risk factors surpasses the available resources—are particularly susceptible to burnout. Specifically, when the balance score falls below zero, parents are categorised as experiencing daily symptoms of parental burnout. This correlation was consistently observed in both cross-sectional and prospective analyses. Thus, it can be asserted that parental burnout arises from a chronic imbalance between parenting stressors and available resources. It is important to note, however, that while this common aetiological mechanism exists, each case of parental burnout is unique. The specific stressors and resources involved in shaping the balance vary from parent to parent (Mikolajczak et al., 2023; Roskam & Mikolajczak, 2018).

The BR2 model has two important implications. The first is that the factors serving as antecedents of parental burnout have their effects moderated by other factors present in the balance. These moderating factors can either exacerbate the risk of burnout or, conversely, act as a buffer against the antecedent factors. A few studies have been dedicated to examining these interaction effects between stress factors and resources within the balance (e.g., Bayot et al., 2021; Lin et al., 2021; Lin et al., 2022). The second implication is that parental burnout is not a static condition. The number and nature of the factors present in the balance, as well as the imbalance between stress factors and resources, fluctuate over time, leading to changes

Figure 2

The prospective association between the Balance of Risks and Resources (BR2) and levels of parental burnout 6 months later (in Wave 2)



Note. Category 1 = 2/3 of the symptoms never to a few times a year; Category 2 = once a month or less; Category 3 = a few times a month; Category 4 = a few times a week; Category 5 = every day.

in the level of parental burnout at the intrapersonal level. Therefore, parental burnout should be viewed as a dynamic concept rather than a static experience. This has been demonstrated in longitudinal studies (e.g., Piotrowski, 2023; Roskam & Mikolajczak, 2021).

The Consequences of Parental Burnout

What happens when a parent experiences burnout? The consequences of parental burnout are multifaceted, impacting the parent, their children, and their relationship with their spouse.

Research indicates that parents suffering from burnout often experience increased irritability, guilt, and shame, which can lead to feelings of self-hatred. Additionally, they may develop physical symptoms. The connection between physical symptoms and parental burnout is not surprising because elevated cortisol levels in the body are a significant contributor to these outcomes. Chronic stress, as indicated by high cortisol levels, can exacerbate pre-existing health issues. Burnt-out parents sometimes engage in increased alcohol consumption.

Moreover, the overwhelming desire to escape their unbearable parenting situation may manifest as thoughts of self-harm or suicidal ideation (Mikolajczak, Brianda, et al., 2018; Mikolajczak et al., 2019). It is noteworthy that suicidal ideation is considerably more prevalent among parents experiencing burnout than in cases of job burnout or depression. While individuals facing professional burnout may access medical leave or change careers to alleviate stress, parents lack similar avenues for distancing themselves from the source of their stress—their parental responsibilities. Since parenting is a role from which one cannot resign, some parents may develop extreme thoughts, including suicidal ideation.

The effects of parental burnout extend to children as well, leading to increased neglectful and aggressive behaviours. While depression raises the risk of neglect by a factor of six and violence by a factor of seven, parental burnout escalates the risk of neglect by a staggering factor of 13 and the likelihood of violence by a factor of 20 (Mikolajczak et al., 2019).

One poignant testimony comes from Elisabeth, a mother of two, who shared: "I experienced very intense outbursts directed at my daughter. Her screams and cries were unbearable, and I found myself engaging in severe verbal aggression. Although I strived to avoid physical violence, there were moments when I felt an overwhelming desire to harm her. One incident involved putting her down for a nap; when she resisted, I lost control and, at that moment, realised I could jeopardise my children's safety." (Hubert & Aujoulat, 2018).

Our recent research also identified a bidirectional effect of parental burnout on children's behaviour, noting a transactional relationship where difficult behaviours increase parental burnout and vice versa. This cyclical dynamic creates a vicious cycle, exacerbating the challenges faced by both the parent and child (Woine et al., under review; Woine et al., 2024).

Lastly, parental burnout adversely affects romantic relationships, leading to increased distance and conflict between partners (Mikolajczak, Brianda, et al., 2018). Elisabeth further recounted her experiences, stating, "I felt exasperated by my children constantly demanding my attention. When my husband arrived home, I would often explode with frustration, venting my feelings towards him rather than addressing my children. Unfortunately, this resulted in verbal insults and created a toxic atmosphere."

Prevention and Treatment of Parental Burnout

Addressing parental burnout is crucial, given both the disorder's prevalence and its serious consequences. Prevention is the first line of defence against parental burnout. Our research indicates that the symptoms of burnout do not manifest simultaneously; rather, they appear in a specific sequence (Roskam & Mikolajczak, 2021). The initial signs typically involve exhaustion, followed by emotional distancing and a loss of pleasure in parenting. Importantly, neglectful and aggressive behaviours towards children emerge when parents begin to emotionally detach (Hansotte et al., 2021). This detachment often results in diminished empathy, further exacerbating the cycle of disconnection.

Currently, many parents seek help only after reaching a severe level of burnout, which complicates prevention efforts. There is a societal stigma associated with acknowledging feelings of exhaustion or dissatisfaction in parenting. Being associated with positive emotions such as joy and pride, it is difficult for parents to admit when they feel overwhelmed in their role as mother or father. If we could enhance early identification and support for parents in the initial stages of exhaustion, we could mitigate the risks associated with burnout, including suicidal ideation, child neglect, and behavioural issues in children. This requires breaking the taboo surrounding parental burnout, and conferences like this one play a vital role in changing the conversation.

Both prevention and treatment of parental burnout require a compassionate approach. Burnt-out parents need a safe space where their feelings can be acknowledged without judgment. Admitting distress in parenting—feeling overwhelmed or losing enjoyment in time spent with children—can evoke shame and guilt. Thus, it is essential to approach these negative emotions with utmost respect and understanding.

Once parents feel heard, we can address the roots of parental burnout. Our aetiological model serves as a framework through which we can effectively reduce burnout symptoms (Mikolajczak & Roskam, 2018). Longitudinal research indicates that parental burnout does not resolve spontaneously; therefore, we cannot rely on time alone for recovery. However, results from our randomised controlled trials demonstrated that targeted psychological interventions can significantly help parents by "rebalancing" their stressors and resources (Bayot et al., revised; Brianda, Roskam, Gross, et al., 2020).

The intervention begins with identifying the specific stressors and resources unique to each parent's situation. The nature, quantity, and impact of these factors vary among individuals. The goal of the

treatment is twofold: to reduce the number and weight of accumulated stressors—such as high parental standards—and to enhance the resources available to cope with current challenges.

Our findings revealed that an 8-week group intervention is effective in facilitating this rebalancing process. Participants reported an average symptom reduction of 37%, sustained even three months post-treatment. Additionally, there was a 30% decrease in instances of child neglect among parents who underwent the intervention, illustrating the direct link between parental burnout and neglectful behaviour (Brianda, Roskam, Gross, et al., 2020).

Furthermore, the intervention significantly decreased parental violence against children. Reductions in burnout symptoms corresponded with lower incidences of violence, reinforcing the notion that these behaviours stem directly from parental burnout.

Notably, we also observed a 52% decrease in hair cortisol levels among parents, suggesting a physiological recovery as their stress levels normalised to those of control parents. This objective measure corroborates the subjective improvements reported by both parents and therapists, highlighting that treatment yields tangible results.

Conclusion

In conclusion, while parental burnout is a severe and prevalent disorder with concerning consequences for families, it is also a condition that can be effectively treated. With appropriate support, parents can recover, experience a reduction in symptoms, and see a corresponding decline in physiological stress. Importantly, the risk of child misbehaviour, neglect, and abuse can be significantly diminished with effective intervention.

References

Bayot, M., Brianda, M. E., Van Der Straten Waillet, N., Mikolajczak, M., Shankland, R., & Roskam, I. (revised). Mindfulness and compassion-based approach to parental burnout: A randomized controlled trial with validated group treatment. *Journal of Child and Family Studies*.

Bayot, M., Roskam, I., Gallée, L., & Mikolajczak, M. (2021). When Emotional Intelligence Backfires. *Journal of Individual Differences*, *42*(1), 1-8. https://doi.org/10.1027/1614-0001/a000324

Bradley, H. B. (1969). Community-based treatment for young adult offenders. *Crime & Delinquency*, *15*(3), 359-370.

Brianda, M. E., Roskam, I., Gross, J. J., Franssen, A., Kapala , F., Gérard , F., & Mikolajczak, M. (2020). Treating Parental Burnout: Impact of Two Treatment Modalities on Burnout Symptoms, Emotions, Hair Cortisol, and Parental Neglect and Violence. *Psychotherapy and Psychosomatics*, *89*(5), 330-332. https://doi.org/10.1159/000506354

Brianda, M. E., Roskam, I., & Mikolajczak, M. (2020). Hair cortisol concentration as a biomarker of parental burnout. *Psychoneuroendocrinology*, *117*, 104681. https://doi.org/https://doi.org/10.1016/j.psyneuen.2020.104681

Deater-Deckard, K. (2014). *Parenting Stress*. University Press. https://books.google.be/books?id=GWfdngEACAAJ

Freudenberger, H. J. (1974). Staff Burn-Out. *Journal of Social Issues*, *30*(1), 159-165. https://doi.org/10.1111/j.1540-4560.1974.tb00706.x

Garcia, C. L., Abreu, L. C., Ramos, J. L. S., Castro, C. F. D., Smiderle, F. R. N., Santos, J. A. D., & Bezerra, I. M. P. (2019). Influence of Burnout on Patient Safety: Systematic Review and Meta-Analysis. *Medicina (Kaunas, Lithuania)*, *55*(9). https://doi.org/10.3390/medicina55090553

Greff, M. J. E., Levine, J. M., Abuzgaia, A. M., Elzagallaai, A. A., Rieder, M. J., & van Uum, S. H. M. (2019). Hair cortisol analysis: An update on methodological considerations and clinical applications. *Clinical Biochemistry*, *63*, 1-9. https://doi.org/10.1016/j.clinbiochem.2018.09.010

Hansotte, L., Nguyen, N., Roskam, I., Stinglhamber, F., & Mikolajczak, M. (2021). Are all Burned Out Parents Neglectful and Violent? A Latent Profile Analysis. *Journal of Child and Family Studies*, *30*(1), 158-168. https://doi.org/10.1007/s10826-020-01850-x

Hubert, S., & Aujoulat, I. (2018). Parental Burnout: When Exhausted Mothers Open Up. *Frontiers in Psychology*, *9*, 1021. https://doi.org/https://doi.org/10.3389/fpsyg.2018.01021

Lin, G.-X., Szczygieł, D., Hansotte, L., Roskam, I., & Mikolajczak, M. (2021). Aiming to be perfect parents increases the risk of parental burnout, but emotional competence mitigates it. *Current Psychology*, 1-9.

Lin, G.-X., Szczygieł, D., & Piotrowski, K. (2022). Child-oriented perfectionism and parental burnout: The moderating role of parents' emotional intelligence. *Personality and Individual Differences*, *198*, 1-6. https://doi.org/10.1016/j.paid.2022.111805

Maslach, C. (1976). Burnout. *Human Behavior*, *5*(9), 16-22.

Mikolajczak, M., Aunola, K., Sorkkila, M., & Roskam, I. (2023). 15 Years of Parental Burnout Research: Systematic Review and Agenda. *Current Directions in Psychological Science*, *32*(4), 09637214221142777. https://doi.org/10.1177/09637214221142777

Mikolajczak, M., Brianda, M. E., Avalosse, H., & Roskam, I. (2018). Consequences of parental burnout: Its specific effect on child neglect and violence. *Child Abuse and Neglect*, *80*, 134-145. https://doi.org/10.1016/j.chiabu.2018.03.025

Mikolajczak, M., Gross, J. J., & Roskam, I. (2019). Parental Burnout: What Is It, and Why Does It Matter? *Clinical Psychological Science*, *7*(6), 1319-1329. https://doi.org/10.1177/2167702619858430

Mikolajczak, M., Gross, J. J., Stinghlamber, F., Lindahl Norberg, A., & Roskam, I. (2020). Is parental burnout distinct from job burnout and depressive symptomatology? . *Clinical Psychological Science*, *8*(4), 673-689. https://doi.org/10.1177/2167702620917447

Mikolajczak, M., Raes, M.-E., Avalosse, H., & Roskam, I. (2018). Exhausted parents: Sociodemographic, child-related, parent-related, parenting and family-functioning correlates of parental burnout. *Journal of Child and Family Studies*, *27*(2), 602-614. https://doi.org/https://doi.org/10.1007/s10826-017-0892-4

Mikolajczak, M., & Roskam, I. (2018). A Theoretical and Clinical Framework for Parental Burnout: The Balance Between Risks and Resources (BR2). *Frontiers in Psychology*, *9*(886). https://doi.org/10.3389/fpsyg.2018.00886

Pelsma, D. M. (1989). Parent Burnout: Validation of the Maslach Burnout Inventory with a Sample of Mothers. *Measurement and Evaluation in Counseling and Development*, *22*(2), 81-87.

Piotrowski, K. (2023). Trajectories of parental burnout in the first year of the COVID-19 pandemic. *Family Relations*, *72*(3), 680-696. https://doi.org/https://doi.org/10.1111/fare.12819

Procaccini, J., & Kiefaver, M. W. (1983). *Parent burnout*. Double day.

Roskam, I., Aguiar, J., Akgun, E., Arena, A. F., Arikan, G., Aunola, K., Besson, E., Beyers, W., Boujut, E., Brianda, M. E., Brytek-Matera, A., Budak, A. M., Carbonneau, N., César, F., Chen, B. B., Dorard, G., Dos Santos Elias, L. C., Dunsmuir, S., Egorova, N.,…Mikolajczak, M. (2023). Three reasons why parental burnout is more prevalent in individualistic countries: a mediation study in 36 countries. *Social Psychiatry and Psychiatric Epidemiology*. https://doi.org/10.1007/s00127-023-02487-z

Roskam, I., Aguiar, J., Akgun, E., Arikan, G., Artavia, M., Avalosse, H., Aunola, K., Bader, M., Bahati, C., Barham, E. J., Besson, E., Beyers, W., Boujut, E., Brianda, M. E., Brytek-Matera, A., Carbonneau, N., César, F., Chen, B.-B., Dorard, G.,…Mikolajczak, M. (2021). Parental Burnout Around the Globe: a 42-Country Study. *Affective Science*, *2*(1), 58-79. https://doi.org/10.1007/s42761-020-00028-4

Roskam, I., Bayot, M., & Mikolajczak, M. (2022). Parental Burnout Assessment (PBA). In O. N. Medvedev, C. U. Krägeloh, R. J. Siegert, & N. N. Singh (Eds.), *Handbook of Assessment in Mindfulness Research* (pp. 1-22). Springer International Publishing. https://doi.org/10.1007/978-3-030-77644-2_81-1

Roskam, I., Brianda, M.-E., & Mikolajczak, M. (2018). A Step Forward in the Conceptualization and Measurement of Parental Burnout: The Parental Burnout Assessment (PBA) [Original Research]. *Frontiers in Psychology*, *9*(758). https://doi.org/10.3389/fpsyg.2018.00758

Roskam, I., & Mikolajczak, M. (2018). La Balance des Risques et des Ressources (BR^2): un modèle explicatif et clinique du burn-out parental. In *Le burn-out parental: comprendre, diagnostiquer et prendre en charge*. de Boeck.

Roskam, I., & Mikolajczak, M. (2021). The slippery slope of parental exhaustion: A process model of parental burnout. *Journal of Applied Developmental Psychology*, *77*, 101354. https://doi.org/https://doi.org/10.1016/j.appdev.2021.101354

Roskam, I., & Mikolajczak, M. (2023). Parental Burnout in the Context of Special Needs, Adoption, and Single Parenthood. *Children (Basel)*, *10*(7). https://doi.org/10.3390/children10071131

Roskam, I., Mikolajczak, M., & Piotrowski, K. (in revision). Unveiling the Uniqueness of Parental Burnout and Parenthood Regret: Impact on Parents and Children. *Clinical Psychological Science*.

Woine, A., Dardier, V., Mikolajczak, M., & Roskam I. (under review). Bidirectional and transactional effects between parental burnout and child internalizing and externalizing behaviors.

Woine, A., Escobar, M. J., Panesso, C., Szczygiel, D., Mikolajczak, M., & Roskam, I. (2024). Parental burnout and child behavior: A preliminary analysis of mediating and moderating effects of positive parenting. *Children*, *11*(3). https://doi.org/10.3390/children11030353

2

Functions of Emotional Crying[1]

Asmir Gračanin
Department of Psychology,
Faculty of Humanities and Social Sciences,
University of Rijeka
agracanin@ffri.uniri.hr

Abstract

Over the last three decades, research on emotional crying has increased significantly. Nevertheless, our understanding of the functions and consequences of this fascinating behaviour remains limited. This paper provides a brief overview of the latest theoretical contributions and empirical findings on the intra- and inter-individual effects of crying. Next, it attempts to disentangle the evolved functions of crying from its effects, which are not necessarily part of these functions. The central premise is that crying represents an evolved mechanism designed to transmit information in a limited set of ancestrally recurrent situations. The paper offers both theoretical analysis and empirical support for the idea that emotional crying functions as an emotional expression or signal that: (a) primarily operates through the visual channel, (b) evolved during recent human evolution to transmit information related to attachment processes, help-seeking, and the expression of friendly intentions, all of which can be structured around the main dimensions of social perception: warmth and competence, (c) often elicits positive reactions from observers, and (d) has positive consequences for the individual that are best understood within the context of its signalling function. In addition to the effects of crying on well-being of the crying individual that are mediated through positive responses from others, this contribution also discusses how crying might directly influence an individual's well-being. Understanding these direct effects requires a consideration of the evolved signalling functions of tears. Based on these proposed functions, the paper attempts a theoretical integration of the inter-individual functions and intra-individual effects of crying.

Keywords: *tears, emotional crying, evolution, communication, mood*

[1] This work is funded by the project The Dimensions of Warmth and Competence, Social Ranking, and Mental Health, uniri-iz-25-292, Financira Europska unija – NextGenerationEU.

Introduction

Emotional crying (hereafter "crying") can be defined as welling tears in or shedding tears from the eye (i.e., tearing up) that is not a consequence of any irritation of ocular structures. It typically occurs during intensive emotional reactions, mostly negative, but also positive ones. Crying is a multifaceted response: it is typically accompanied by muscular facial expressions, and sometimes includes vocal crying and sobbing, whereas the latter refers to convulsive inhaling and exhaling (Patel, 1993). Tears *always* preclude other forms of (sincere) crying, and this overlap allows for a certain degree of joint analysis of different aspects of the crying phenomenon. However, it is important to have in mind that some studies showed that aspects of crying other than tearing up (e.g., vocal) can result in different, often less positive reactions from observers (see Zickfeld & Wróbel, 2023). The current contribution will review relevant research and offer theoretical insights that are primarily intended to explain the occurrence of tearing up, as a key component of crying, which clearly has implications for other aspects of crying. Throughout the text, the term tears will be used only when referring specifically to this core aspect of crying (e.g., when it is clear that research dealt with tears only or when I refer to functional properties of tears), while the term crying will be used when specific research findings (likely) involved other aspects of crying, or when referring to functions of crying as a more general response.

Crying can be regarded as a form of emotional expression, but it is different from expressions of so-called primary emotions such as fear or disgust. While those expressions are relatively consistently linked to specific emotions and their functions (Ekman & Friesen, 2003), it is much harder to link crying with any specific emotion. There is an overall agreement among researchers that each emotion consists of cognitive appraisals, physiological component, behavioural tendencies or motivation, subjective feelings, and expressive behaviour. Each of these components differs between specific emotions to various extent. In some cases, some of these components have similar features across different emotions, which may also explain why different emotions co-occur in the same situations, and which makes borders between emotional categories relatively fuzzy (see Lange & Zickfeld, 2023). Crying represents the expressive component of different emotions and, as I argue in the current contribution, it is functionally linked to specific patterns of other basic components of different emotions (e.g., appraisals), which allows it to be coupled with various subjective emotional experiences.

Currently, there is no systematic empirical data providing evidence for any animal species other than humans producing emotional tears (but see Murata et al., 2022). Crying is a relatively frequent human behaviour. Across the globe, males cry, on average, once every two months, while females cry two to five times per month (Bylsma et al., 2011; Vingerhoets, 2013). Nevertheless, there are large individual differences in crying frequency and proneness other than gender. For example, neuroticism, extraversion, and habitual prosocial tendencies play an important role in whether someone will react with tears when exposed to various situations (Barthelmäs & Keller, 2021; Laan et al., 2012; Rottenberg et al., 2008b; Vingerhoets et al., 2018). The strongest triggers of crying are the events that occur rarely, such as social rejection, ending love relationship, and separation from and death of close others. More common situations that trigger crying are physical pain or injury, experiencing any kind of loss, interpersonal conflicts, being harshly criticised, feelings of personal inadequacy, as well as witnessing suffering of other beings (Denckla et al., 2014; Vingerhoets et al., 1997, 2009). Crying co-occurs with a variety of subjective emotional feelings, more often negative, including sadness, grief, self-pity, hopelessness, anger, frustration, but also positive ones, such as joy, elation, and relief, which is only a small part of the large number of different types of feelings reported across different studies (Crile, 1915; Frijda, 1986; Miceli & Castelfranchi, 2003; Murube et al., 1999; Vingerhoets et al., 1997). When it comes to specific emotions, people most often link crying with sadness (Choti et al., 1987; Vingerhoets et al., 1997), and less frequently with fear and anger (Barthelmäs et al., 2022; Gračа-

nin et al., 2021; Vingerhoets, 2013). A common factor linking all these situations seems to be the appraisal reflecting the state of helplessness (Gračanin et al., 2018a; Miceli & Castelfranchi, 2003; Murube et al., 1999; Vingerhoets et al., 1997; Zickfeld & Wróbel, 2024). As it will be presented below, one of the main proposed inter-individual functions of tears consists of promotion of helping behaviours by observers. Finally, tears seem to occur not just when requesting help, but also when offering help. Specifically, Murube and colleagues (1999) in their account on the functions of crying considered all episodes of positive crying to be related to offering help, which is a possibility that I will briefly discuss below.

Notably, helplessness is based on the appraisal of low coping potential or low control/agency (Moors et al., 2013), which is not typical of happiness and anger (Smith & Ellsworth, 1985) during which crying sometimes occurs (Gračanin et al., 2021; Vingerhoets et al., 1997). One of the explanations of this seeming paradox is that people cry when they feel overwhelmed with (intensive) emotion (Vingerhoets & Bylsma, 2016). Here, the intensity rather than the type of the emotion (and typical appraisal linked to it) appears to be important. Thus, even in case of anger or happiness, being overwhelmed with emotion results in the appraisal of very low, rather than high agency, eventually resulting in crying. The remaining question is whether there are other key precursors of crying in addition to high emotion intensity. It has been proposed that other factors, such as the characteristics of the receivers of the crying signal, interact with emotion intensity in predicting the likelihood of crying response (see Sznycer et al., 2025). These features are either stable or context dependent, for example: the observer is in a better position to help the individual than the individual themselves; the observer can easily fight back if attacked by the potentially crying individual—both of which certainly decrease the level of one's control, irrespective of the intensity of emotion. In addition, it is important whether the observer knows the individual. All such features fit within the major dimensions of social perception—warmth and competence (Abele et al., 2021; Fiske et al., 2007), which could represent a strong unifying framework for the understanding of the triggers and context of crying episodes. Zickfeld and Wróbel (2024) systematise observers' responses to emotional tears by structuring them around the two overarching dimensions of social perception—*communion* and *agency* (Abele et al., 2021). I further argue that the more specific aspects of these two dimensions—*warmth* and *competence*—and their interaction are especially relevant for understanding the *triggers* of emotional tears. Warmth refers to features such as friendliness and social closeness, but is not concerned with morality, which is included in the more general factor of communion. By assessing one's competence, we learn about one's ability to carry out one's intentions, which makes it more specific than agency, as the latter also involves assertiveness (Abele et al., 2021; Abele & Wojcizske, 2007; Fiske et al., 2007). The idea is that the individual perception of *relational warmth* (e.g., one perceives certain degree of social closeness or mutual stakes in well-being of the two individuals) interacts with *relational competence* (one perceives that the other individual has greater control over the situation) in prediction of the likelihood of emotional tearing (Gračanin et al., 2023). The current contribution relies on the idea that crying is typically triggered by the appraisals involving high warmth and low competence and their interaction.

Despite a recent increase in both basic and applied research on crying (see Bylsma et al., 2021; Gračanin et al. 2018a; Sznycer et al., 2025; Vingerhoets, 2013; Vingerhoets & Bylsma, 2016; Zickfeld & Wróbel, 2024), we still know very little about its functions and evolutionary origins. Why and how did crying evolve? How exactly does it affect crying individuals and what are its social consequences? The aim of the current contribution is to describe and disentangle the evolved functions of crying from the effects of emotional crying that are not necessarily fulfilling these functions. Specifically, the current contribution provides and critically evaluates the arguments for the idea that emotional crying represents an evolved mechanism whose function is to transmit information in a limited set of ancestrally reoccurring situations. Traditionally, the functions of crying are divided into *intra-individual*, referring to the effects of crying on the crying individual, and *inter-individual,* referring to its effects on observers. These functions are presented in Table 1, together with a brief overview of the currently available relevant empirical evidence. Overall, the

current contribution offers theoretical analysis and evaluates empirical support for the idea that emotional crying represents an emotional expression or signal that (a) most likely operates through the visual channel, (b) evolved during recent human evolution as a means of transmitting information, primarily related to attachment processes, seeking help, and showing friendly intentions, (c) most often results in positive reactions from observers, and (d) with its positive consequences for the crying individual best understood in the context of its signalling function. The idea that crying represents an evolved signalling mechanism is relevant for understanding all the effects that crying can have on the individual expressing this behaviour. Before I turn to the proposed evolved function of crying, I will briefly summarise the most important issues and findings from the research on the intra-individual effects of crying.

Table 1

Intra- and inter-individual functions of crying proposed in the literature and current state of the empirical evidence

Function of crying	Specific empirical evidence for the effects of crying	Strength of the evidence	References (examples)
Intra-individual functions of crying			
Psychological well-being improvements	Mood improvement:		
	Retrospective studies and lay conceptions	Strong	Bindra (1972); Cornelius (1986, 1997, 2001); Hanser et al. (2016); Hendriks et al. (2008); Rottenberg et al. (2008b)
	Studies in therapeutic context	Strong	Nelson (2008)
	Diary studies	Weak	Bylsma et al. (2011)
	Quasi experimental studies	Very weak	Gračanin et al. (2015, 2022); Martin & Labbot (1991); Gross et al. (1994); Rottenberg et al. (2002)
Physiological homeostasis regulation	*Quasi-experimental studies:*		
	Parasympathetic increases following crying	Moderate	Zickfeld & Grüning (2021)
	Changes in oxytocin and opioid levels	No evidence	Gračanin et al. (2022); Vingerhoets (2013)
Removing toxins from the body	*Quasi-experimental studies:*		
	Changes in cortisol levels	No evidence	Vingerhoets & Kirschbaum (1997)
Inter-individual functions of crying			
Elicitation of proximity, caregiving, and help	*Retrospective and diary studies:*		
	Tears promoting helping behaviour	Strong	Barthelmäs et al. (2024); Bylsma et al. (2008); Hendriks et al. (2008)
	Experiments with pictures/vignettes:		
	Tears promoting:		
	Perception of helplessness	Very strong	Balsters et al. (2013); Cornelius & Lubliner (2003); Vingerhoets et al. (2016); Zickfeld et al. (2021)
	Perception of emotions reflecting helplessness-related cognitive appraisal (e.g., sadness)	Very strong	Balsters et al. (2013); Gračanin et al. (2021); Ito & Ong (2023); Provine et al. (2009); Reed et al. (2015)
	Helping intentions	Very strong	Bobowik et al. (2023); Provine et al. (2009); Vingerhoets et al. (2016); Zickfeld et al. (2021)
	Feelings of empathy	Very strong	Gračanin et al. (2021); Zeifman & Brown (2011); Zickfeld et al. (2021)
	Approach intentions	Moderate	Hendriks & Vingerhoets (2006)
	Automatic approach/blocking avoidance	Weak	Gračanin et al. (2018b)
Aggression reduction	*Experiments with pictures/vignettes:*		
	Intentions for decreasing punishment	Weak	Picó et al. (2020)
Social bonding and collaboration	*Experiments with pictures/vignettes:*		
	Perception of friendliness, warmth, and non-aggressiveness	Strong	Vingerhoets et al. (2016); Zickfeld et al. (2021)
	Feelings of connectedness	Strong	Vingerhoets et al. (2016); Zickfeld et al. (2021)

Note. This table summarises the research findings on the effects of crying. A more complete analysis of the arguments supporting the main functions of crying also needs to be informed by the knowledge about its triggers.

Intra-Individual Effects of Crying

Earlier studies on crying focused exclusively on the direct effects of crying on the individual who exerts this behaviour, specifically, its effects on psychological well-being, and more specifically, changes in psychophysiological reactions and mood. Due to the great potential for conceptual refinements and critical approach to the existing research, this section will place a greater focus on the subjective aspects of potential intra-individual effects of crying. Nevertheless, I will also briefly introduce the major questions and findings on the effects of crying on psychophysiological processes, which is a topic that received less empirical attention. The idea that shedding emotional tears has beneficial (i.e., cathartic) effects on the crying individual has its roots in early psychodynamic theory, and is based on the assumption that holding back one's emotion has a negative impact on the organism (Freud, 1917; Frey, 1985). It is important to note that such a theoretical position mostly assumes a removal of negative impact of emotional inhibition, rather than the positive impact of emotional expression itself. Nevertheless, this group of studies tried to assess the effects of crying per se on decreases in mental and physical stress responses, and on mood improvement. These studies have shown contradictory results, with differences between findings being related to differences in the methodologies used. Systematic research on lay theories about strategies for decreasing negative emotion, analyses of popular-scientific and fiction literature, as well as media content in general revealed that people consider crying very beneficial in terms of bringing emotional relief (Cornelius, 1986, 1997, 2001; Hanser et al., 2016; Hendriks et al., 2008). In retrospective studies, i.e., when participants report on a previous (most often the last) crying episode, a large majority of them respond that their mood was less negative and more positive after, in comparison to before the crying episode (Bindra, 1972; Rottenberg et al., 2008b). Such findings are corroborated by studies in which therapists report on beneficial effects of their clients' crying (Nelson, 2008). Diary studies are expected to be more reliable when it comes to remembering crying events and less affected by lay theories about crying. In the only systematic diary study on crying, a mere 30% of participants reported mood improvements and 10% reported mood deterioration following crying, while the remaining majority did not report any mood changes (Bylsma et al., 2011). Finally, quasi-experimental research typically showed significant mood deterioration immediately following crying (e.g., Martin & Labbot, 1991; Gross et al., 1994; Rottenberg et al., 2002). One study that assessed mood at several time points has indicated that for the mood improvement to occur it might take a longer period (up to 90 minutes) after a crying episode (Gračanin et al., 2015; but see Gračanin et al., 2022). However, much larger effects observed by using such approach were those related to mood's returning to baseline. Specifically, after the initial large deterioration of mood, which was also observed in other quasi-experimental studies mentioned above, the mood of crying participants quickly returned to the initial level, corresponding to the one before the crying episode.

The existing explanations for the inconsistencies between the results of quasi-experimental and other studies mostly focus on methodological issues, ranging from low ecological validity of laboratory studies to the fact that these are quasi-experiments, rather than experiments. Specifically, crying elicitation in laboratory was limited to film clips and it was coupled with complete absence of (otherwise normative; see below) positive reactions from other people (i.e., experimenters) witnessing participants' crying. At the same time, a quasi-experimental approach does not allow researchers to disentangle the effects of stable individual differences or other forms of inter-individual variation from the effects of the crying process. Even if stable individual differences and emotional states are controlled for (see e.g., Gračanin et al., 2022), this type of design simply cannot de-couple differences in the spontaneous occurrence of crying from differences in *reasons* why crying does or does not occur—that is, from the differences in the triggers of crying. Unfortunately, the same problem is pertinent to diary studies as well. In addition to these basic methodological issues, there are important conceptual differences between different types of studies that could potentially

resolve this apparent paradox. In laboratory studies, mood is typically measured as a composite of different aspects of emotional states, such as sadness, anger, cheerfulness, or tension (e.g., Gračanin et al., 2022). A similar mood parameter was used in the above-mentioned diary study that also failed to find consistent beneficial effects of crying on mood (Bylsma et al. 2011). It might be worth considering the possibility that crying increases certain aspects of negative mood (e.g., sadness), while at the same time it decreases other aspects of mood, such as tension or anger. In retrospective studies, participants might typically refer to some of these aspects of mood due to various reasons (e.g., some negative emotional states are more arousing than others). To my knowledge, no published studies explored the possibility that different aspects of one's negative mood are differently affected by one's crying. Possibly, focusing on more specific aspects of mood or other psychological processes that are important for psychological well-being could help resolve the issue of inconsistent findings regarding the subjective well-being consequences of crying. On the other hand, it is possible that crying relatively rarely results in well-being improvements relative to a period immediately before its occurrence, as shown in the above-mentioned diary study (Bylsma et al., 2011), which was in many instances methodologically superior to many of the other studies mentioned. In that case, consistent results of studies other than quasi-experimental and diary studies, showing the beneficial effects of crying on mood, could be an artefact of what we focus on, how we experience, interpret, and remember mood changes. It is possible that people focus on large mood improvements that occur immediately after large mood deterioration following crying, such as those observed in Gračanin et al. (2015). As suggested by the authors of that study, such sudden mood improvements, that are nevertheless merely a return to baseline, might represent a salient experience that is remembered more easily in comparison to other mood changes following crying.

Whatever the typical or average effects of crying on mood, there is strong evidence of significant variation in these effects. Different characteristics of situations in which one cries influence mood changes following crying. For example, in situations that are less controllable, or in which observers react negatively to one's crying, mood improvement is less likely (Bylsma et al., 2008). Personality also plays an important role in determining whether someone will experience mood improvement following crying. More neurotic and conscientious individuals report less, and more extraverted and agreeable individuals report more mood benefits from crying (De Fruyt, 1997; Barthelmäs & Keller, 2021; Peter et al., 2001).

In cases when crying results in well-being improvements, these effects were proposed to be mediated by several possible mechanisms. The most important mechanism, for which there is also the most empirical support, is an *indirect* one based on observers' responses. Specifically, the strongest evidence regarding the effects of crying on mood pertains to mood changes that are the consequence of other individuals' positive responses to one's tears (Bylsma et al., 2008). In this case, all intra-individual effects of crying are merely byproducts of its inter-individual functions. On the other hand, empirical support for the more *direct* effects of crying on mood is inconsistent, as shown above. More direct mechanisms by which crying might bring benefits to crying individuals are in most cases those for which I propose to be functionally linked to its inter-individual functions. I propose that the intra-individual effects of crying are always based on its inter-individual functions. However, some of these are mere byproducts of the inter-individual functions, while others can have direct functional effects on the crying individual, as I will elaborate below.

Earlier accounts of the psychophysiological processes that might directly facilitate or reflect well-being improvements following crying promoted the idea that tears eliminate stress hormones and toxins from the blood (Frey, 1985). However, theoretical rationale for such claims is highly debatable and empirical evidence is missing (Vingerhoets, 2013; Vingerhoets & Kirschbaum, 1997). The majority of studies on psychophysiological concomitants of crying (as potential mediators between crying and well-being or mood improvements) focused on the (de)activation of sympathetic and parasympathet-

ic nervous systems. Activation of both systems has an important role for psychological well-being in domains of mood and prosocial behaviour (Porges, 2003; Thayer et al., 1996). A dozen of quasi-experiments indicated that the initial sympathetic increase at the onset of tearing turns into sympathetic decrease and parasympathetic increase around the offset of tearing with the latter showing relatively week effects (e.g., Hendriks et al., 2007; Rottenberg et al., 2003; see Zickfeld & Grüning, 2021 for meta-analysis). Notably, these quasi-experiments suffer from similar conceptual and methodological problems as the above-described laboratory studies on mood changes. They also suffer from additional issues, such as the lack of a clear definition of the onset and offset of crying (Zickfeld & Grüning, 2021). Other potential psychophysiological mechanisms involve changes in levels of oxytocin and endogenous opioids (see Vingerhoets, 2013). Few studies explored the role of these neurochemical substances, and they failed to support the hypotheses about their links with crying (see Gračanin et al., 2022). Largely missing from studies on psychophysiological concomitants or effects of crying are distinctions between various aspects of this multifaceted behaviour. It seems reasonable to ask whether different types of crying (e.g., tearing up vs vocal or sobbing) have different psychophysiological (and potentially resulting mood) consequences.

Finally, there is a possibility that in some individuals crying improves mood via various idiosyncratic factors, such as one's personal attitudes towards crying (Gračanin, 2023; Gračanin et al., 2014). For example, a person might assume that crying is beneficial or that "only good people cry" (see Zickfeld et al., 2022). Similarly, as stated above, the "effects" of crying that one remembers (and thus reports on) possibly depend on the salience of specific aspects of mood dynamics, which is simply part of the process of mood's returning to baseline after a stressful event, as explained above (Gračanin et al., 2015).

If our ancestors did not develop crying responses because these were useful for directly improving their mood, a possibility which any serious functional analysis could easily refute, then what was the reason for this curious behaviour to become part of our nature? The evolutionary perspective offers two straightforward, mutually exclusive hypotheses about the origin of crying and their effects on crying individuals and observers. Crying and its normative consequences (both intra- and inter-individual) exist because they either, on average, increased the chances for survival and reproduction in our ancestors, or they exist as a byproduct of other adaptive mechanisms. As we proposed earlier (Gračanin et al., 2017c, 2018a; Sznycer et al., 2025), inter-individual functions of crying can be understood on the basis of its ability to promote gene replication of crying individuals by serving communication functions. In contrast, intra-individual effects of crying in a large part (especially those related to mood changes) seem to be byproducts of its evolved inter-individual effects. However, some of these intra-individual effects are tightly related to inter-individual functions of crying and, as I will elaborate below, they should not be considered byproducts since they might represent part of the more global crying response, whose function is the preparation for and coordination of pro-social interactions.

Inter-Individual Functions of Crying

Theorising about evolved functions of tears started with Darwin (1872) who proposed that crying of infants and small children is functional, while crying in adults serves no purpose. Much later, Tomkins (1963) offered one of the first accounts of the signalling functions of crying by proposing that it signals distress and vulnerability, and promotes caregiving and social bonding behaviour. Two decades later, Frijda (1986) discussed communicative functions of this behaviour in terms of its ability to elicit sympathy and support. Bowlby's attachment theory offered one of the most elaborated and straightforward accounts of the adaptive value of crying (Bowlby, 1969, 1980). He considered crying as an important

component of the innate attachment mechanism. By this account, crying is a main "weapon" of infants, and it fulfils two main functions: (1) regulation of proximity with caregivers and (2) regulation of feeding and receiving other types of care. Finally, over the last three decades, a series of empirical contributions about the inter-individual effect of tears, rather than the more general crying behaviour (for a detailed overview, see Vingerhoets, 2013 and Zickfeld & Wróbel, 2024) has created a solid basis for theorising on their adaptive functions. Tears, as a key aspect of crying, were proposed to represent a signal of submission (Hasson, 2009), neediness and helplessness (Gračanin et al., 2018a; Miceli & Castelfranchi, 2003; Murube et al., 1999; Vingerhoets 2013; Zickfeld & Wróbel, 2024), and prosocial intentions (Gračanin et al., 2018a). Recently, we offered a comprehensive evolutionary account of tears, substantiated by strong evidence for their specific adaptive functions (Sznycer et al., 2025). We focused exclusively on tears, rather than the crying response as a whole, and we postulated that tears are a means of informing observers about the value one attaches to specific things. The transmission of information about one's specific valuation occurs mostly in negative situations when the individual (potentially) suffers damage to their reproductive prospects. It also happens when the observers have greater power to solve the situation in favour of the suffering individual than the individual does, and when there is a sufficient amount of shared interest between the individual and the observer. In positive situations where tearing sometimes occurs, the information about one's valuation is transmitted in order to create more lasting changes in observers' cognition, leading to better coordination between the tearful individual and observers over longer periods. This theoretical account is consistent with the notion that crying is typically triggered by the potential crier's perception of high relational warmth and competence.

Why and How Did Crying Evolve?

Crying likely evolved from distress calls, an infant behaviour that is omnipresent across bird and mammal species. Their function is to communicate distress and elicit care from caregivers. These calls are limited to sounds in all these species except in humans, who also produce emotional tears. Gračanin et al. (2017c, 2018a) proposed that during recent human evolution these calls initially consisted of vocal crying only. Such communication is useful because it transmits information (e.g., about a baby's hunger) on relatively large distances, in all directions, and behind obstacles. However, it is disadvantageous since it can attract predators and assaulters (Vingerhoets, 2013; Walter, 2006), and can be extremely annoying, resulting in aggression from others. Such a signal becomes less useful from the very moment small children become able to move towards caregivers. At the same time, combining physical approach with visual signal such as tears has an advantage of directing the distress call to specific individuals, which may also explain the decline in vocal crying with age (Gračanin et al., 2018). In addition, avoiding the spread of information about one's neediness and distress can be important for one's social rank even in small children (see Strayer & Trudel, 1984), which ultimately becomes one of the most important factors influencing the individual's survival and reproduction. A salient signal such as tears thus becomes crucial.

How did emotional tears evolve? Gračanin et al. (2017c, 2018a) proposed that tears evolved as a signal of neediness and distress first in infants, subsequently in older children, and finally in adults. A simple path that led to the emergence of this adaptation might have been based on the fact that animals produce reflex tears when muscles and other tissue around the eyes are mechanically pressed, which is typically observed in yawning or laughing in humans. It was proposed that such behaviours exert pressure on nerve fibres that innervate the lacrimal gland, resulting in excessive tear production (Murube, 2009). However, I could not find any studies comparing the amount of reflex tears produced in humans and other animals. I was also not able to find any scientific accounts of animals other than humans tearing up when yawning. Nevertheless, we do know that chimps, as our closest relatives, have much more robust facial musculature

around the eyes compared to ours. The exceptionally delicate and fragile human facial anatomy, including that of the eyes—especially in infants—could explain the more excessive production of reflex tears in response to mechanical pressure in humans. In our past, the regularity of tearing that was occurring during vocal crying led to the process known as *ritualisation*, which was proposed to be at the basis of the evolution of muscular facial expressions (e.g., anger displays; Eibl-Eibesfeldt, 1989). Through this process tears gradually became a source of information that is not necessarily tied to strong muscular movements around the eyes (Gračanin et al., 2018a).

Signal Value of Tears

What is the signal value of tears, that is, what information is carried by this signal? To answer that, it is first important to determine what the prototypical triggers of tears are. People most often cry when they incur costs, e.g., when experiencing losses, and less often when they enjoy benefits. In addition to external triggers, another important factor is the state of helplessness (e.g., Vingerhoets, 2013). In other words, the receiver of the signal also gains information about the crier's relative inability to deal with the problem or defend themselves (Sznycer et al., 2025; Gračanin et al., 2018a, Hasson, 2009). For example, people expect individuals to be more likely to tear up during negotiation about resources not only when the outcome of negotiation is more likely to be unfavourable but also when their hierarchical status and ability to return favours are lower rather than higher (Gračanin et al., 2023). When it comes to the role of tears in conflict and aggression, a notable finding is that individuals who show more nonverbal cues of dominance tend to cry less frequently (Gračanin et al., 2019). In short, tears carry information about one's low ability to act (i.e., *competence*) in relation to a certain situation or other individuals. Second, as I elaborate below, people more likely cry in the presence of emotionally closer individuals (e.g. Vingerhoets, 2013). In other words, crying seems to more likely occur in relationships characterised by more *warmth*. The remaining question is how exactly these aspects of social perception trigger the expressive (i.e., crying) component of emotion. Are there some prototypical aspects of physiological and motivational processes functionally related to appraisals of warmth and competence that precede or co-occur with crying?

Finally, for tears to be a signal, on average they must elicit a specific response in observers. Thus, it is also necessary to consider the normative impact of tears on observers' behaviour. Most studies trying to answer this question are experiments that typically consist of the presentation of tearful and non-tearful faces to participants, where tears were either removed from pictures of originally crying faces or artificially added to otherwise non-crying faces. When exposed to tearful faces, observers report feeling more readiness to help the depicted individual than when exposed to non-tearful faces (Bobowik et al., 2023; Provine et al., 2009; Vingerhoets et al., 2016; Zickfeld et al., 2021). Among few others, a recent study based on real life experience rather than on reactions to pictures corroborated and extended those results by showing the effects of tears on actual helping behaviour (Barthelmäs et al., 2024). Tears also decrease aggressive responses in observers, although the supporting evidence is still relatively sparse. In our study in which we exposed participants to vignettes describing mock trials, tearful defendants were punished less harshly than non-tearful ones (Picó et al., 2020). Observers' reactions in dependence of the different characteristics of crying were investigated to a much lesser extent. Tearing up in males was found to elicit observers' more positive reactions compared to the more intense response labelled as crying (Warner & Shields, 2007), or sobbing (Wong et al. 2011). These findings imply that not just more intensive tearing, but also other aspects of crying might be less functional in terms of facilitation of observers' positive responses. However, the research testing such possibility across various contexts is largely missing

Why and How Exactly Tears Affect Observers' Pro-Social Responses

Based on all the listed findings, it can be inferred that tears signal a need for physical proximity, help, and support. Similarly, tears may act as a plea for reduction of aggression in observers. These messages share a common theme: tearful individuals ask observers for some sort of investment, or more generally, for a "better treatment". Why would observers be inclined to forego certain resources in order to pass them to the crying individual? A crucial factor is the degree of stake in one's well-being—be it genetic relatedness or a history of cooperation. Indeed, people are more likely to cry in presence of close individuals than not close individuals (Bylsma et al., 2008; Lombardo et al., 1983; Sharman et al., 2019; Vingerhoets, 2013), and close individuals are more likely to respond with providing help or succour than non-close individuals or strangers (Barthelmäs et al., 2024; Bylsma et al., 2008; Hendriks et al., 2008). However, there is another implicit reason for observers to pass resources to the crying individual: the prospects of future cooperation, at least during the specific situation. Helping somebody who is not going to appreciate it or ceasing aggression towards somebody who will immediately attack is not a viable response. Thus, for the crying signal to fulfil its function, it has to carry information about the tearful individual's willingness to cooperate, both in the given situation as well as in the long run, so that the investment of the helper will be somehow paid back. Accordingly, tears were proposed to act as a signal of submissiveness (Hasson, 2009) and friendly intentions (Gračanin et al., 2018a). This is in line with recent research consistently showing that tearful individuals are perceived as friendlier, less aggressive, and warmer than non-tearful individuals (Hendriks et al., 2008; van de Veen et al., 2016; Vingerhoets et al., 2016). It is thus not surprising that when tears are present on one's face, observers also report that they feel more ready to approach (Hendriks & Vingerhoets, 2006), which ultimately results in helping (Vingerhoets et al., 2016), than when tears are not present. As mentioned earlier, some authors proposed that in addition to requesting help, crying also occurs when providing help, which was provided as an explanation for why people cry when experiencing positive emotion (Murube et al., 1999). However, it might be challenging to incorporate various instances of tears of joy, such us crying when witnessing beauty (see Zickfeld et al., 2020), into the function of offering help. A more general category that involves requesting help and offering help, that also has a potential for explain various forms of tears of joy, is the category of prosocial intentions. I speculate that witnessing any highly valued event activates a psychological system designed to facilitate collaboration by sharing the information about one's valuation of things (see also Sznycer et al., 2025). Clearly, the idea that tears may act as a signal of prosocial intentions or an invitation to cooperate has the potential to explain why tears occur in response to both negative and positive events, which otherwise poses a serious theoretical challenge.

What are the most basic cognitive mechanisms through which visual exposure to tears affects observers' prosocial responses? For example, do tears affect observers' behavioural tendencies at the early stages of cognitive processing? Gračanin et al. (2018b) investigated whether tears facilitate approach relative to avoidance tendencies in two different approach–avoidance tasks. Participants were instructed to either move towards tearful faces and away from non-tearful faces, or vice versa, using a joystick or pressing buttons. Tears promoted approach relative to avoidance behaviour, although it was unclear whether tears specifically facilitate approach, block avoidance tendencies in observers, or have both effects.

The presence of tears facilitates recognition or attribution of certain specific emotions. When tears are present on neutral, sad, or on faces expressing other primary emotion, tearful individuals are perceived as more sad than non-tearful individuals (Balsters et al., 2013; Gračanin et al., 2021; Ito et al., 2019; Provine et al., 2009; Reed et al., 2015). The perception of sadness from tears added to faces expressing different emotions can occur automatically, as shown in experiments with reaction-time based Implicit Association Test (Ong & Ito, 2022). Furthermore, tears on faces reliably indicate sadness even

in the presence of contextual information such as body postures, visual scenes, and written scenarios that suggest non-sadness emotions (Ito & Ong, 2023). However, the presence of tears facilitates, or at least does not block the perception of the expressions of fear and anger, compared to other expressions, such as surprise or disgust, the perception of which it strongly decreases. More specifically, we recently explored how tears act in combination with different muscular facial expressions (Gračanin et al., 2021). We tested three different hypotheses stating that the presence of tears increases the perception of: (1) the expressed emotion in general, (2) the expression of one specific emotion—sadness, or (3) helplessness-related cognitive appraisals and behavioural intentions of the crier. The experiment provided the most support for the third hypothesis, implying that tears do not carry information about one specific emotion, but rather inform observers about the costs (potentially) incurred by the crier and their inability to deal with those costs. These messages are typically conveyed by more than one type of emotional expression, and tears seem to facilitate the perception of such expressions. In line with that, in another study we found that observers perceive more intense guilt in tearful individuals than in non-tearful ones (Picó et al., 2020). Apart from affecting the perception of specific emotions, tears have the capacity to make various (but not all) expressions appear more intensive as well as more sincere. Again, this primarily applies to expressions linked to helplessness (Gračanin et al., 2021; Picó et al., 2020). Not surprisingly, tearful individuals are perceived as more helpless (Balsters et al., 2013) and needing more help, care, and succour than non-tearful individuals, which is one of the most straightforward findings in the research on crying (e.g., Cornelius & Lubliner, 2003; Vingerhoets et al., 2016; Zickfeld et al., 2021). When exposed to tearful faces, observers report experiencing more negative emotions themselves (Hendriks & Vingerhoets, 2006), but also more empathy and feelings of connectedness with the depicted individual, than when exposed to non-tearful faces (Provine et al., 2009; Vingerhoets et al., 2016; Zeifman & Brown, 2011; Zickfeld et al., 2021). Most importantly, the effects of tears on social support were found to be mediated by feelings of empathic concern, as well as by perceived helplessness, friendliness, and connectedness with crying individuals (Vingerhoets et al., 2016; Zickfeld et al., 2021).

Pro-Social Effects via Olfactory Channel?

In a paper from 2011, Gelstein and colleagues reported on the results of three studies showing decreased male sexual interest when exposed to fresh female tears through nasal inhalation. Importantly, they also observed meaningful hormonal changes during male exposure to tears, as well as less theoretically clear changes in the activity of the autonomic nervous system. Our team attempted to replicate the findings pertaining to the sexual interest, but we failed to observe any effects across three studies (Gračanin et al., 2017a; see also Gračanin et al., 2017b and Sobel, 2017). We concluded that sniffing female tears has little effect on male sexual interest and that it appears that tears act primarily as a visual signal. This conclusion is in line with the fact that evidence for the existence of pheromones in humans is weak (Wyatt, 2015). In addition, even if tears contained a chemosignal, there remains a question of its functionality in comparison to tears as a visual signal. While sniffing tears seems to require close proximity (note that in this line of studies participants sniffed tears directly from the test tubes), visual presentation of tears works at larger distance, which represents a more functional signal. Importantly, no other research teams reported on the systematic effects of sniffing female tears on male sexual behaviour. Interestingly, Agron et al. (2023) found that human female tears decrease aggressive behaviour of males, accompanied by theoretically relevant changes in brain activity, which might indirectly support the idea of olfactory effects of tears on sexual behaviour, and more generally, point to the role of olfaction in the evolution of tears as a signal. In the future, research in this field could provide valuable insights into the evolution and functions of tears.

Intra-Individual Effects and Inter-Individual Functions of Crying: An Integration

As stated above, potential direct intra-individual effects of crying that increase individual well-being are related to the inter-individual functions of tears. It is only through understanding the evolved signalling functions of tears that we can understand the processes through which crying impacts individuals' well-being. In this section, I elaborate on the link between the signalling function of tears—which implies that tears are a signal of helplessness and friendly intentions—and the inter-individual processes that occur during crying and that have the capacity to improve the well-being of the crier.

As I proposed above, crying is typically triggered by the appraisal of low agency (and more specifically, one's low relational competence) and high relational warmth. Its proposed function consists of motivating people who are potentially willing and more capable to provide positive value (or less negative value) by acting in favour of crying individuals than these individuals are capable themselves. Switching from active to passive coping is an adaptive solution in situations when resource expenditure is too high, futile, or dangerous. For example, giving up a task that is too demanding or impossible to solve, or ceasing active confrontation is potentially a better solution than insisting on behaviours that might decrease one's prospects (e.g., Nesse, 2000; Price et al., 1994). I propose that crying behaviour represents the signal component of a broader set of reactions in such situations, whose function is to prepare the organism for passive coping and reception of prosocial responses from nearby individuals (Gračanin et al., 2018a). Relatively loosely bundled emotional mechanisms (see Barrett, 2006) activated in such situations include crying as a signal component, along with other key aspects of emotions—cognitive appraisal, physiological changes, behavioural tendencies, and subjective feelings (see also Lange & Zickfeld, 2023). It was proposed that this specific pattern of reactions that follow the appraisal of low agency/competence together with the appraisal of a sufficient (typically large or suddenly increased; see Zickfeld et al., 2021) amount of relational warmth is coordinated by the *Prosocial Crying System* (PCS; Gračanin, 2023). The system is likely based on brain structures involved in generation of distress calls that we share with other mammals, labelled as the *cry circuit* (Newman, 2007). I further speculate that in humans such system may also overlap with or represent part of other, more general systems, such as the *Social Engagement System* that, among other functions, coordinates the activation of the parasympathetic nervous system and prepares the organism for pro-social interactions (Porges, 2003). PCS is proposed to coordinate tearing and other signalling crying responses, such as vocal crying, with specific patterns of autonomic activation, such as parasympathetic increases and sympathetic withdrawal, together with changes in endocrine system, such as those in oxytocin levels (Zickfeld & Grüning, 2021; Vingerhoets, 2013), which are all related to increases in positive and decreases in negative affect (e.g., Gianaros et al., 2002). In addition to promotion of passive pro-social behaviour, which also includes submissive and helpless responses, the workings of the PCS system thus may promote less negative and more positive subjective emotional states. The existence of such system may further help explain the apparent paradox of people crying when experiencing both positive and negative events. Its output may be useful across both positive and negative situations in which the crying individual communicates friendly intentions, with the function of diminishing losses (e.g., in case of events that provoke sadness, fear, or guilt) and/or fostering collaboration (e.g., during events that provoke joy, gratitude, or kama muta; see Zickfeld et al., 2019).

The initial activation of PCS together with its beneficial effects on well-being (e.g., in terms of increased parasympathetic activation or emotional relief) should occur relatively independently from the reactions of (or in the absence of) observers. However, once its activation is triggered, observers' response to its output might further modulate the effects of PCS on one's well-being, making them less beneficial in case of observers' negative responses, as shown by Bylsma et al. (2008). Specifically, certain social norms and (expected) observers' responses to one's crying (see MacArthur & Shields, 2019) can result in inhibition of

crying or one's negative reactions to one's own tears, such as experiencing feelings of shame or embarrassment (Becht & Vingerhoets, 2002; Bylsma et al., 2008). Modulation of the effects of PCS might occur relatively frequently, especially when it comes to observers' (expected) reactions to more conspicuous aspects of crying, such as vocal. Nevertheless, in the absence of such modulation, the *direct* outputs of PCS are expected to be beneficial in terms of physiology and subjective feelings, even when it comes to those outputs of PCS that sometimes evoke negative responses in observers. This may also explain why crying so often occurs in private (Vingerhoets, 2013) and why it is nevertheless generally considered as psychologically beneficial (Cornelius, 1986, 1997, 2001; Hanser et al., 2016; Hendriks et al., 2008). It is also in line with the above-cited findings about increases in parasympathethic activation following crying in laboratory experiments, in which participants do not experience reactions of other individuals (Zickfeld & Grüning, 2021). Mood changes might indeed represent a byproduct, while other intra-individual effects, such as physiological changes, energy conservation, and ceasing of dangerous activity might represent fulfilled intra-individual function. Clearly, these latter processes activated during switching to passive /prosocial mode could represent mediators between crying and mood improvements (see Gračanin et al., 2014; Vingerhoets, 2013). In short, the existence of PCS may explain why people report feeling better following crying even without observers' beneficial intervention. Admittedly, the presented integration model is in a very initial phase, while at the same time it is rather complex, which makes any attempt of its empirical evaluation a huge challenge. A general approach for testing the model could involve experiments that would manipulate appraisals of certain aspects of warmth and competence, measure spontaneously elicited tears and patterns of psychophysiological (ANS and hormonal) activity as mediators, and specific coping responses including motivation (e.g., for collaboration/submission), as well as subjective feelings and mood dynamics as outcomes.

Conclusion

I provided a brief overview of the latest knowledge on the intra- and inter-individual effects of emotional crying. If crying improves the well-being of the crying individual, it most often occurs through observers' responses. However, there are some indications that this might also happen directly and internally, through the crying process itself. In the final part of the current contribution, I offered a brief argumentation for why and how this might occur. Undoubtedly, much stronger or at least more documented beneficial effects of crying for one's well-being are those based on an indirect mechanism: eliciting help and support from others. There is now strong empirical support for the claim that tears can be seen as a signal of one's helplessness and willingness to cooperate, whose function is to promote helping responses in observers and to foster cooperation. These messages seem to be mediated by the tears' ability to affect observers' cognition by altering attributions of the states, intentions, and more stable characteristics of crying individuals. In both negative and positive situations, tears are highly informative regarding one's motivations, emotional states, coping abilities, and intentions, all of which can be subsumed within the dimensions of warmth and competence. Tears carry information about one's valuation of things, allowing observers to more efficiently provide, protect, or even benefit from what the crying individual values. The role of other aspects of crying, which can also be regarded as more intensive forms, is less clear and deserves more research.

References

Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. *Psychological Review*, *128*(2), 290–314. https://doi.org/10.1037/rev0000262

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, *93*(5), 751–763. https://doi.org/10.1037/0022-3514.93.5.751

Agron, S., de March, C.A., Weissgross, R., Mishor, E., Gorodisky, L., Weiss, T., Furman-Haran, E., Matsunami, H., & Sobel, N. A. (2023). Chemical signal in human female tears lowers aggression in males. *PLoS Biolology, 21*(12):e3002442. https://doi.org/10.1371/journal.pbio.3002442

Balsters, M. J., Krahmer, E. J., Swerts, M. G., & Vingerhoets, A. J. J. M. (2013). Emotional tears facilitate the recognition of sadness and the perceived need for social support. *Evolutionary Psychology, 11*(1), 148-158. https://doi.org/10.1177/147470491301100114

Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on Psychological Science, 1*, 28–58. https://doi.org/10.1111/j.1745-6916.2006.00003.x

Barthelmäs, M., & Keller, J. (2021). Adult emotional crying: Relations to personality traits and subjective well-being. *Personality and Individual Differences, 177*, 110790. https://doi.org/10.1016/j.paid.2021.110790.

Barthelmäs, M., Kesberg, R., Hermann, A., & Keller, J. (2022). Five reasons to cry—FRC: A taxonomy for common antecedents of emotional crying. *Motivation and Emotion, 46*(3), 404–427. https://doi.org/10.1007/s11031-022-09938-1

Barthelmäs, M., Stöckle, D., & Keller, J. (2024). On the social signal function of emotional crying: Broadening the perspective to social interactions in daily life. *Emotion, 24(4)*, 960–974. https://doi.org/10.1037/emo0001313

Becht, M. C., & Vingerhoets, A. J. J. M. (2002). Crying and mood change: A cross-cultural study. *Cognition and Emotion, 16*, 87–101. https://doi.org/10.1080/02699930143000149

Bindra, D. (1972). Weeping: A problem of many facets. *Bulletin of the British Psychological Society, 25*, 281–285.

Bobowik, M., Doroszuk, M., Slawuta, P., & Basabe, N. (2023). When they cry: Tears facilitate responses toward members of socially disadvantaged groups. *Emotion, 23(2)*, 521-537. https://doi.org/10.1037/emo0001085

Bowlby, J. (1969). *Attachment and loss (Vol. 1)*. Basic Books.

Bowlby, J. (1980). *Attachment and loss* (Vol. 3*): Loss, sadness, and depression*. Basic Books.

Bylsma, L. M., Croon, M. A., Vingerhoets, A. J. J. M., & Rottenberg, J. (2011). When and for whom does crying improve mood? A daily diary study of 1004 crying episodes. *Journal of Research in Personality, 45*, 385–392. http://dx.doi.org/10.1016/j.jrp.2011.04.007

Bylsma, L. M., Gračanin, A., & Vingerhoets, A. J. J. M. (2021). A clinical practice review of crying research. *Psychotherapy, 58*(1), 133–149. https://doi.org/10.1037/pst0000342

Bylsma, L. M., Vingerhoets, A. J. J. M., & Rottenberg, J. (2008). When crying is cathartic? An international study. *Journal of Social and Clinical Psychology, 27,* 1165–1187. http://dx.doi.org/10.1521/jscp.2008.27.10.1165

Choti, S. E., Marston, A. R., Holston, S. G., & Hart, J. T. (1987). Gender and personality variables in film-induced sadness and crying. *Journal of Social and Clinical Psychology, 5*(4), 535–544. https://doi.org/10.1521/jscp.1987.5.4.535

Cornelius, R. R. (1986). Prescience in the pre-scientific study of weeping: A history of weeping in the popular

press from the mid-1800's to the present. *75th Annual Meeting of the Eastern Psychological Association*. NewYork, USA.

Cornelius, R. R. (1997). Toward a new understanding of weeping and catharsis? In: A. J. J. M. Vingerhoets, F. J. van Bussel, & A. J. W. Boelhouwer (Eds.), *The (non)expression of emotions in health and disease* (pp. 303–321). Tilburg University Press.

Cornelius, R. R. (2001). Crying and catharsis. In: A. J. J. M. Vingerhoets, & R. R. Cornelius (Eds.), *Adult crying: A biopsychosocial approach* (pp. 199–212). Brunner-Routledge.

Cornelius, R. R., & Lubliner, E. (2003). The what and why of others' responses to our tears: Adult crying as an attachment behavior. *3rd international conference on The (Non)Expression of Emotions in Health and Disease*, Tilburg, the Netherlands.

Crile, G. W. (1915). *The origin and nature of the emotions*. Saunders.

Darwin, C. (1872). *The expression of emotions in animals and man*. Oxford University Press (1998 edition, with an introduction, afterword, and commentaries by P. Ekman).

De Fruyt, F. (1997). Gender and individual differences in adult crying. *Personality and Individual Differences, 22*, 937–940. https://doi.org/10.1016/S0191-8869(96)00264-4

Denckla, C. A., Fiori, K. L., & Vingerhoets, A. J. J. M. (2014). Development of the crying proneness scale: Associations among crying proneness, empathy, attachment, and age. *Journal of Personality Assessment, 96*(6), 619–631. https://doi.org/10.1080/00223891.2014.899498

Eibl-Eibesfeldt, I. (1989). *Human ethology.* Aldine de Gruyter.

Ekman, P., & Friesen, W. V. (2003). *Unmasking the face*. ISHK/Malor Books.

Freud, S., (1917). The history of the Psychoanalityc movement. *Nervous and mental disease monograph series, 25*. Nervous and Mental Disease Pub. Co.

Frey, W. H. (1985). *Crying: The mystery of tears*. Winston Press.

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77-83. https://doi.org/10.1016/j.tics.2006.11.005

Frijda, N. H. (1986). *The emotions*. Cambridge University Press.

Gelstein, S., Yeshurun, Y., Rozenkrantz, L., Shushan, S., Frumin, I., Roth, Y., & Sobel N. (2011). Human tears contain a chemosignal. *Science*, 14; 331(6014), 226-230. https://doi.org/10.1126/science.1198331

Gianaros, P. J., Quigley, K. S., Mordkoff, J. T., & Stern, R. M. (2001). Gastric myoelectrical and autonomic cardiac reactivity to laboratory stressors. *Psychophysiology, 38*, 642-652. https://doi.org/10.1017/S0048577201001093

Gračanin, A. (2023). Suze i raspoloženje: intraindividualni učinci emocionalnoga plakanja (Tears and Mood: Intra-Individual Effects of Emotional Crying). *Psihologijske teme, 32*, 663-686. https://doi.org/10.31820/pt.32.3.13

Gračanin, A., Alaburić, N., Kardum, I., & Vingerhoets, A. J. J. M. (2023, October). *Why we cry: Factors that influence our preparedness to cry in negotiations*. 8th international conference on The (Non)Expression of Emotions in Health and Disease, Tilburg, the Netherlands.

Gračanin, A., Bylsma, L., & Vingerhoets, A. J. J. M. (2014). Is crying a self-soothing behavior. *Frontiers in Psychology, 5,* 502. https://doi.org/10.3389/fpsyg.2014.00502

Gračanin, A., Bylsma, L. M., & Vingerhoets, A. J. J. M. (2017c). The communicative and social functions of human crying. In J.-M. Fernández-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 217–233). Oxford University Press.

Gračanin, A., Bylsma, L., & Vingerhoets, A. J. J. M. (2018a). Why only humans shed emotional tears: Evolutionary and cultural perspectives. *Human Nature, 29*(2), 104–133. https://doi.org/10.1007/s12110-018-9312-8

Gračanin, A., Hendriks, M. C. P., & Vingerhoets, A. J. J. M. (2022). Crying does not alleviate acute pain percep-

tion: Evidence from an experimental study. *European Journal of Pain, 26*(6), 1380-1394. https://doi.org/10.1002/ejp.1970

Gračanin, A., Krahmer, E., Balsters, M., Küster, D., & Vingerhoets, A. J. J. M. (2021). How weeping influences the perception of facial expressions: The signal value of tears. *Journal of Nonverbal Behavior, 45*, 83-105. https://doi.org/10.1007/s10919-020-00347-x

Gračanin, A., Krahmer, E., Rinck, M., & Vingerhoets, A. J. J. M. (2018b). The effects of tears on approach-avoidance tendencies in observers. *Evolutionary Psychology, 16*(3), 1-10. https://doi.org/10.1177/1474704918791058

Gračanin A., van Assen. M. A., Omrčen, V., Koraj, I., & Vingerhoets, A. J. J. M. (2017a) Chemosignalling effects of human tears revisited: Does exposure to female tears decrease males' perception of female sexual attractiveness? *Cognition and Emotion, 31*(1), 139-150. https://doi.org/10.1080/02699931.2016.1151402

Gračanin, A., van Emden, R., Kardum, I., & Vingerhoets, A. J. J. M. (2019, October). *Why am I treated so bad? The effects of tears in the context of submission/dominance and welfare-tradeoff ratios*. 7th international conference on The (Non)Expression of Emotions in Health and Disease, Tilburg, the Netherlands.

Gračanin, A., Vingerhoets, A. J. J. M., & van Assen, M. A. (2017b). Response to comment on "Chemosignalling effects of human tears revisited: Does exposure to female tears decrease males' perception of female sexual attractiveness?". *Cognition and Emotion, 31*(1), 158-159. https://doi.org/10.1080/02699931.2016.1182471

Gračanin, A., Vingerhoets, A. J. J. M., Kardum, I., Zupčić, M,. Šimić, M., & Šantek, M. (2015). Why crying does and sometimes does not seem to alleviate mood: A quasi-experimental study. *Motivation and Emotion, 39,* 953–960. https://doi.org/10.1007/s11031-015-9507-9

Gross, J. J., Fredrickson, B. L., & Levenson, R. W. (1994).The psychophysiology of crying. *Psychophysiology, 31*, 460–468. https://doi.org/10.1111/j.1469-8986.1994.tb01049.x

Hanser, W. E., ter Bogt, T. F. M., Van den Tol, A. J. M., Mark, R. E., & Vingerhoets, A. J. J. M. (2016). Consolation through music: A survey study. *Musicae Scientiae, 20*(1), 122–137. https://doi.org/10.1177/1029864915620264

Hasson, O. (2009). Emotional tears as biological signals. *Evolutionary Psychology,* 7, 363–370. https://doi.org/10.1177/147470490900700302

Hendriks, M. C. P., Nelson, J. K., Cornelius, R. R., &Vingerhoets, A. J. J. M. (2008). Why crying improves our well-being: An attachment-theory perspective on the functions of adult crying. In: A. J. J. M. Vingerhoets, I. Nyklicek, & J. Denollet (Eds.), *Emotion regulation: Conceptual and clinical issues* (pp. 87–96). Springer.

Hendriks, M. C. P., Rottenberg, J., & Vingerhoets, A. J. J. M. (2007). Can the distress- signal and arousal-reduction views of crying be reconciled? Evidence from the cardiovascular system. *Emotion, 7*(2), 458–463. https://doi.org/10.1037/1528- 3542.7.2.458

Hendriks, M. C. P., & Vingerhoets, A. J. J. M. (2006). Social messages of crying faces: Their influence on anticipated person perception, emotional and behavioral responses. *Cognition and Emotion*, 20(6), 878–886. https://doi.org/10.1080/02699930500450218

Ito, K., & Ong, C. W. (2023). Perception of emotional tears with body postures, visual scenes, and written scenarios. *Asian Journal of Social Psychology, 26*(1), 52–68. https://doi.org/10.1111/ajsp.12544

Ito, K., Ong, C. W., & Kitada, R. (2019). Emotional tears communicate sadness but not excessive emotions without other contextual knowledge. *Frontiers in Psychology*, 10:878. https://doi.org/10.3389/fpsyg.2019.00878

Laan, A., Van Assen, M. A. L. M., & Vingerhoets, A. J. J. M. (2012). Individual differences in adult crying: The

role of attachment styles. *Social Behavior and Personality, 40*, 453–471. http://dx.doi.org/10.2224/sbp.2012.40.3.453

Lange, J., & Zickfeld, J. H. (2023). Comparing implications of distinct emotion, network, and dimensional approaches for co-occurring emotions. *Emotion, 23*(8), 2300–2321. https://doi.org/10.1037/emo0001214

MacArthur, H. J., & Shields, S. A. (2019). How you cry, when you cry, why you cry, and who you are: Responses to adult crying in social contexts. In U. Hess & S. Hareli (Eds.), *The social nature of emotion expression: What emotions can tell us about the world* (pp. 209–225). Springer. https://doi.org/10.1007/978-3-030-32968-6_12a

Martin, R. B., & Labott, S. M.(1991). Mood following emotional crying: Effects of the situation. *Journal of Research in Personality, 25*, 218–244. http://dx.doi.org/10.1016/0092-6566(91)90017-K

Miceli, M., & Castelfranchi, C. (2003). Crying: Discussing its basic reasons and uses. *New Ideas in Psychology, 21*, 247–273. https://doi.org/10.1016/j.newideapsych.2003.09.001

Murata, K., Nagasawa, M., Onaka, T., Kanemaki, N., Nakamura, S., Tsubota, K., Mogi, K., & Kikusui, T. (2022). Increase of tear volume in dogs after reunion with owners is mediated by oxytocin. *Current Biology, 32*(16), 869-870. https://doi.org/10.1016/j.cub.2022.07.031

Murube, J. (2009). Basal, Reflex, and Psycho-emotional Tears. *The Ocular Surface, 7*(2), 60-66. https://doi.org/10.1016/s1542-0124(12)70296-3

Murube, J., Murube, L., & Murube, A. (1999). Origin and types of emotional tearing. *European Journal of Ophthalmology, 9*, 77–84.

Nelson, J. K. (2008). Crying in psychotherapy: Its meaning, assessment and management based on attachment theory. In: A. J. J. M. Vingerhoets, I. Nyklicek, & J. Denollet (Eds.), *Emotion regulation: Conceptual and clinical issues* (pp. 202–214). Springer.

Nesse, R. M. (2000). Is depression an adaptation? *Archives of General Psychiatry, 57*(1), 14-20. https://doi.org/10.1001/archpsyc.57.1.14

Newman, J. D. (2007). Neural circuits underlying crying and cry responding in mammals. *Behavioural Brain Research, 182*(2), 155-165. https://doi.org/10.1016/j.bbr.2007.02.011

Ong, C. W., & Ito, K. (2022). Can't fight seeing sadness in tears: Measuring the implicit association between tears and sadness. *British Journal of Social Psychology, 61*(2), 672-687. https://doi.org/10.1111/bjso.12503

Patel, V. (1993). Crying behavior and psychiatric disorder in adults: A review. *Comprehensive Psychiatry, 34*, 206–211. https://doi.org/10.1016/0010-440X(93)90049-A

Peter, M., Vingerhoets, A. J. J. M., & Van Heck, G. L. (2001). Personality, gender, and crying. *European Journal of Personality, 15*, 19–28. https://doi.org/10.1002/per.386

Picó, A., Gračanin, A., Gadea, M., Boeren, A., Aliño, M., & Vingerhoets, A. J. J. M. (2020). How visible tears affect observers' judgements and behavioral intentions: Sincerity, remorse, and punishment. *Journal of Nonverbal Behavior, 44(2)*, 215–232. https://doi.org/10.1007/s10919-019-00328-9

Porges, S. W. (2003). Social engagement and attachment: A phylogenetic perspective. *Annals of the New York Academy of Sciences, 1008,* 31–47. https://doi.org/10.1196/annals.1301.004

Price, J., Sloman, L., Gardner, R. Jr, Gilbert, P., & Rohde, P. (1994). The social competition hypothesis of depression. *British Journal of Psychiatry, 164*(3), 309-315. https://doi.org/10.1192/bjp.164.3.309

Provine, R. R., Krosnowski, K. A., & Brocato, N. W. (2009). Tearing: Breakthrough in human emotional signaling. *Evolutionary Psychology, 7*(1), 52–56. https://doi.org/10.1177/147470490900700107

Reed, L. I., Deutchman, P., & Schmidt, K. L. (2015). Effects of tearing on the perception of facial expressions of emotion. *Evolutionary Psychology, 3(4)*, 1–5. https://doi.org/10.1111/ajsp.12544

Rottenberg, J., Bylsma, L. M., Wolvin, V., & Vingerhoets, A. J. J. M. (2008). Tears of sorrow, tears of joy: An in-

dividual differences approach to crying in Dutch females. *Personality and Individual Differences, 45*, 367–372. https://doi.org/10.1016/j.paid.2008.05.006

Rottenberg, J., Wilhelm, F. H., Gross, J. J., & Gotlib, I. H. (2003). Vagal rebound during resolution of tearful crying among depressed and nondepressed individuals. *Psychophysiology, 40*(1), 1–6. https://doi.org/10.1111/1469-8986.00001

Rottenberg, J., Gross, J. J., Wilhelm, F. H., Najmi, S., & Gotlib, I. H. (2002). Crying threshold and intensity in major depressive disorder. *Journal of Abnormal Psychology, 111*, 302–312. https://doi.org/10.1037/0021-843X.111.2.302

Sharman, L. S., Dingle, G.A., Baker, M., Fischer, A., Gračanin, A., Kardum, I., Manley, H., Manokara, K., Pattara-Angkoon, S., Vingerhoets, A. J. J. M., & Vanman, E. J. (2019). The relationship of gender roles and beliefs to crying in an international sample. *Frontiers in Psychology*, 10:2288. https://doi.org/10.3389/fpsyg.2019.02288

Sobel, N. (2017). Revisiting the revisit: added evidence for a social chemosignal in human emotional tears. *Cognition and Emotion, 31*(1), 151-157. https://doi.org/10.1080/02699931.2016.1177488

Strayer, F. F., & Trudel, M. (1984). Developmental changes in the nature and function of social dominance among young children. *Ethology and Sociobiology, 5*(4), 279–295. https://doi.org/10.1016/0162-3095(84)90007-4

Sznycer, D., Gračanin, A., & Lieberman, D. (2025). Emotional tears: What they are and how they work. *Evolution and Human Behavior, 46*(1), 1-12. https://doi.org/10.1016/j.evolhumbehav.2025.106652

Thayer, J. F., Friedman, B. H., & Borkovec, T. D. (1996). Autonomic characteristics of generalized anxiety disorder and worry. *Biological Psychiatry, 39*, 255–266. https://doi.org/10.1016/0006-3223(95)00136-0

Tomkins, S. S. (1963). *Affect, imagery, consciousness: II. The Negative Affects*. Springer.

Van de Ven, N., Meijs, M., & Vingerhoets, A. J. J. M. (2016). What emotional tears convey: Tearful individuals are seen as warmer, but also as less competent. *British Journal of Social Psychology, 56*(1), 146–160. https://doi.org/10.1111/bjso.12162

Vingerhoets, A. J. J. M. (2013). *Why only humans weep: Unravelling the mysteries of tears*. Oxford University Press.

Vingerhoets, A. J. J. M., & Bylsma, L. M. (2016). The riddle of human emotional crying: A challenge for emotion researchers. *Emotion Review, 8,* 207–217. http://dx.doi.org/10.1177/1754073915586226

Vingerhoets, A. J. J. M., Bylsma, L. M., & Rottenberg, J. (2009). Crying: A biopsychosocial phenomenon. In: T. Fögen (Ed.), *Tears in the Graeco-Roman World* (pp. 439-476). De Gruyter. https://doi.org/10.1515/9783110214024.439

Vingerhoets, A. J. J. M., Hoevenaar, N. M., & Gračanin, A. (2018, November). Tears and disgust: Emotional basis of our moral compass? *Consortium of European Research on Emotion* conference. Glasgow, United Kingdom.

Vingerhoets, A. J. J. M., & Kirschbaum, C. (1997). Crying, mood and cortisol. *Paper Presented at the Annual Meeting of the American Psychosomatic Society*. Santa Fe, NM, USA (abstracted in *Psychosomatic Medicine, 59*, 92–93).

Vingerhoets, A. J. J. M., van de Ven, N., & van der Velden, Y. (2016). The social impact of emotional tears. *Motivation and Emotion, 40*, 455–463. https://doi.org/10.1007/s11031-016-9543-0

Vingerhoets, A. J. J. M., van Geleuken, A. J. M. L., van Tilburg, M. A. L., & van Heck, G. L. (1997). The psychological context of crying episodes: Toward a model of adult crying. In: A. J. J. M. Vingerhoets, F. J. Van Bussel, & A. J. W. Boelhouwer (Eds.), *The (non)expression of emotions in health and disease* (str. 323 – 336). Tilburg University Press.

Walter, C. (2006). *Thumbs, toes, and tears: And other traits that make us human*. Walker & Co.

Warner, L. R., & Shields, S. A. (2007). The perception of crying in women and men: Angry tears, sad tears, and the "right way" to cry. In U. Hess & P. Philippot (Eds.), *Group dynamics and emotional expression* (pp. 92–117). Cambridge University Press. https://doi.org/10.1017/CBO9780511499838.006

Wong, Y. J., Steinfeldt, J. A., LaFollette, J. R., & Tsao, S.-C. (2011). Men's tears: Football players' evaluations of crying behavior. *Psychology of Men & Masculinity, 12*(4), 297–310. https://doi.org/10.1037/a0020576

Wyatt, T. D. (2015). The search for human pheromones: The lost decades and the necessity of returning to first principles. *Proceedings of the Royal Society: Biological Sciences, 282*, 1–9. https://doi.org/10.1098/rspb.2014.2994

Zeifman, D. M., & Brown, S. A. (2011). Age-related changes in the signal value of tears. *Evolutionary Psychology, 9*(3), 313–324. https://doi.org/10.1177/147470491100900304

Zickfeld, J. H., & Grüning, D. J. (2021). *A meta-analysis on the autonomic nervous system correlates of human emotional crying*. https://doi.org/10.31234/osf.io/axjd5

Zickfeld, J. H., Kamble, S. V., Oostelbos, M. C. J. M., & Vingerhoets, A. J. J. M. (2022). Only the good cry: Investigating the relationship between crying proneness and moral judgments and behavior. *Social Psychological Bulletin*, *17*, Article e6475. https://doi.org/10.32872/spb.6475

Zickfeld, J. H., Schubert, T. W., Seibt, B., Blomster, J. K., Arriaga, P., Basabe, N., Blaut, A., Caballero, A., Carrera, P., Dalgar, I., Ding, Y., Dumont, K., Gaulhofer, V., Gračanin, A., Gyenis, R., Hu, C.-P., Kardum, I., Lazarević, L. B., Mathew, L., . . . Fiske, A. P. (2019). Kama muta: Conceptualizing and measuring the experience often labelled being moved across 19 nations and 15 languages. *Emotion, 19*(3), 402–424. https://doi.org/10.1037/emo0000450

Zickfeld, J., Seibt, B., Lazarevic, L., Zezelj, I., and Vingerhoets, A. J. J. M. (2020). *A model of positive tears*. PsyArXiv [Preprint]. doi: 10.31234/osf.io/sf7pe

Zickfeld, J. H., van de Ven, N., Pich, O., Schubert, T. W., Berkessel, J. B., Pizarro, J. J., Bhushan, B., Mateo, N. J., Barbosa, S., Sharman, L., Kökönyei, G., Schrover, E., Kardum, I., Aruta, J. J. B., Lazarevic, L., Escobar, M. J., Stadel, M., Arriaga, P., Dodaj, A., … Vingerhoets, A. (2021). Tears evoke the intention to offer social support: A systematic investigation of the interpersonal effects of emotional crying across 41 countries. *Journal of Experimental Social Psychology, 95*, 104-137. https://doi.org/10.1016/j.jesp.2021.104137

Zickfeld, J. H., & Wróbel, M. (2024). Emotional tears as social motivators: When and how tearing up motivates social support. *Social and Personality Psychology Compass*, e12921. https://doi.org/10.1111/spc3.12921

3

Determinants of Successful Ageing and Longevity: How to Age Well and Achieve a Good Old Age?

Ivana Tucak Junaković
Department of Psychology, University of Zadar
itucakju@unizd.hr

Abstract

The extension of the average human lifespan and the increasingly better quality of life in the newer generations of older people have increased scientific interest in the positive aspects of ageing and old age and the factors that contribute to longevity and successful ageing. Successful ageing is one of the most represented but also the most controversial topics in contemporary gerontology. Researchers in this field have reached a consensus that successful ageing is a complex multidimensional construct that incorporates various domains of functioning (not just physical) and requires a holistic approach. However, there is no consensus regarding the components of this complex construct and the most appropriate way to measure it. In the first part of the article, the concept of successful ageing and the differences compared to similar concepts such as healthy, active, productive, or quality ageing were explained. Furthermore, different theoretical approaches, such as biomedical, psychosocial, and lay, and recent holistically oriented multidimensional models of successful ageing, along with methods of research on successful ageing were described. In the second part of the article, the key determinants of successful ageing and longevity were presented. These were confirmed by numerous studies, including those conducted in samples of Croatian older people. An attempt was made to explain in more detail the role of different socio-demographic characteristics, lifestyle factors, personality traits, psychological resources, social relationships, and social participation, as well as cognitively stimulating activities in the context of successful/active/ageing well and improving the quality of life in old age.

Keywords: *determinants of successful ageing, longevity, research methods on successful ageing, theoretical approaches to successful ageing, successful ageing*

Introduction

Demographic trends towards an increase in the average human lifespan and an increasingly better quality of life in newer generations of older people have led to an increase in scientific interest in factors that contribute to longevity and good or successful ageing. Croatia joins these demographic trends because we are at the very top of Europe in terms of the proportion of older people. According to the latest Census from 2021, people aged 65 and over made up even 22.45% of the total population of the Republic of Croatia (Croatian Bureau of Statistics, 2022), and a comparison with the results of previous Censuses shows that the proportion of older people is continuously increasing.

In the context of these demographic trends and contemporary shift in scientific interest towards positive aspects of ageing, the first aim of this study is to present the concept of successful ageing, and to explain different theoretical approaches and methods of research on successful ageing. The second aim is to present the key determinants of successful ageing and longevity, including socio-demographic characteristics, lifestyle factors, personality traits, psychological resources, social relationships, and social participation, as well as cognitively stimulating activities.

Successful Ageing and Related Concepts

Ageing is a term that usually has a negative connotation and is traditionally defined as a universal progressive process of decline in all functions that ends in death. However, according to the lifespan developmental perspective, ageing and development are intertwined processes, where successful adaptation to increasing losses in ageing process is key to successful ageing (Baltes, 1987). Both development and ageing are multidimensional and multidirectional processes that include physical, cognitive, and psychosocial dimensions that are not affected to the same extent by degenerative processes associated with increasing age. Therefore, there is a space for developmental gains and for the preservation of psychological vitality and social engagement in old age (Achenbaum, 2001). Newer generations of older people, despite their great heterogeneity, are generally living longer, are healthier, better informed, more educated, and more active in social and public life, and this trend is likely to continue in the future.

Some abilities and traits are very well preserved or progress until a very late age (e.g., satisfaction with life, positive emotionality, generativity, wisdom, crystallised knowledge and intelligence, semantic memory, etc.). Therefore, negative stereotypes about older people are less and less justified, and facts support a much more favourable image of old age and ageing than that held by many people. Therefore, over the past two decades, there has been a visible shift in scientific and professional interest towards positive aspects of ageing and old age and factors of longevity. In this context, concepts such as successful, active, healthy, vital, productive, quality ageing, etc. have emerged (Bowling, 2007; Depp & Jeste, 2006; Fernandez-Ballesteros et al., 2013). They are all very similar and relate to the fundamental idea of ageing well (Fernandez-Ballesteros et al., 2013). However, each of them emphasises different components of ageing well. Thus, the term *healthy ageing* primarily refers to good health and preserved functional ability (Depp & Jeste, 2006). The term *active ageing*, which has often been used lately, was introduced by the World Health Organisation in 2002. It encompasses the health, active social participation, and safety of older people and is frequently used in a socio-political context for the purpose of influencing public policies (Fernandez-Ballesteros et al., 2013). *Successful ageing* is a concept that has a broader meaning since it encompasses both the biomedical and psychosocial dimensions. Some authors believe that it is an umbrella construct that encompasses healthy, productive, and active ageing as narrower constructs (Fernandez-Ballesteros et al., 2013; Fernandez-Ballesteros, 2019; Urtamo et al., 2019). Compared to other similar terms, the term successful ageing is the most frequently used in gerontological research.

Criticism of the Concept of Successful Ageing

Despite its contribution in the context of emphasising individual differences in the ageing process and the possibilities of more quality ageing, the concept of successful ageing has also been criticised (e.g., Liang and Luo, 2012; Masoro, 2001). The leading objection to the very idea of "success" in the ageing process is that it indirectly implies that there are also individuals who age unsuccessfully or "losers" in the ageing process. Furthermore, some believe that the paradigm of successful ageing imposes excessive individual responsibility for one's own ageing process, ignoring the influence of genetic, environmental, and other factors that the individual cannot influence. Also, emphasising the importance of successful ageing can create pressure on individuals to be healthy, productive, and active, with the ultimate goal of reducing their idependence on social institutions (Rubinstein & de Medeiros, 2014).

These criticisms are not entirely unfounded, and it is important to emphasise that the inevitable losses in the ageing process should not be underestimated, nor should exceptionally successful ageing be made an imperative. On the other hand, it is a great merit of research on successful, active, quality ageing, etc., that it has emphasised the underestimated positive aspects of ageing. It has also helped identify those factors (biological, psychosocial, lifestyle factors, etc.) that contribute most to a quality life in old age and longevity.

Theoretical Approaches and Methods of Research on Successful Ageing

Scientific interest in successful ageing emerged together with the beginnings of gerontology as a scientific discipline (Tucak Junaković, 2022). The major gerontological theories in the field of social sciences first deal with successful ageing in a more systematic and scientifically based manner. Thus, *the disengagement theory* (Cumming & Henry, 1961) presents the idea that the best way to age successfully for an elderly person is to gradually withdraw from society through a reduction in social contacts, activities, social roles, and responsibilities. On the other hand, according to another influential gerontological theory that emerged at the same time, *the activity theory* (Havighurst, 1961), an active life and productive involvement in various types of activities are associated with greater life satisfaction in old age. Similarly, *the continuity theory* (Atchley, 1989) proposes that the continuity of previous activities, lifestyle, social relationships, personal characteristics, etc. in later life is crucial for successful ageing.

In the context of the first major theories dealing with successful ageing, it is necessary to mention Erik Erikson and his *theory of psychosocial development*. According to this theory, in the last eighth stage of the lifecycle, a person tries to find meaning in their past life and accept it with all its good and bad sides. If they succeed, they will develop ego integrity in old age or the sense that their life has been meaningful. Such a person can rightly be said to have aged successfully. Therefore, the concept of ego integrity can be considered an earlier form of the concept of successful ageing (Martin et al., 2015).

Although the first major gerontological theories mentioned tried to answer the question of how to age successfully, it was not until the late 1980s that physician John Rowe and psychologist Robert Kahn popularised the concept of successful ageing in the scientific community. In their influential article, Rowe and Kahn (1987) described the difference between pathological, normal, and successful ageing. Rowe and Kahn's well-known model of successful ageing will be discussed in the subsequent section that addresses biomedical approach to successful ageing. Namely, in the contemporary scientific literature, three basic theoretical approaches to successful ageing can be identified: biomedical, psychosocial, and lay.

Biomedical Model of Successful Ageing

The biomedical approach to successful ageing is still dominant, within which Rowe and Kahn (1997) proposed the most influential model of successful ageing to date. According to this model, successful ageing includes three components: (1) the absence or low probability of disease and disease-related disability, (2) the maintenance of physical and cognitive functioning at a high level, and (3) continued engagement with life. The model was tested in the well-known longitudinal MacArthur study of successful ageing (Rowe & Kahn, 1998). Its results emphasised the key role of physical activity and involvement in productive and social activities in preserving good physical and cognitive functioning over the years.

Within the biomedical model, successful ageing is usually conceptualised as a dichotomous variable, i.e., individuals are identified as "successful" or "unsuccessful". It is examined with objective measures such as the presence/absence of chronic conditions typical of the elderly population (hypertension, diabetes, arthritis, cardiovascular diseases, etc.), and measures of functional ability, i.e., degree of independence in performing daily activities (e.g., Activities of Daily Living Scale-ADLS, Katz et al., 1970; The Instrumental Activities of Daily Living Scale-IADLS, Lawton & Brody 1969). In research based on the biomedical model, an individual's cognitive status is often assessed through short cognitive assessment measures, such as the Mini-Mental State Examination-MMSE (Folstein et al., 2011) or the Clifton Assessment Procedures for the Elderly-CAPE (Pattie & Gilleard, 1996).

However, it is not easy for older people to meet the strict criteria of successful ageing according to the biomedical model, especially the one that refers to the absence or low incidence of disease and functional limitations (Bowling & Dieppe, 2005; Strawbridge et al., 2002). It is (almost) impossible to live to a very old age without chronic conditions and with preserved independence in daily functioning (Anderson-Ranberg et al., 2001; Cho et al., 2012). The biomedical approach has also been objected for neglecting other dimensions of successful ageing, such as psychosocial and spiritual dimensions (Young et al., 2009).

Psychosocial Model of Successful Ageing

In response to the limitations of the biomedical model, psychosocial and lay approaches have emerged. According to the psychosocial approach, the basic components of successful ageing are quality social relationships and social participation, as well as positive psychological states and resources (satisfaction with life, self-esteem, resilience, optimism, autonomy, good adaptation to losses in the ageing process, effective coping strategies, etc.). According to one of the most influential psychological models of successful ageing, *the model of selective optimisation with compensation* (Baltes & Baltes, 1990), successful ageing is a process of good adaptation to changes throughout life, that is, the process of maximising developmental gains and regulating and minimising losses through the processes of selection, optimisation, and compensation.

Studies within the psychosocial approach capture successful ageing with measures of life satisfaction or subjective well-being, social participation, and personal resources such as resilience, self-efficacy, or optimism (Tucak Junaković et al., 2024). For example, the Satisfaction with Life Scale-SWLS (Diener et al., 1985), The Connor-Davidson Resilience Scale-CD-RISC (Connor & Davidson, 2003), Life Orientation Test-Revised-LOT-R (Scheier et al., 1994), Self-Efficacy Scale–SES (Sherer et al., 1982), and other measures have been used in these studies.

According to the psychosocial approach, even people with impaired health can age successfully if they are satisfied with life, relationships with loved ones, and if they perceive their life as fulfilling and meaningful (Stewart et al., 2019; Strawbridge et al., 2002). This implies that it is very important to consider how older people themselves see their own ageing process. This subjective perception often deviates significantly from the so-called objective or biomedical criteria of successful ageing.

Lay Model of Successful Ageing

Scientists have eventually realised that when developing research definitions of successful ageing, it is useful to consider the perspective of older people themselves. Insight into how older people in a particular socio-cultural setting perceive successful ageing can also be useful when developing public health and other intervention programmes aimed at promoting successful ageing in that specific setting (Bowling, 2006; Tucak Junaković, 2022).

Lay concepts of successful ageing have been examined in research using qualitative methods such as focus groups or interviews with smaller groups of older people who are asked how they perceive successful ageing and the contributing factors. Previous research has shown that older people view successful ageing in a much more complex way than researchers do (Ambrosi-Randić & Plavšić, 2008; Bowling, 2006; Reichstadt et al., 2007; Tucak Junaković & Ambrosi-Randić, 2022). In addition to the components of successful ageing that biomedical and psychosocial models include, lay concepts also include many other components to which researchers have paid less attention in their theoretical concepts and operationalisations. These include, for example, material status, a sense of achievement, contribution to others and society as a whole, a sense of meaning in life, concern for one's physical appearance, etc. A recent meta-analysis that compared lay definitions of successful ageing identified in studies conducted in 13 countries across Western Europe, North America, the Middle East, Asia, and Oceania, from 2010 to 2020, showed that older people most often cite social participation and a positive attitude as components of successful ageing. They have also frequently mentioned independence and physical health (Reich et al., 2020). In one of more recent qualitative studies that examined how older people in Croatia perceive successful ageing and its determinants, it was found that Croatian older people see physical/physiological factors (e.g., general health, physical mobility, healthy habits such as a healthy diet and non-smoking), then psychological factors (e.g., a good life, life satisfaction, subjective well-being, psychological resources such as adaptability and a positive attitude), and community involvement and activity (e.g., being generally active, socialising, continuing previous activities in old age, engaging in enjoyable hobbies) as key components of successful ageing (Tucak Junaković & Ambrosi-Randić, 2022).

While the influence of cultural norms and social expectations on shaping perception of (successful) ageing should not be ignored, there appear to be some universal attributes of successful ageing that people across different cultures value as important, such as good health, quality social relationships, or financial security.

The Newer Generation Models of Successful Ageing

The recent generation models of successful ageing are multidimensional and holistically oriented. They capture the physical, functional, and psychosocial dimensions of successful ageing, as well as the so-called objective indicators (e.g., chronic diseases and functional ability), and subjective assessment of "success" (Cosco et al., 2015; Fernandez-Ballesteros et al., 2013; Kleineidam et al., 2019; Parslow et al., 2011; Pruchno et al., 2010; Vahia et al., 2012; Young et al., 2009). For example, Young et al. (2009) have presented a multidimensional model with physiological, psychological, and social component. The physiological component captures chronic diseases and functional limitations. The psychological component refers to the assessment of cognitive status, emotional vitality, and depression. The social one includes social engagement and spirituality. The model assumes large individual differences and various possible paths to successful ageing, as well as the possibility to compensate for physical decline through preserved psychological and social resources.

Furthermore, Pruchno et al. (2010) have presented *a two-factor model of successful ageing* that

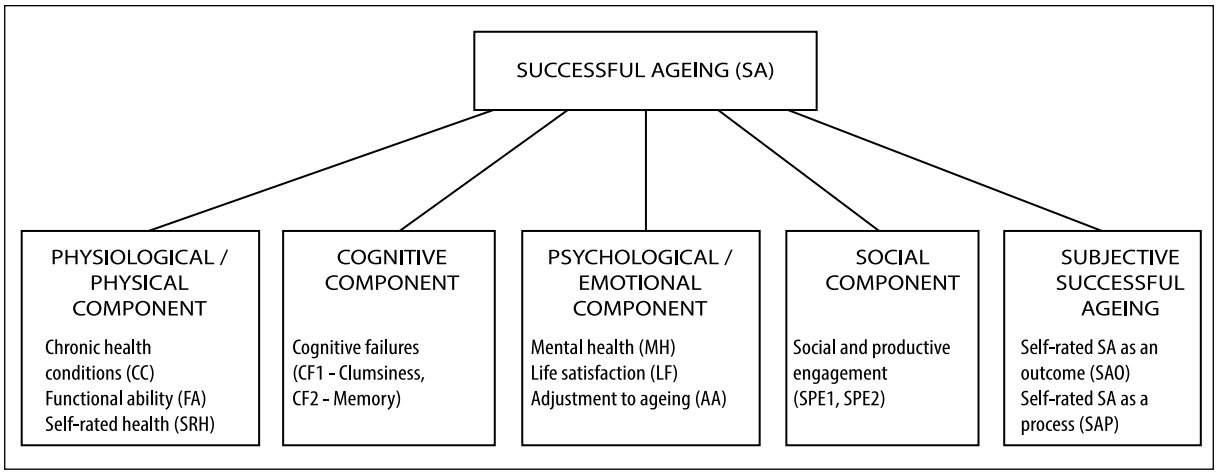
includes subjective and objective components. The subjective component captures a) self-assessment of successful ageing as an outcome assessed at a certain point in time, b) self-assessment of good ageing as an ongoing process, and c) life satisfaction. The objective component includes a) the number of chronic conditions, b) the assessment of functional ability, and c) the pain assessment. Vahia et al. (2012) supplemented the model of Pruchno et al. (2010) by adding the following components: cognitive ability, mood, and certain psychosocial characteristics (i.e., resilience, self-efficacy, optimism, and attitude towards own ageing). They have called this extended model *the model of successful cognitive and emotional ageing.*

Furthermore, one of the more recent models of successful ageing proposed by Fernandez-Ballesteros (2019) also includes biomedical components (i.e., health and independence in performing daily activities, physical and cognitive functioning) and psychosocial components (i.e., good psychological adaptation and involvement in social and productive activities), as well as subjective and objective dimensions of successful ageing. According to this model, as mentioned earlier, successful ageing is an overarching construct that encompasses healthy, active, and productive ageing, and the underlying thread of all of them is the idea of *ageing-well*.

Finally, it is worth mentioning another recent model of successful ageing developed within the research project *Successful Ageing: Development and Validation of an Integrative Multidimensional Model* (IP.01.2021.21) at the University of Zadar, Croatia. This multidimensional model of successful ageing, presented in Figure 1, includes five latent factors (with their indicators specified in brackets). These are: 1) Physiological/physical (Chronic health conditions (CC)—reverse coded, Functional ability (FA), Self-rated health (SRH)), 2) Cognitive (Cognitive failures 1 (CF1)—Clumsiness factor, Cognitive failures 2 (CF2)—Memory factor; both reverse coded), 3) Psychological/emotional (Mental health (MH), Life satisfaction (LF), Adjustment to ageing (AA)), 4) Social (Social and productive engagement 1 (SPE1)—activities in which older people had more opportunities to participate (e. g., helping friends or family members, attending religious activities), Social and productive engagement 2 (SPE2)—activities in which older people had probably less opportunity to participate (e.g., attending cultural activities, attending various educations, courses, or public lectures), and 5) Subjective (Self-rated SA as an outcome (SAO)—self-assessment on a scale from 0 to 10, Self-rated SA as a process

Figure 1

An integrative multidimensional model of successful ageing (adapted according to Tucak Junaković et al., 2024)



(SAP)—self-assessment on a scale from 0 to 10) successful ageing factors (Figure 1).

The construct validity of the five-component model was tested using a confirmatory factor analysis, and by comparing the model with several other theoretical models with two, three, or four components of successful ageing. The results confirmed that the five-component model, presented in Figure 1, fits the data best (Tucak Junaković et al., 2024). There is also initial evidence (Tucak Junaković, 2024) of the predictive validity of the proposed five-component model in predicting some objective and subjective outcome criteria

of successful ageing, such as health care and informal care utilisation, and ego-integrity or sense of meaning in life as the outcome of the last stage in the lifecycle (Erikson & Erikson, 1998). Therefore, this five-factor model could be seen as a step forward in the development of a sound and comprehensive definition of successful ageing (Tucak Junaković et al., 2024), which could be relevant in predicting specific outcomes.

In conclusion on theoretical conceptions and operationalisations of successful ageing, we can state that researchers in this field have proposed numerous definitions and measures of successful ageing. They reached an agreement that successful ageing is a complex multidimensional construct. However, they have not reached a consensus regarding components that it includes and the best method of measurement.

Depending on the theoretical conceptions and methods of assessing successful ageing, estimates of the proportion of older people who are ageing successfully (when successful ageing is viewed as a process) or have aged successfully (when successful ageing is viewed as an outcome) vary greatly across different studies. For example, in a systematic review of research on successful ageing by Cosco et al. (2014), the percentage of individuals identified as those ageing successfully ranged widely across different studies, from less than 1% to more than 90%. In general, a small percentage of older people, especially the oldest old, age successfully if the so-called objective criteria of absence of disease and disease-related limitations are taken as criteria for successful ageing. On the other hand, most older people assess themselves as successful agers (Montross et al., 2006). This is because they meet their own subjective and psychosocial criteria for successful ageing, such as good psychological adjustment, life satisfaction, good relationships with others, and the sense of purpose. For example, in a recent qualitative study conducted on a sample of Croatian older adults, even 92% of interviewed older people rated their ageing as successful (Tucak Junaković & Ambrosi-Randić, 2022).

The following chapters will provide a systematic overview of the various factors that influence successful ageing, regardless of its theoretical conceptualisation and operationalisation. Systematising knowledge about these factors is also important for practical reasons, as various preventive and intervention measures aimed at promoting healthy and successful ageing and quality of life in old age are based on these factors (i.e., those that we can influence, such as physical activity or social participation).

Determinants of Successful Ageing and Longevity: What Facilitates Good Ageing?

Longevity is determined by numerous factors such as hereditary and environmental factors, diseases, and lifestyle (Depp et al., 2007). The role of genetic factors in the process of (successful) ageing and determining the length of life is important and cannot be ignored. This is well demonstrated by the fact that parents and siblings of individuals who lived a long time were often long-living persons. It is estimated that hereditary factors can explain between 20 and 35% of individual differences in successful ageing, or in longevity as an indicator of successful ageing (Finch & Tanzi, 1997; Glatt et al., 2007). From the above, it can be indirectly concluded that contextual and lifestyle factors, which are subject to change, play a very important role in the process of successful ageing and longevity. Here, research in the field of epigenetics is very promising because it shows that our lifestyle can influence gene expression in such a way that, under favourable circumstances, genes that would otherwise prompt the onset of a disease can remain latent. The role of lifestyle, as well as other factors such as sociodemographic characteristics, personality traits, psychological strengths, social participation and support, in the process of (successful) ageing will be discussed next.

Sociodemographic Characteristics

Sociodemographic characteristics such as gender, age, education, marital status, or finances certainly influence the success with which a person adapts to changes in the ageing process. The results of

various studies are not entirely consistent, but they quite convincingly support the fact that in the older population, the proportion of successful agers decreases with age (Depp & Jeste, 2006; Hank, 2011; Meng & D'Arcy, 2014; Nosraty et al., 2012; Parslow et al., 2011; Pruchno et al., 2010). This may be attributed to better health and functional independence of younger people within older population.

There is still no clear answer to the question of whether men or women age more successfully. Women live on average four to seven years longer than men, so if we take life expectancy as an indicator of successful ageing, then we can say that women are ageing more successfully. Studies conducted in groups of older people of a wide age range have confirmed that women age more successfully (Ng et al., 2009). On the other hand, studies show that in the over-90s and centenarians' groups, men are more likely to be classified as successful (Araujo et al., 2016; Nosraty et al., 2012). Although contrary to the findings mentioned earlier, these gender differences are consistent with the well-known paradox that women have higher overall morbidity rates despite living longer than men. In the oldest age group of people over 85, there are more (sick) women because most men have already died by then, and those who have reached this late age are probably the most resilient. Furthermore, older women are at greater risk of poverty and widowhood than older men and generally have lower socioeconomic status, and research clearly shows that lower socioeconomic status and widowhood are associated with poorer health and lower health-related quality of life (Cherepanov et al., 2010).

Some studies, however, did not confirm gender differences in successful ageing, mainly those that used psychosocial criteria for successful ageing (Montross et al., 2006).

Finances and education also play an important role in the process of successful ageing and achieving a long life. Good financial status, objective or perceived, is associated with more successful ageing in groups of centenarians as well as in more age-heterogeneous samples of older adults (Araujo et al., 2016; Feng et al., 2014). Similarly, better education contributes to successful ageing (Depp & Jeste, 2006; Hank, 2011; Ng et al., 2009; Nosraty et al., 2012; Pruchno et al., 2010; Tucak Junaković, Ambrosi-Randić et al., 2023). Higher education is associated with better psychological resources such as more adaptive ways of coping with life stressors or greater awareness of the importance of a healthy lifestyle. It is also related to a greater cognitive reserve that will enable better cognitive functioning in old age, that is, more successful cognitive ageing. Furthermore, better finances enable better health care and easier satisfaction of basic existential and other needs that are important for successful ageing. Lay descriptions of successful ageing also place great importance on financial security as an essential component or determinant of successful ageing (Cosco et al., 2013; Tucak Junaković & Ambrosi-Randić, 2022).

Regarding the role of partner/marital status, research most often shows that older people who live with a partner are healthier, more satisfied with life, and live longer, that is, age more successfully, compared to single, divorced, or widowed people (Bowling & Iliffe, 2006; Meng & D'Arcy, 2014; Nosraty et al., 2012; Pruchno et al., 2010). The quality of the relationship, that is, social support and closeness with the partner, is probably more important here than marital status per se.

Lifestyle Factors

One of the most important determinants of successful ageing is a healthy lifestyle that includes physical activity, a healthy diet (e.g., Mediterranean), non-smoking, avoiding excessive alcohol consumption, etc. Regular physical activity and a healthy, balanced, and moderate diet are foundations of healthy and successful ageing. There is a lot of evidence about beneficial effects of physical activity on both physical and mental health. Physical activity in older people reduces the risk of a number of health problems, such as hypertension, cardiovascular diseases, osteoporosis, type 2 diabetes, high cholesterol, some types of cancer, etc. (Vuori, 2005). It is also associated with a lower incidence of mental health problems, such as depressive and anxiety disorders (Kadariya et al., 2019; Maier et al., 2021; Tucak Junaković, Nekić et al., 2023). Physical

activity also reduces the risk of developing dementia and slows cognitive decline (Larson et al., 2006; Sofi et al., 2011). It also slows down the decline of functional ability and reduces the frequency of falls in old age, thereby enabling a longer independent living (Spirduso & Cronin, 2001; Vuori, 2005). Epidemiological follow-up studies show that regular physical activity reduces mortality risk (Kokkinos et al., 2010). It can also contribute to successful ageing and longevity in other indirect ways. For example, through organised group exercise, older adults socialise and feel less lonely and isolated; exercise strengthens their sense of self-efficacy, independence, and personal control (Lepan & Leutar, 2012; Štambuk & Tomičić, 2020), which facilitate successful ageing.

Many studies have confirmed the association between physical activity and successful ageing (e.g., Almeida et al., 2014; Depp & Jeste, 2006; Meng & D'Arcy, 2014; Parslow et al., 2011; Pruchno et al., 2010), with physical activity being one of its most important protective factors. It should be emphasised that in these studies successful ageing is operationalised as a broader construct that includes not only physical and health dimensions, but also psychosocial dimension.

In the context of health behaviour, we cannot ignore the impact of smoking and alcohol consumption on healthy and successful ageing. The proven harmful effects of smoking on health are beyond discussion, and studies clearly confirm that smoking also negatively contributes to successful ageing in general (Depp & Jeste, 2006; Gureje et al., 2014). Regarding alcohol consumption, the findings are not entirely unambiguous, but they are closer to the conclusion that people who consume alcohol moderately (but not excessively) are healthier and age more successfully compared to those who do not drink alcoholic beverages at all (Feng et al., 2014; Maraldi et al., 2009; Pruchno et al., 2010; Tucak Junaković, Nekić et al., 2023). This is usually explained by moderate drinking as a social ritual that maintains social contacts, because in this case people drink while socialising with others, and social interaction significantly contributes to successful ageing, as we will see later in this paper.

Personality Traits and Psychological Resources

How successfully an individual adapts to changes in the ageing process and faces different life challenges is largely determined by their personality. Traits from the five-factor model (McCrae & Costa, 2004) were most often examined in relation to different developmental outcomes. These studies have shown that the traits of extraversion, emotional stability, conscientiousness, and openness to experience are associated with more successful ageing (Ambrosi-Randić, 2022; Chapman et al., 2010; Martin et al., 2006; Pocnet et al., 2020; Tucak Junaković et al., 2018). They achieve this connection mainly because of their influence on the tendency to experience stress and on health and other behaviours that are associated with successful ageing and longevity. For example, conscientious people practice a healthier lifestyle, have more stable relationships and better coping strategies, while extraverted people are less likely to experience stress, are more sociable and are at lower risk of loneliness, which makes it easier for both to age successfully and reduces the risk of mortality (Chapman et al., 2010; Pocnet et al., 2020). On the other hand, neurotic people have a harder time dealing with life problems, are more prone to experiencing stress and practising unhealthy behaviours, which threatens their quality of life and increases the risk of mortality (Chapman et al., 2010; Pocnet et al., 2020).

Furthermore, there is a series of evidence that the personal resources or psychological strengths such as optimism, resilience, sense of control, adaptability, effective coping with stress, self-esteem, self-efficacy, spirituality, sense of meaning in life, etc. facilitate successful ageing (Baltes & Baltes, 1990; Cosco et al., 2015; Heckhausen & Schultz, 1993; Martin-Joy & Vaillant, 2010; Ng et al., 2009; Tucak Junaković et al., 2017; Tucak Junaković, Ambrosi-Randić et al., 2023; Vahia et al., 2012). Optimism (as long as it is not excessive and unrealistic) plays a particularly important role in making life's challenges easier to deal with. It is often mentioned, along with physical activity and nutrition, as one of the pillars of successful ageing (Ambrosi-Randić

& Plavšić, 2008). Optimism has a positive effect on successful ageing and longevity primarily due to better emotional regulation in optimists and less focus on the negative aspects of different situations, which helps them cope better with unpleasant events (Ambrosi-Randić, 2022; Carver & Scheier, 2014). Similarly, self-efficacy, or an individual's belief in their own ability to achieve set goals (Bandura, 1994), is associated with healthier behaviours and the avoidance of unhealthy ones, and with better functioning of the immune system (Ambrosi-Randić, 2022). All of this links self-efficacy with successful ageing and longevity.

Social Participation, Social Relationships, and Social Support

A lot has already been written about the importance of forming and maintaining quality social relationships and the importance of the social support we receive from such relationships for our health, quality and long life. It seems important to mention here that older people compensate for the narrowing of their social network that occurs in old age by investing more in and nurturing those very close and emotionally fulfilling relationships (English & Carstensen, 2014). Social relationships, social participation, and social support are associated with better health and functional status, greater life satisfaction, higher self-esteem, and a sense of personal control in older people; they facilitate adjustment to retirement and, in general, promote good ageing and improve the quality of life in old age (Barbosa et al., 2016; Cosco et al., 2015; Douglas et al., 2017; Montross et al., 2006; Pruchno et al., 2010). Social interactions and other productive activities in the community that strengthen a positive self-image and a sense of self-efficacy also have a positive impact on the cognitive functioning of older adults. They thus reduce the risk of developing dementia, while experience of loneliness increases this risk (Wang et al., 2002; Wilson et al., 2007). Social support and socialising with friends and family members can even reduce the risk of death, while social exclusion and a lack of social support are associated with higher mortality rates, as evidenced by longitudinal studies (e.g., Dalgard & Lund Håheim, 1998).

Cognitively Stimulating Activities and Lifelong Learning

A good path to successful (cognitive) ageing is cognitive investment in various activities throughout life and training cognitive abilities in old age (Vranić, 2022). Numerous studies show that an active lifestyle and stimulating everyday activities (not only intellectual, but also physical and social) significantly contribute to the preservation of cognitive functions in old age and successful ageing in general. Greater investment in formal and informal education (e.g., studying, learning foreign languages, using new information and communication technologies) and cognitively stimulating activities in leisure time (e.g., playing board games or reading) build up a cognitive reserve and reduce the risk of dementia (Cujzek & Vranić, 2017; Lustig et al., 2009; Vranić, 2022). *Cognitive reserve* is a term that refers to accumulation of various experiences, knowledge, and skills that enable good cognitive functioning in the presence of negative neural changes (Vranić, 2022). A high cognitive reserve reduces the risk of developing dementia. Different types of cognitive training or structured exercises aimed at strengthening specific cognitive abilities are also associated with successful (cognitive) ageing (see more on this topic in Vranić, 2022).

Conclusion

Newer generations of older people are living longer and better-quality lives. They are more active in social and public life. Therefore, scientists and experts in the field of gerontology are increasingly interested in the positive aspects of ageing and old age and the factors contributing to longevity. There are more and

more facts that support the conclusion that adverse changes in the ageing process are often not as great or limiting for the daily functioning of older people as is usually believed. Also, despite health and functional limitations, it is shown that most older people are satisfied with life and perceive it as meaningful, which is, ultimately, probably the best indicator of the "success" of ageing process.

There are numerous determinants of successful or good ageing and longevity that can be influenced and thus improve the quality of life in old age. These are, first and foremost, healthy habits and behaviours, but also adaptive psychological traits and strengths, such as optimism and self-efficacy. A supportive social environment that provides an individual with a sense of belonging and support is also important. In this context, all public health programmes and individual psychological and other interventions aimed at strengthening the protective factors of successful ageing and a long quality life are extremely important. However, it should be emphasised that these interventions must be adapted to the condition and needs of each individual because there are different individual paths to successful ageing.

As a summary of all the previously said about the factors of good ageing and longevity, the experiences of long-living people in the so-called Blue Zones can be used. These are places on Earth where people live the longest, such as Sardinia in Italy, Okinawa in Japan, or Ikaria in Greece. What they all have in common and what contributes to their long and quality lives are some common features of their lifestyle. First, they are moderate in eating habits, practising a semi-vegetarian diet rich in plants, legumes, whole grains, and fish, with occasional fasting. Then, moderate physical activity is included in their natural, everyday movements such as gardening or walking. Furthermore, residents of these areas attach great importance to good family relationships. They are also actively involved in the community and experience their lives as purposeful and meaningful (Buettner, 2012).

In the end, it can be concluded that there is a significant body of findings on factors that contribute to successful ageing and longevity. Still, there are many open questions to which scientists have yet to find answers with the ultimate goal of achieving the best possible quality of life in old age (and in earlier stages of the life cycle). In this context, research in the field of epigenetics, which integrates insights from behavioural sciences and genetics, seems particularly promising.

Acknowledgements

This work is conducted within the research project "Successful Ageing: Development and Validation of an Integrative Multidimensional Model" (IP.01.2021.21), supported by the University of Zadar, Croatia.

References

Achenbaum, W. A. (2001). Productive aging in historical perspective. In N. Morrow-Howell, J. Hinterlong, M. Sherraden (Eds.), *Productive aging: Concepts and challenges* (pp. 19–36). The Johns Hopkins University Press.

Almeida, O. P., Karim, M. K., Hankey, G. J., Bu, B. Y., Golledge, J., & Flicker, L. (2014). 150 minutes of vigorous physical activity per week predicts survival and successful ageing: A population-based 11-year longitudinal study of 12201 older Australian men. *British Journal of Sports Medicine*, *48*(3), 220–225. https://doi.org/10.1136/bjsports-2013-092814.

Ambrosi-Randić, N. (2022). Emocije, ličnost i socijalni odnosi u kontekstu uspješnog starenja (Emotions, personality and social relations in the context of successful aging). In I. Tucak Junaković (Eds.), *Uspješno starenje: teorijski pristupi, metode istraživanja i čimbenici doprinosa (Successful aging: theoretical approaches, research methods and contributing factors)* (pp. 101-143). University of Zadar.

Ambrosi-Randić, N., & Plavšić, M. (2008). *Uspješno starenje (Successful aging).* Društvo psihologa Istre, Istarska županija, Sveučilište Jurja Dobrile u Puli.

Anderson-Ranberg, K., Schroll, M., & Jeune, B. (2001). Healthy centenarians do not exist, but autonomous centenarians do: A population-based study of morbidity among Danish centenarians. *Journal of American Geriatric Society*, *49*, 900–908. https://doi.org/10.1046/j.1532-5415.2001.49180.x

Araujo, L., Ribeiro, O., Teixeira, L., & Paul, C. (2016). Predicting successful aging at one hundred years of age. *Research on Aging*, *38*(6), 689–709. https://doi.org/10.1177/0164027515603771

Atchley, R. C. (1989). A continuity theory of normal aging. *The Gerontologist*, *29*(2), 183–190. https://doi.org/10.1093/geront/29.2.183

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, *23*(5), 611–626. https://doi.org/10.1037/0012-1649.23.5.611

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In P. B. Baltes, M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1–34). Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511665684

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Eds.), *Encyclopedia of human behavior* (pp. 71–81). Academic Press.

Barbosa, L. M., Monteiro, B., & Murta, S. G. (2016). Retirement adjustment predictors – A systematic review. *Work, Aging and Retirement*, *2*(2), 262–280. https://doi.org/10.1093/workar/waw008

Bowling, A. (2006). Lay perceptions of successful ageing: Findings from a national survey of middle aged and older adults in Britain. *European Journal of Ageing*, *3*(3), 123–136. https://doi.org/10.1007/s10433-006-0032-2

Bowling, A. (2007). Aspirations for older age in the 21st century: What is successful aging? *International Journal of Aging & Human Development*, *64*(3), 263–297. https://doi.org/10.2190/L0K1-87W4-9R01-7127

Bowling, A., & Dieppe, P. (2005). What is successful ageing and who should define it? *British Medical Journal*, 331, 24–31. https://doi.org/10.1136/bmj.331.7531.1548

Bowling, A., & Iliffe, S. (2006). Which model of successful aging should be used? Baseline findings from a British longitudinal survey of ageing. *Age and Ageing*, *35*(6), 607–614. https://doi.org/10.1093/ageing/afl100

Buettner, D. (2012). *The Blue Zones, Second Edition: 9 Lessons for Living Longer From the People Who've Lived the Longest.* Disney Publishing Group.

Carver, S. C., & Scheier, M. F. (2014). Dispositional optimism. *Trends in Cognitive Sciences*, *18,* 293–299. https://doi.org/10.1016/j.tics.2014.02.003

Chapman, B. P., Fiscella, K., Kawachi, I., & Duberstein, P. (2010). Personality, socioeconomic status, and all-cause mortality in the United States. *American Journal of Epidemiology*, *171,* 83–92. https://doi.org/10.1093/aje/kwp323

Cherepanov, D., Palta, M., Fryback, D. G., & Robert, S. A. (2010). Gender differences in health-related quality-of-life are partly explained by sociodemographic and socioeconomic variation between adult men and women in the US: Evidence from four US nationally representative data sets. *Quality of Life Research*, *19*(8), 1115–1124. https://doi.org/10.1007/s11136-010-9673-x

Cho, J., Martin, P., & Poon, L. W. (2012). The older they are, the less successful they become? Findings from the Georgia Centenarian study. *Journal of Aging Research*, 2012, Article ID 695854. https://doi.org/10.1155/2012/695854

Connor, K. M., & Davidson, J. R. T. (2003). Development of a new resilience scale: The Connor-Davidson Resilience Scale (CD-RISC), *Depression and Anxiety, 18*(2), 76–82. https://doi.org/10.1002/da.10113

Cosco, T. D., Prina, A. M., Perales, J., Stephan, B. C. M., & Brayne, C. (2013). Lay perspectives of successful ageing: A systematic review and metaethnography. *BMJ Open, 3*(6). https://doi.org/10.1136/bmjopen-2013-002710

Cosco, T. D., Prina, A. M., Perales, J., Stephan, B. C. M., & Brayne, C. (2014). Operational definitions of successful aging: A systematic review. *International Psychogeriatrics, 26* (3), 373–381. https://doi.org/10.1017/S1041610213002287

Cosco, T. D., Stephan, B. C. M., & Bryne, C. (2015). Validation of an a priori, index model of successful aging in a population-based cohort study: The successful aging index. *International Psychogeriatrics*, *27*(12), 1971–1977. https://doi.org/10.1017/S1041610215000708

Cujzek, M., & Vranić, A. (2017). Computerized tabletop games as a form of a video game training for old-old. *Aging. Neuropsychology and Cognition, 24*(6), 631–648. https://doi.org/10.1080/13825585.2016.1246649

Cumming, E., & Henry, W. E. (1961). *Growing old: The process of disengagement*. Basic Books.

Dalgard, O. S., & Lund Håheim, L. (1998). Psychosocial risk factors and mortality: A prospective study with special focus on social support, social participation, and locus of control in Norway. *Journal of Epidemiology and Community Health, 52*(8), 476-81. https://doi.org/10.1136/jech.52.8.476

Depp, C. A., Glatt, S. J., & Jeste, D. V. (2007). Recent advances in research on successful or healthy aging. *Current Psychiatry Reports*, *9* (1), 7–13. https://doi.org/10.1007/s11920-007-0003-0

Depp, C. A., & Jeste, D. V. (2006). Definitions and predictors of successful aging: A comprehensive review of larger quantitative studies. *The American Journal of Geriatric Psychiatry*, *14*, 6–20. https://doi.org/10.1097/01.JGP.0000192501.03069.bc

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Jornal of Personality Assessment, 49,* 71–76. https://doi.org/10.1207/s15327752jpa4901_13

Douglas, H. B., Georgiou, A., & Westbrook, J. (2017). Social participation as an indicator of successful aging: An overview of concepts and their associations with health. *Australian Health Review*, *41*, 455–462. https://doi.org/10.1071/AH16038

Državni zavod za statistiku. Konačni rezultati Popisa 2021: Stanovništvo prema starosti, Popis 2021. (Croatian Bureau of Statistics. 2021 Census) <https://dzs.gov.hr/vijesti/objavljeni-konacni-rezultati-popisa-2021/1270>

English, T., & Carstensen, L. L. (2014). Selective narrowing of social networks across adulthood is associated with improved emotional experience in daily life. *International Journal of Behavioral Development*, *38*, 195–202. https://doi.org/10.1177/01650254135154

Erikson, & Erikson (1998). *Life cycle completed.* W. W. Norton & Company.

Feng, Q., Son, J., & Zeng, Y. (2014). Prevalence and correlates of successful ageing: A comparative study between China and South Korea, *European Journal of Ageing, 12*(2), 83–94. https://doi.org/10.1007/s10433-014-0329-5

Fernandez-Ballesteros, R. (2019). The concept of successful aging and related terms. In R. Fernandez-Ballesteros, A. Benetos, J-M. Robine (Eds.), *The Cambridge handbook of successful aging* (pp. 6–22). Cambridge University Press. https://doi.org/10.1017/9781316677018.002

Fernandez-Ballesteros, R., Molina, M., Schettini, R., & Santacreu, M. (2013). The semantic network of aging well. In J.-M. Robine, C. Jagger, E. M. Crimmins (Eds.), *Annual review of gerontology and geriatrics, Vol. 33. Healthy longevity: A global approach* (pp. 79–107). Springer Publishing Company.

Finch, C. E., & Tanzi, R. E. (1997). Genetics of aging. *Science*, *278*(5337), 407–411. https://doi.org/10.1126/science.278.5337.407

Folstein, M. F., Folstein, S. E., White, T., & Messer, M. A. (2011). *Kratko ispitivanje mentalnog statusa, Priručnik (Mini-Mental State Examination, Manual).* Naklada Slap.

Glatt, S. J., Chayavichitsilp, P., Depp, C., Schork, N. J., & Jeste, D. V. (2007). Successful aging: From phenotype to genotype. *Biological Psychiatry*, *62*(4), 282–293. https://doi.org/10.1016/j.biopsych.2006.09.015

Gureje, O., Oladeji, B. D., Abiona, T., & Chatterji, S. (2014). Profile and determinants of successful aging in the Ibadan Study of Ageing. *Journal of the American Geriatrics Society*, *62*(5), 836–842. https://doi.org/10.1111/jgs.12802

Hank, K. (2011). How "successful" do older Europeans age? Findings from SHARE. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 66B(2),* 230–236. https://doi.org/10.1093%2Fgeronb%2Fgbq089

Havighurst, R. (1961). Successful aging. *Gerontologist*, *1*(1), 8–13. https://doi.org/10.1093/geront/1.1.8

Heckhausen, J., & Schulz, R. (1993). Optimization by selection and compensation: Balancing primary and secondary control in life span development. *International Journal of Behavioral Development*, *16*, 287–303. https://doi.org/10.1177/016502549301600210

Kadariya, S., Gautam, R., & Aro, A. R. (2019). Physical activity, mental health, and wellbeing among older adults in south and southeast Asia: A scoping review. *BioMed Research International.* 2019, Article ID: 6752182. https://doi.org/10.1155/2019/6752182

Katz, S., Downs, T. D., Cash, H. R., & Grotz, R. C. (1970). Progress in development of the index of ADL. *Gerontologist*, *10*(1), 20–30. https://doi.org/10.1093/geront/10.1_part_1.20

Kleineidam, L., Thoma, M. V., Maercker, A., Bickel, H., Mösch, E., Hajek, A., König, H. H., Eisele, M., Mallon, T., Luck, T., Röhr, S., Weyerer, S., Werle, J., Pentzek, M., Fuchs, A., Wiese, B., Mamone, S., Scherer, M., Maier, W., Riedel-Heller, S. G., & Wagner, M. (2019). What is successful aging? A psychometric validation study of different construct definitions*. Gerontologist, 59*(4), 738-748. https://doi.org/10.1093/geront/gny083

Kokkinos, P., Myers, J., Faselis C., Panagiotakos, D. B., Doumas, M., Pittaras, A., & Fletcher, R. (2010). Exercise capacity and mortality in older men: A 20 year old follow–up study. *Circulation*, *122*, 790–797. https://doi.org/10.1161/CIRCULATIONAHA.110.938852

Larson, E. B., Wang, L., Bowen, J. D., McCormick, W. C., Teri, L., Crane P., & Kukull, W. (2006). Exercise is associated with reduced risk for incident dementia among persons 65 years of age and older. *Annals of Internal Medicine*, *144*(2), 7–81. https://doi.org/10.7326/0003-4819-144-2-200601170-00004

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist*, *9*(3), 179–186. https://doi.org/10.1093/geront/9.3_Part_1.179

Lepan, Ž., & Leutar, Z. (2012). Važnost tjelesne aktivnosti u starijoj životnoj dobi (The importance of physical activity for older adults). *Socijalna* ekologija, *21*(2), 203–223.

Liang, J., & Luo, B. (2012). Toward a discourse shift in social gerontology: From successful aging to harmonious aging. *Journal of Aging Studies*, *26*(3), 327–334. https://doi.org/10.1016/j.jaging.2012.03.001

Lustig, C., Shah, P., Seidler, R., & Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: A review and future directions. *Neuropsychology Review*, *19*(4), 504–522. https://doi.org/10.1007/s11065-009-9119-9

Maier, A., Riedel-Heller, S. G., Pabst, A., & Luppa, M. (2021). Risk factors and protective factors of depression in older people 65+. A systematic review. *PLoS One, 16*(5):e0251326. https://doi.org/10.1371/journal.pone.0251326

Maraldi, C., Harris, T. B., Newman, A. B., Kritchevsky, S. B., Pahor, M., Koster, A., Satterfield, S., Ayonayon, H. N., Fellin, R., & Volpato, S. (2009). Moderate alcohol intake and risk of functional decline: The Health, Aging, and Body Composition Study. *Journal of American Geriatrics Society*, *57*(10), 1767–1775. https://doi.org/10.1111/j.1532-5415.2009.02479.x

Martin, P., da Rosa, G., Siegler, I. C., Davey, A., MacDonald, M., Poon, L. W. for the Georgia Centenarian Study (2006). Personality and longevity: Findings from the Georgia Centenarian Study. *Age (Omaha) 28*, 343–352. https://doi.org/10.1007/s11357-006-9022-8

Martin, P., Kelly, N., Kahana, B., Kahana, E., Willcox, B. J., Willcox, D. C., & Poon, L. W. (2015). Defining successful aging: A tangible or elusive concept? *The Gerontologist*, *55*(1), 14–25. https://doi.org/10.1093/geront/gnu044

Martin-Joy, J., & Vaillant, G. E. (2010). Recognizing and promoting resilience. In C. A. Depp, D. V. Jeste (Eds.), *Succesfull emotional and cognitive aging* (pp. 363–382). American Psychiatric Publishing.

Masoro, E. J. (2001). "Successful aging": useful or misleading concept? *The Gerontologist*, *41*(3), 415-418. https://doi.org/10.1093/geront/41.3.415.

McCrae, R. R., & Costa, P. T. (2004). A contemplated revision of the NEO five-factor inventory. *Personality and Individual Differences*, *36*(3), 587–596. https://doi.org/10.1016/S0191-8869(03)00118-1

Meng, X., & D'Arcy, C. (2014). Successful aging in Canada: Prevalence and predictors from a population-based sample of older adults. *Gerontology*, *60*(1), 65–72. https://doi.org/10.1159/000354538

Montross, L. P., Depp, C., Daly, J., Reichstadt, J., Golshan, S., Moore, D., Sitzer, D., & Jeste, D. V. (2006). Correlates of self-rated successful aging among community dwelling older adults. *American Journal of Geriatric Psychiatry*, *14*(1), 43–51. https://doi.org/10.1097/01.JGP.0000192489.43179.31

Ng, T. P., Broekman, B. F. P., Niti, M., Gwee, X., & Kua, E. H. (2009). Determinants of successful aging using a multidimensional definition among Chinese elderly in Singapore. *The American Journal of Geriatric Psychiatry*, *17*(5), 407–416. https://doi.org/10.1097/JGP.0b013e31819a808e

Nosraty, L., Sarkeala, T., Hervonen, A., & Jylhä, M. (2012). Is there successful aging for nonagenarians? The vitality 90+ study. *Journal of Aging Research*, *2012*, Article ID: 868797. https://doi.org/10.1155/2012/86879

Parslow, R. A., Lewis, V. J., & Nay, R. (2011). Successful aging: development and testing of a multidimensional model using data from a large sample of older Australians. *Journal of the American Geriatrics Society*, *59*(11), 2077–2083. https://doi.org/10.1111/j.1532-5415.2011.03665.x

Pattie, A. H., & Gilleard, C. J. (1996). *Priručnik za primjenu Clifton postupaka za procjenu starijih osoba CAPE (Application Manual for the Clifton Assessment Procedures for the Elderly CAPE).* Naklada Slap.

Pocnet, C., Popp, J., & Jopp, D. (2020). The power of personality in successful ageing: A comprehensive review of larger quantitative studies. *European Journal of Ageing*, *18*, 269–285. https://doi.org/10.1007/s10433-020-00575-6

Pruchno, R. A., Wilson-Genderson, M., Rose, M., & Cartwright, F. (2010). Successful aging: Early influences and contemporary characteristics. *The Gerontologist*, *50*(6), 821–833. https://doi.org/10.1093/geront/gnq041

Reich, A. J., Claunch, K. D., Verdeja, M. A., Dungan, M. T., Anderson, S., Clayton, C. K., Goates, M. C., & Thacker, E. L. (2020). What does "successful aging" mean to you? – Systematic review and cross-cultural comparison of lay perspectives of older adults in 13 countries, 2010-2020. *Journal of Cross-Cultural Gerontology*, *35*(4), 455–478. https://doi.org/10.1007/s10823-020-09416-6.

Reichstadt, J., Depp, C. A., Palinkas, L. A., Folsom, D. P., & Jeste, D. V. (2007). Building blocks of successful aging: A focus group study of older adults' perceived contributors to successful aging. *American Journal of Geriatric Psychiatry*, *15*(3), 194–201. https://doi.org/10.1097/JGP.0b013e318030255f.

Rowe, J. W., & Kahn, R. L. (1987). Human aging: Usual and successful. *Science*, *237*, 143–149. https://doi.org/10.1126/science.3299702

Rowe, J. W., & Kahn, R. L. (1997). Successful aging. *The Gerontologist*, *37*, 433–440. https://doi.org/10.1093/geront/37.4.433

Rowe J. W., & Kahn, R. L. (1998). *Successful aging.* Pantheon Books.

Rubinstein, R. L., & de Medeiros, K. (2014). "Successful aging," gerontological theory and neoliberalism: A qualitative critique. *The Gerontologist*, *55*(1), 34–42. https://doi.org/10.1093/geront/gnu080

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A re-evaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, *67*, 1063–1078. https://doi.org/10.1037/0022-3514.67.6.1063

Sherer, M., Maddux, J. E., Mercandante, B., Prentice-dunn, S., Jacobs, B., & Rogers, R. W. (1982). The Self-Efficacy Scale: Construction and validation. *Psychological Reports, 51*(2), 663–671. https://doi.org/10.2466/pr0.1982.51.2.663

Sofi, F., Velecchi, D., Bacci, D., Abbate, R., Gensini, G.F., & Macchi, C. (2011). Physical activity and risk of cognitive decline: A meta-analyses of prospective studies. *Journal of Internal Medicine*, *269*(1), 107–117. https://doi.org/10.1111/j.1365-2796.2010.02281.x

Spirduso, W. W., & Cronin, D. L. (2001). Exercise dose-response effects on quality of life and independent living in older adults. *Medicine & Science in Sports & Exercise*, *33*(suppl 6), S598–S608. https://doi.org/10.1097/00005768-200106001-00028.

Stewart, J. M., Auais, M., Belanger, E., & Phillips, S. P. (2019). Comparison of self-rated and objective successful ageing in an international cohort. *Ageing and Society*, *39*(7), 1317–1334. https://doi.org/10.1017/S0144686X17001489

Strawbridge, W. J., Wallhagen, M. I., & Cohen, R. D. (2002). Successful aging and well-being: Self–rated compared with Rowe and Kahn. *The Gerontologist*, *42*(6), 727–733. https://doi.org/10.1093/geront/42.6.727

Štambuk, A., & Tomičić, V. (2020). Iskustva starijih osoba s plesom kao oblikom fizičke aktivnosti (Experiences of older people with dancing as a form of physical activity). *Croatian Journal of Education, 22*(4), 1255-1281. https://doi.org/10.15516/cje.v22i4.3805

Tucak Junaković, I. (2022). Teorijski pristupi i metode istraživanja uspješnog starenja: Što je uspješno starenje i kako ga mjeriti? (Theoretical approaches and methods of research on successful aging: What is successful aging and how to measure it?) In I. Tucak Junaković (Eds.), *Uspješno starenje: teorijski pristupi, metode istraživanja i čimbenici doprinosa (Successful aging: theoretical approaches, research methods and contributing factors)* (pp. 7-36). University of Zadar.

Tucak Junaković, I. (2024). Odrednice uspješnog starenja i dugovječnosti: kako dobro (o)starjeti (Determinants of successful aging and longevity: how to age well and achieve a good old age). In I. Macuka, I. Burić, A. Slišković, A. Šimunić, P. Valerjev (Eds.), *24th Psychology Days in Zadar-Book of Abstracts* (p. 18). Department of Psychology, University of Zadar.

Tucak Junaković, I., & Ambrosi-Randić, N. (2022). Lay Definitions of Successful Aging and Contributing Factors among Croatian Older Adults: A Thematic Analysis of Qualitative Data. *Psihologijske teme,*

31(3), 685-708. https://doi.org/10.31820/pt.31.3.11

Tucak Junaković, I., Ambrosi-Randić, N., & Macuka, I. (2023). Psihosocijalne determinante objektivnih i subjektivnih pokazatelja uspješnoga starenja (Psychosocial determinants of objective and subjective indicators of successful ageing). *Društvena istraživanja, 32*(3), 449-469. https://doi.org/10.5559/di.32.3.04

Tucak Junaković, I., Ambrosi-Randić, N., & Nekić, M. (2018). Odnos otvorenosti prema iskustvu i uspješnog starenja: posredujuća uloga životnih žaljenja (The Relationship between Openness to Experience and Successful Aging: Testing the Mediation Role of the Life Regrets). *Psihologijske teme*, *27*(3), 499–518. https://doi.org/10.31820/pt.27.3.8

Tucak Junaković, I., Martinčević, M., & Vranić, A. (2024). Validation of an integrative multidimensional model of successful aging in community-dwelling older adults: what is successful aging all about? *Current Psychology*, *43,* 24977–24993 (2024). https://doi.org/10.1007/s12144-024-06165-4

Tucak Junaković, I., Nekić, M., & Ambrosi-Randić, N. (2017). Some psychosocial predictors of successful aging among Croatian older adults. In I. Burić (Eds.), *Book of selected proceedings of the 20th Psychology Days in Zadar* (pp. 275–286). University of Zadar.

Tucak Junaković, I., Nekić, M., & Šimić, N. (2023). Lifestyle factors, psychosocial resources, and mental health in older people. *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry.* Advance online publication. https://doi.org/10.1024/1662-9647/a000330

Urtamo, A., Jyväkorpi, S. K., & Strandberg, T. E. (2019). Definitions of successful ageing: A brief review of a multidimensional concept. *Acta BioMedica*, *90*(2), 359-363. https://doi.org/10.23750/abm.v90i2.8376

Vahia, I. V., Thompson, W. K., Depp, C. A., Allison, M., & Jeste, D. V. (2012). Developing a dimensional model for successful cognitive and emotional aging. *International Psychogeriatrics*, *24*(4), 515-523. https://doi.org/10.1017/S1041610211002055.

Vranić, A. (2022). Uspješno kognitivno starenje: kako očuvati zdravlje mozga? (Successful cognitive aging: how to maintain brain health?). In I. Tucak Junaković (Eds.), *Uspješno starenje: teorijski pristupi, metode istraživanja i čimbenici doprinosa (Successful aging: theoretical approaches, research methods and contributing factors)* (pp. 73-100). University of Zadar.

Vuori, I. (2005). Tjelesna aktivnost kao učinkovito sredstvo protiv nepovoljnog zdravstvenog djelovanja tjelesne neaktivnosti (Physical activity as an effective means against the adverse health effects of physical inactivity). *Glasnik Hrvatskog saveza sportske rekreacije*, 3–12.

Wang, H. X., Karp, A., Winblad, B., & Fratiglioni, L. (2002). Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: A longitudinal study from the Kungsholmen project. *American Journal of Epidemiology*, *155*(12), 1081–1087. https://doi.org/10.1093/aje/155.12.1081

Wilson, R. S., Krueger, K. R., Arnold, S. E., Schneider, J. A., Kelly, J. F., Barnes, L. L., Tang, Y., & Bennett, D. A. (2007). Loneliness and risk of Alzheimer disease. *Archives of General Psychiatry*, *64*(2), 234–240. https://doi.org/10.1001/archpsyc.64.2.234

Young, Y., Frick, K. D., & Phelan, E. A. (2009). Can successful aging and chronic illness coexist in the same individual? A multidimensional concept of successful aging. *American Medical Directors Association*, *10*, 87–92. https://doi.org/10.1016/j/jamda.2008.11.003

4

Meaning in Life: Is Humour the Answer?

Paula Pedić Duić
Faculty of Teacher Education, University of Zagreb
paula.pedic-duic@ufzg.hr

Tea Pavin Ivanec
Faculty of Teacher Education, University of Zagreb

Zvjezdan Penezić
Department of Psychology, University of Zadar

Abstract

This study aimed to examine the mediating role of neuroticism in the relationship between materialism and meaning in life. Additionally, the study explored the moderating role of humour, conceptualised as a dispositional positive emotion, in explaining the connection between neuroticism and meaning in life and the proposed mediation mechanism. The research was conducted on a convenience sample of 322 adults, of whom 86% were women, with an average age of 32 years. Materialism (happiness dimension) was measured using the Short Form of the Materialism Scale (Richins, 2004), neuroticism was assessed with the BFI questionnaire (neuroticism subscale; Benet-Martinez & John, 1998), meaning in life was measured with the Meaning in Life Scale (Vulić-Prtorić & Bubalo, 2006), and humour was operationalised as a dispositional tendency to experience humour, as measured by the Humour subscale of the Dispositional Positive Emotions Scale (Shiota et al., 2006). The results revealed that neuroticism mediates the relationship between materialism and life meaning. Individuals inclined towards materialism showed higher levels of neuroticism, which was associated with a lower meaning in life. Humour moderated the relationship between neuroticism and meaning in life and the proposed mediation mechanism. The negative association between neuroticism and meaning in life was weaker among those more inclined to experience humour daily. Similarly, the negative relationship between materialism and meaning in life through neuroticism was less pronounced among those with a greater tendency to experience humour in everyday life.

Keywords: *humour, materialism, neuroticism, life meaning, happiness in possession*

Introduction

Reker and Wong (1988) conceptualise meaning in life as the comprehension of order, harmony, and purpose in one's existence. Their model identifies three components of meaning in life: affective, cognitive, and motivational. The affective component relates to the sense of satisfaction and fulfilment derived from finding personal meaning; the cognitive component involves interpreting life events as meaningful; and the motivational component reflects behaviours directed towards activities that provide life with meaning. One possible negative correlate of meaning in life is materialism. Although studies that link these two constructs are relatively limited, existing literature suggests a negative relationship between materialism and well-being (Kasser & Ryan, 1996; Dittmar et al., 2014; Roberts & Clement, 2007). Given the conceptual overlap between well-being and meaning in life—where meaning is often considered a crucial determinant of well-being (Steger, 2009)—it is reasonable to expect that materialism would also negatively impact meaning in life. Self-Determination Theory (SDT; Ryan & Deci, 2000) offers a valuable framework for understanding the detrimental effects of materialism on well-being and meaning in life. According to SDT, individuals have three fundamental psychological needs: autonomy (a sense of control over one's actions), competence (the feeling of effectiveness and capability), and relatedness (a sense of connection with others). When these needs are fulfilled, individuals experience greater well-being and a stronger sense of meaning in life (Marwa & Xiaosong, 2023). However, materialistic values prioritise extrinsic goals such as wealth, social status, and image, which do not inherently satisfy these psychological needs (Kasser & Ryan, 1996). In other words, materialism may diminish one's sense of meaning by diverting attention from pursuits that authentically contribute to well-being and fulfilment. Accordingly, and building on these theoretical and empirical insights, this study proposes that materialism is negatively associated with meaning in life.

Materialism has most commonly been conceptualised in three ways (Lind et al., 2014). Belk (1985) defined materialism as a personality trait, characterised by possessiveness, envy, and nongenerosity. Drawing on Self-Determination Theory (Ryan & Deci, 2000), Kaser (2002) described materialism as part of a broader value system that prioritises extrinsic goals (e.g., financial success, fame, social recognition) at the expense of intrinsic goals (e.g., self-acceptance, meaningful relationships). Finally, Richins (2004) and Richins and Dawson (1992) conceptualised materialism as a set of personal values, focusing on three distinct dimensions: (1) success, the belief that material wealth is a key measure of life achievement; (2) centrality, the extent to which possessions occupy a central role in one's life and identity; and (3) happiness, the belief that material goods are essential for personal well-being and life satisfaction. These dimensions reflect different ways in which individuals derive meaning from material possessions. While this framework has been widely used, research on the distinct contributions of these dimensions to well-being remains relatively limited. Among the existing studies, Ahuvia and Wong (2002) found that the happiness dimension of materialism is the strongest negative predictor of life satisfaction among the three dimensions. The success dimension also negatively correlates with life satisfaction, but only in specific domains (e.g., social relationships), while the centrality dimension does not significantly impact life satisfaction. These findings suggest that the happiness dimension of materialism is the most substantial negative predictor of well-being among all three dimensions.

Recent studies emphasise the complex role of personality traits in shaping the relationship between materialism and well-being. Specifically, materialists with high neuroticism and low grandiose narcissism experience diminished well-being, whereas those with low neuroticism and high grandiose narcissism report higher well-being (Górnik-Durose, 2020). This suggests that neuroticism and grandiose narcissism act as opposing mediators—neuroticism exacerbates the negative effects of materialism, while grandiose narcissism mitigates them. In contrast, Górnik-Durose & Pyszkowska (2020) reported that neuroticism alone

does not mediate this relationship, while narcissism exerts a suppressing effect. The authors propose that the absence of neuroticism as a mediator in their study can be explained by differences in measurement and sample composition (Górnik-Durose & Pyszkowska, 2020). They also suggest that their use of the emotionality scale from HEXACO-PI-R, instead of traditional neuroticism measures (e.g., EPQ, NEO-FFI), may have influenced the findings, as HEXACO emotionality, while related to neuroticism, captures a broader range of traits (Ashton et al., 2014). Additionally, they emphasise that their study focused solely on young adults, a group in which materialism may be more influenced by societal norms than by early-life insecurity (Kasser et al., 2004). Given that personality traits in this developmental stage are still undergoing change (Roberts et al., 2006), they propose that neuroticism may play a less pronounced regulatory role in comparison to older populations.

The relationship between meaning in life and neuroticism has been studied across various samples, including students (Halama, 2005), middle-aged adults (Halama, 2005), and individuals with psychiatric diagnoses, including alcohol addiction (Pearson & Sheffield, 1974; Bupić & Dijaković, 2019). Individuals with high levels of neuroticism are generally more emotionally reactive, irritable, anxious, and insecure compared to those with lower levels of neuroticism (Tackett & Lahey, 2017). These neurotic traits appear to particularly hinder the achievement of meaning in life (Halama, 2005). However, this pattern varies with age. Halama (2005) suggests that in younger individuals, neuroticism is more strongly associated with the cognitive and motivational components of meaning in life, while in middle-aged individuals, it is more closely linked to the affective components. For younger people, neuroticism may impede the formation of an optimistic worldview and achieving goals, leading to a lower sense of meaning in life. In contrast, for older individuals, neuroticism affects meaning in life by disrupting experiences of satisfaction and happiness, ultimately resulting in reduced meaning in life.

Given that materialism is a positive predictor of neuroticism (Górnik-Durose & Boroń, 2018; Ardabili et al., 2022), and neuroticism is a negative predictor of meaning in life (Halama, 2005; Pearson & Sheffield, 1974; Bupić & Dijaković, 2019), this study hypothesises that neuroticism will negatively mediate the relationship between materialism and meaning in life. However, due to the lack of studies on specific associations of the subcomponents of materialism with neuroticism and meaning in life—and given that previous research has demonstrated the happiness dimension of materialism to be the strongest negative component of well-being (Ahuvia & Wong, 2002)—this study will specifically focus on the happiness dimension of materialism.

Besides examining the aforementioned mediation model, this study also explores humour, as a potential moderator of this mediation effect and the direct relationship between neuroticism and meaning in life. By investigating this dual moderating role, the study seeks to clarify the conditions under which the relationship between materialism, neuroticism, and meaning in life varies according to individuals' levels of humour. Humour has been conceptualised in multiple ways, including as a temperament, a coping strategy, an ability, an attitude, and character strength (Ruch & McGhee, 2014). While there is no single universally accepted definition, research suggests that humour is a relatively stable trait (Ruch, 1998) that induces amusement (Auerbach et al., 2016), which is a key component of positive emotional experiences (Platt et al., 2013). Research consistently shows that humour plays a critical role in enhancing physical and psychological well-being, primarily through its moderating effect on stress (Martin et al., 2003). Studies indicate that individuals with a stronger sense of humour are more resilient to psychological distress in stressful situations and less vulnerable to mood disorders (Martin & Lefcourt, 1983; Kuiper, 2012). Similarly, humour buffers the relationship between negative life events and depressive symptoms, reinforcing the concept of a "stress-buffering effect" (Nezu et al., 1988). One recent study specifically examined the moderating role of humour on the relationship between neuroticism and life satisfaction (Pinchot, 2020). Individuals with high levels of neuroticism

who frequently employed self-enhancing humour reported greater life satisfaction than those with high neuroticism who rarely used self-enhancing humour. For highly neurotic individuals, who are typically more prone to experiencing negative emotions, humour might help them reframe adverse events in a less threatening way, thereby reducing the emotional intensity of these experiences. This cognitive reappraisal process can help diminish the overall impact of stress and promote a more positive outlook on life, ultimately leading to higher life satisfaction.

A similar outcome can also be expected for meaning in life. For individuals with high neuroticism, who may struggle to find coherence and purpose due to their tendency towards negative thinking and emotional distress, humour can offer a different perspective (Martin, 2007; Kuiper et al., 1993). By enabling them to see humour in challenging situations, they may reinterpret these experiences in a way that aligns with a broader life purpose or values (Folkman & Moskowitz, 2000), so the third hypothesis of this study is that humour moderates the relationship between neuroticism and meaning in life. It is expected that the negative effect of neuroticism on meaning in life is weaker for individuals who are more prone to experiencing humour in daily life.

Humour might also moderate the negative mediation effect of neuroticism on the relationship between materialism and meaning in life. Specifically, a higher disposition towards humour might mitigate the adverse emotional effects typically associated with materialistic values and neurotic tendencies, thereby sustaining or potentially enhancing one's perceived meaning in life. Seeking happiness through wealth and possessions often undermines a sense of meaning because it focuses on external rewards that do not provide deep or lasting purpose (Kasser, 2002). When people who are high in neuroticism—characterised by frequent anxiety, worry, and emotional instability—also hold materialistic values, they may struggle even more to find meaning, as their negative emotional tendencies make it harder to feel satisfied with their material achievements (Diener et al., 2003). Hence, humour can serve as a coping mechanism by helping these individuals handle negative emotions and reframe their experiences in a more positive light (Simione & Gnagnarella, 2023). By fostering a lighter perspective and reducing the emphasis on material success, humour may enable them to find meaning in their everyday lives, relationships, and internal values rather than solely from external possessions (Martin & Ford, 2018; Peterson & Seligman, 2004). Thus, the fourth hypothesis of this study is that humour weakens the negative effect of neuroticism on the relationship between materialism and meaning in life.

Method

Participants and Procedure

The study included 322 adults, predominantly female (86%). This sample size was deemed sufficient for analysing the index of moderated mediation, as Monte Carlo simulations by Ng et al. (2024) suggest that a sample of N ≥ 300 is required to ensure the robustness of moderated mediation effects. The average age of the participants was 32.38 (*SD* = 11.67), with an age range from 18 to 83 years. Most participants (70%) had at least a bachelor's degree, with a range of educational backgrounds from high school completion (30%) to PhD (4%). Most participants (68%) reported an average SES, 23.9% reported above-average, and 8.1% reported below-average SES. Data was collected via an online questionnaire using Google Forms. The participants were informed about the purpose of the study, the voluntary nature of participation, and data anonymity. The invitation to participate was disseminated through social media platforms by the researchers. The only inclusion criterion was that the participants were required to be adults (18+ years old). The study was conducted from February to March 2022.

Instruments

Meaning in Life

Meaning in life was measured with the Meaning in Life Scale (Vulić-Prtorić & Bubalo, 2006), a Croatian version of the Purpose in Life Test (PIL; Crumbaugh & Maholick, 1969). The scale consists of 23 items that assess both the affective (e.g., "*My life is full of exciting events")* and cognitive (e.g., "*I haven't found any purpose in life"*) aspects of the meaning of life, specifically examining the quality and purpose of existence. The participants rate their agreement with each item on a 5-point Likert scale (1 = *does not apply to me at all*; 5 = *applies to me completely*). Cronbach's alpha for this and other measures used in this study is displayed in Table 1.

Materialism

Materialism was assessed using the nine-item short form of the Material Values Scale (MVS; Richins, 2004). This scale comprises three dimensions: Success, Centrality, and Happiness, each represented by three items. In the present study, only the Happiness subscale was used to evaluate the degree to which individuals perceive that acquiring material possessions enhances their happiness (e.g., "*I'd be happier if I could afford to buy more things*"). The participants responded to items on a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Neuroticism

Neuroticism was assessed using the Neuroticism subscale of the Big Five Inventory (BFI; Benet-Martinez & John, 1998). This subscale consists of eight items measuring tendencies towards anxiety, depression, mood variability, and emotional sensitivity (e.g., *"I see myself as someone who is very worried"*), with responses rated on a 5-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree).*

Humour

The dispositional tendency towards experiencing humour was measured using the Humour subscale of the Dispositional Positive Emotion Scale (DPES; Shiota et al., 2006). The DPES consists of seven subscales, each measuring a different positive emotion: joy, humour, contentment, pride, love, compassion, and awe. The Humour subscale includes five items, such as *"I find humour in almost everything",* and participants rate their agreement with each statement on a 7-point Likert scale (1 = *strongly disagree* to 7 = *strongly agree).*

Analysis Plan

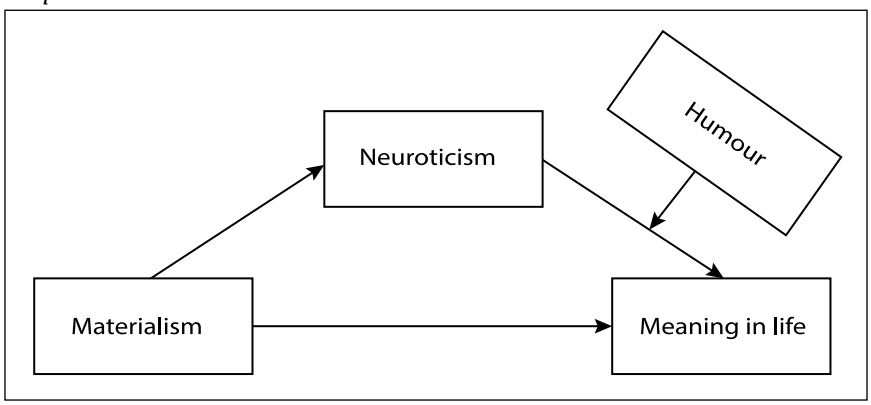
Consistent with previous research that tested mediation and b-path moderated mediation hypotheses using the PROCESS macro in SPSS (Kim & Chen, 2015; Yang et al., 2021; Afrin et al., 2022), this study employed both Model 4 and Model 14 of the PROCESS macro (Hayes, 2017). Specifically, Model 4 was used to test the mediation hypothesis, where neuroticism was hypothesised to mediate the relationship between materialism and meaning in life (as shown in Figure 1). The PROCESS macro generated 5000 bootstrapped estimates for the indirect effect with 95% confidence intervals (CIs), which was considered significant if the CIs did not include zero.

Model 14 was subsequently used to examine the proposed moderated mediation. This analysis assessed the conditional indirect effect of the moderator (humour) on the relationship between neuroticism and meaning in life, as well as on the indirect effect of neuroticism on the relationship between materialism and meaning in life, as illustrated in Figure 1. The significance of the conditional indirect effects in the moderated mediation model was also determined using 5000 bootstrapped estimates provided by the PROCESS macro.

The descriptive statistics and correlation analyses were conducted using IBM SPSS version 25.0.

Figure 1

Proposed mediation and moderated mediation model



Results

Table 1

Descriptive statistics and Pearson correlation coefficients among the study variables (N=322)

	M	*SD*	Cronbach's α	1.	2.	3.	4.
1. Materialism	3.0	1.1	.82	-	-.28**	.28**	.01
2. Meaning in life	83.9	16.0	.82		-	-.58**	.22**
3. Neuroticism	2.8	0.9	.88			-	-.21**
4. Humour	4.8	1.3	.83				-

**$p < 0.01$

As seen in Table 1, materialism and meaning in life are weakly negatively correlated, while materialism and neuroticism are weakly positively correlated. Materialism is not significantly associated with humour. Meaning in life and neuroticism are strongly negatively correlated, whereas meaning in life and humour are weakly positively correlated. Finally, neuroticism and humour are weakly negatively correlated. Less neurotic individuals tend to perceive life as more meaningful, associate happiness less with material possessions, and experience more humour in everyday life. The mean scores for the Meaning in Life Scale, Neuroticism subscale of the Big Five Inventory, and Humour subscale of the Dispositional Positive Emotion Scale are consistent with previous research (e.g., Vulić-Prtorić & Bubalo, 2006; Benet-Martinez & John, 1998; Chirico et al., 2021), while the mean score for the Material Values Scale—Happiness subscale is lower than reported in earlier studies (Trzcińska & Sekścińska, 2022). Reliability, as measured by Cronbach's alpha, was adequate for all scales used in the study.

Table 2

Mediation analysis results showing the effects of materialism on meaning in life through neuroticism (N=322)

Predictors		M Neuroticism				Y Meaning in life		
		B	*SE*	95% *CI*		*B*	*SE*	95% *CI*
X (Materialism, happiness dimension)	a_1	0.22	0.04	[0.14, 0.31]	*c'*	-1.77	0.70	[-3.15, -0.40]
M (Neuroticism)	-	-	-	-	*b*	-10.48	0.87	[-12.18, -8.78]
Constant	i_{M1}	2.10	0.14	[3.79, 4.22]	*iy*	119.83	2.83	[114.27, 125.38]
		R^2 = 0.08				R^2 = 0.37		
		$F(1, 321)$ = 27.07				$F(2,320)$ = 92.36		
		p <0.01				p <0.01		

Note. B – Unstandardised regression coefficients; X – Predictor; M – Mediator; Y – Outcome.

The indirect effect of materialism on meaning in life through neuroticism was negative (*ab* = -2.36, *BootSE* = 0.59, 95% *BootCI* [-3.58, -1.27]). As seen in Table 2, materialism (happiness dimension) is positively associated with neuroticism, which in turn negatively affects meaning in life (*B* = -10.48, *SE* = 0.87, *p* < 0.01). The direct effect of materialism on meaning in life is also negative *(B* = -1.7, *SE* = 0.70, *p* < 0.01). This suggests that materialistic values diminish meaning in life both directly and indirectly by increasing neuroticism, with neuroticism accounting for a substantial part of this effect, explaining 37% of the variance in meaning in life when neuroticism is considered, compared to a much lower variance explained by materialism alone (Table 2).

Moderated Mediation Analysis

Figure 2 illustrates the results of the moderated mediation model, which includes the predictor (materialism), the mediator (neuroticism), the moderator (humour), and the interaction between neuroticism and humour in predicting meaning in life. The findings indicate that materialism (happiness dimension) is positively associated with neuroticism (*B* = 0.22, *SE* = 0.04, *p* < 0.01), whereas both materialism (*B* = -1.66, *SE* = 0.69, *p* < 0.01) and neuroticism (*B* = -18.28, *SE* = 1.75, *p* < .05) are negatively related to meaning in life. Individuals who tend to experience more humour in everyday life reported higher levels of meaning in life (*B* = -3.82, *SE* = 1.67, *p* < 0.01). The statistically significant interaction (*B* = 1.75, *SE* = 0.55, *p* < 0.01) suggests that the effect of neuroticism on meaning in life is moderated by humour, indicating that humour influences the strength of this effect. The index of moderated mediation is also statistically significant (*ab* = 0.39, *BootSE* = 0.18, 95% *BootCI* [0.09, 0.77]), showing that humour moderates the impact of materialism on meaning in life through neuroticism. A simple slopes analysis was performed to further explore this interaction and the moderated mediation. This analysis examined the conditional effect of neuroticism on meaning in life at different levels of humour, as well as the conditional effect of neuroticism on the relationship between materialism and meaning in life at different levels of humour, generating confidence intervals to determine at which levels of the moderator (humour) the conditional effect is significant. The results of this analysis are presented in Table 3.

Figure 2

Results of the moderated mediation model

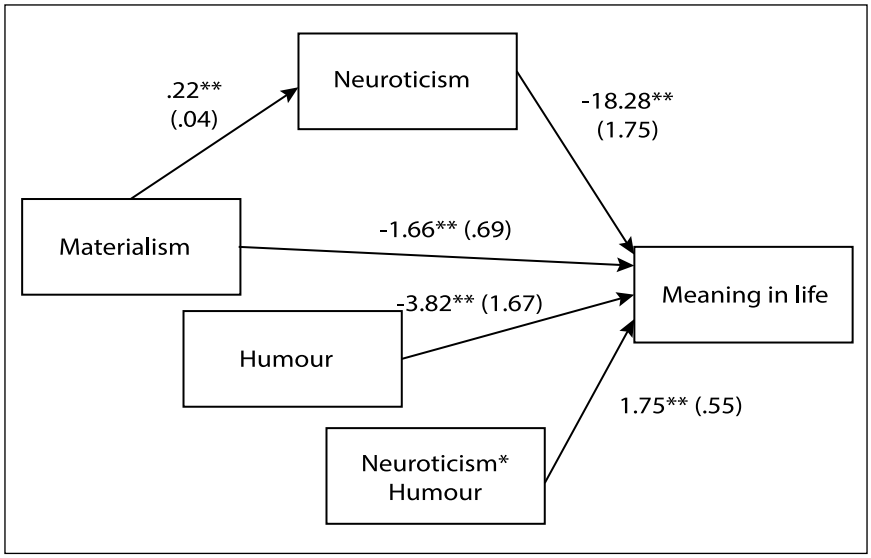


Table 3

Conditional effects of neuroticism on meaning in life and the mediating effect of neuroticism on the relationship between materialism and meaning in life at different levels of humour (N=322)

Humour levels	Effect of neuroticism on meaning in life (*B*)	*SE*	95% *CI*	Mediating effect of neuroticism on the relationship between materialism and meaning in life (*B*)	*BootSE*	95% *BootCI*
-1 *SD* (3.44)	-12.28	1.11	[-14.47, -10.09]	-2.74	0.67	[-4.11, -1.45]
M (4.77)	-9.96	0.87	[-11.67, -8.24]	-2.22	0.55	[-3.31, -1.21]
+1 *SD* (6.09)	-7.63	1.16	[-9.91, -5.36]	-1.70	0.49	[-2.74, -0.83]
Overall model	$R = 0.63$			$R^2 = 0.39$		$F(4, 317) = 51.45$; $p < 0.01$

As shown in Table 3, the negative effect of neuroticism on meaning in life weakens as levels of humour increase. Also, the negative indirect effect of materialism on meaning in life through neuroticism decreases as levels of humour increase. The overall moderated mediation model explains 39% of the variance in meaning in life, indicating a substantial proportion of variance accounted for by the predictor, mediator, and moderator variables together.

Discussion

This study aimed to examine whether neuroticism mediates the relationship between materialism and meaning in life and whether individuals' levels of humour influence this mediation. The first hypothesis proposed that materialism, precisely its happiness dimension, would have a negative effect on meaning in life. The results confirmed this hypothesis: materialism (happiness dimension) was significantly negatively associated with meaning in life. Individuals who placed a higher value on material possessions as a source of happiness reported a lower sense of meaning in their lives. This result is particularly noteworthy considering recent study by Lo et al. (2024), which is one of the first studies to specifically investigate the

relationship between materialism and meaning in life and the distinct effects of different materialism subcomponents. Lo et al. (2024) found that among the three dimensions of materialism, only materialistic happiness was negatively associated with meaning in life. Although the current study did not directly compare the three dimensions of materialism, our findings corroborate Lo et al.'s (2024) conclusion that materialistic happiness is particularly detrimental to one's sense of life meaning. Lo et al. (2024) explain the negative relationship between materialism and meaning in life through the framework of Self-Determination Theory (SDT; Deci & Ryan, 2000). According to SDT, fulfilling three basic psychological needs—autonomy, competence, and relatedness—is essential for well-being and the development of meaning in life. However, materialism prioritises extrinsic goals, such as wealth and status, over intrinsic aspirations, and such prioritisation fails to satisfy these fundamental needs (Kasser & Ryan, 1996). Lo et al. (2024) proposed a serial double mediation model, demonstrating that materialistic happiness leads to basic psychological need frustration, which in turn reduces subjective well-being, ultimately diminishing meaning in life. Their findings indicate that individuals who identify happiness with material possessions experience chronic dissatisfaction and weaker life purpose due to their unmet psychological needs. In addition to this mechanism, the present study proposes neuroticism as another potential mediator in the relationship between materialism and meaning in life. Specifically, it was proposed that neuroticism would act as a mediator in this relationship, such that higher levels of materialistic happiness would be associated with increased neuroticism, which, in turn, would negatively impact meaning in life.

The results confirmed this proposed mediation. More materialistic people tend to experience more anxiety, emotional instability, and negative emotions—traits associated with neuroticism (Suls & Martin, 2005; Górnik-Durose & Pilch, 2016). These negative emotional states can reduce a person's sense of meaning and purpose in life (Addad, 1987). Thus, neuroticism helps to explain why materialistic individuals often feel less fulfilled or purposeful: their focus on material possessions may increase their emotional distress, leading to a diminished sense of life meaning.

However, humour can act as a buffer against the negative effects of neuroticism. Specifically, the third hypothesis of this study, which suggested that humour could weaken the impact of neuroticism on meaning in life, and the fourth hypothesis, which proposed that humour could weaken the mediation effect of neuroticism on the relationship between materialism and meaning in life, were both supported by results. This emphasises the role of humour in helping individuals cope with stress and negative emotions, thereby reducing the harmful effects of neuroticism on meaning in life for those who frequently experience humour (Martin et al., 2003).

Furthermore, humour also moderated the mediation effect, indicating that the negative pathway from materialism to reduced life meaning through neuroticism is less pronounced among those with a greater tendency to experience humour. Thus, humour has a dual role in moderating both the direct effect of neuroticism on meaning in life and the indirect pathway from materialism to reduced life meaning through neuroticism. This dual moderating role of humour shows that individuals who are more inclined to experience humour may be better protected against the detrimental emotional consequences of materialistic values, allowing them to maintain a higher sense of meaning in life despite the presence of neurotic traits or materialistic tendencies (Martin & Ford, 2018; Peterson & Seligman, 2004).

The dual moderating role of humour aligns with the Broaden-and-Build Theory (Fredrickson, 1998), which suggests that positive emotions broaden cognitive and attentional resources, enabling individuals to develop more adaptive coping strategies over time. Unlike negative emotions, which narrow thought processes and focus on stressors, positive emotions—such as those elicited by humour—expand cognitive flexibility, allowing individuals to reframe negative experiences and see alternative perspectives. In the context of neuroticism, this broader perspective can help interrupt rigid patterns of negative thinking and emotional distress, reducing excessive self-focus and worry. Regarding materialism, humour may facilitate

reinterpretation of material possessions and values, preventing them from undermining life meaning. By fostering cognitive reappraisal and emotional regulation, humour thus plays a crucial role in protecting individuals from the harmful psychological effects of both neuroticism and materialism, helping them maintain a stronger sense of meaning in life.

Some limitations of this study should, however, be discussed. The sample was predominantly composed of highly educated women, limiting the generalisability of the findings across different sexes and educational backgrounds. This issue regarding the gender imbalance is a frequently observed social phenomenon in online surveys (Becker, 2022). Further, while the study found significant relationships between materialism, neuroticism, humour, and meaning in life, it cannot be concluded that these relationships are causal. Namely, establishing causality requires specific conditions, such as demonstrating that changes in one variable cause changes in another over time (Hayes, 2017), which cannot be achieved with the current study's cross-sectional design.

Despite these limitations, this study offers valuable insights by using a moderated mediation model to explore how humour may buffer the negative effects of neuroticism and materialism on meaning in life.

Conclusion

This study provides additional insights into how materialism, particularly its happiness dimension, is related to a lower sense of meaning in life. It confirms that neuroticism mediates this relationship, with materialistic values linked to increased neuroticism and, consequently, reduced meaning in life. However, humour appears to mitigate these negative effects, weakening both the direct effect of neuroticism on meaning in life and the indirect effect of neuroticism on the relationship between materialism and meaning in life. While the study has limitations, including a predominantly female, highly educated sample and a cross-sectional design that limits causal conclusions, it highlights the potential protective role of humour against the adverse emotional impacts of materialism and neuroticism on meaning in life. Future research could advance these findings by examining more diverse populations and using longitudinal study designs.

References

Addad, M. (1987). Neuroticism, extraversion and meaning of life: A comparative study of criminals and non-criminals. *Personality and Individual Differences, 8*(6), 879–883. https://doi.org/10.1016/0191-8869(87)90139-5

Afrin, R., Prybutok, G., & Prybutok, V. R. (2022). The pathways from online health information seeking to cyberchondria: A perspective from moderated mediation. *Psychiatry Research Communications, 2*(3), 100069. https://doi.org/10.1016/j.psycom.2022.100069

Ahuvia, A., & Wong, N. Y. (2002). Personality and values-based materialism: Their relationship and origins. *Journal of Consumer Psychology,12*(4), 389–402.

Ardabili, T., Chamorro-Premuzic, T., Grover, S., & Furn ham, A. (2022). Neuroticism, Materialism, Pay and Job Satisfaction. *Psychology, 13*, 1697-1710. https://doi.org/10.4236/psych.2022.1312106

Ashton, M. C., Lee, K., & de Vries, R. E. (2014). The HEXACO Honesty-Humility, agreeableness, and emotionality factors: A review of research and theory. *Personality and Social Psychology Review, 18*(2), 139–152. https://doi.org/10.1177/1088868314523838

Auerbach, S., Ruch, W., & Fehling, A. (2016). Positive emotions elicited by clowns and nurses: an experimental study in a hospital setting. *Translational Issues in Psychological Science 2,* 14–24. https://doi.org/10.1037/tps0000055

Becker, R. (2022). Gender and survey participation. An event history analysis of the gender effects of survey participation in a probability-based multi-wave panel study with a sequential mixed-mode design. *Methods, Data, Analyses, 6*(1), 3-22. https://doi.org/10.12758/mda.2021.08

Belk, R. W. (1985). Materialism: Trait aspects of living in the material world. *Journal of Consumer Research, 12*(3), 265–280. https://doi.org/10.1086/208515

Benet-Martínez, V., & John, O. P. (1998). *Los Cinco Grandes* across cultures and ethnic groups: Multitrait multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology, 75*, 729–750.

Bupić, J. & Bogović Dijaković, A. (2019). Meaning in Life and Personality Traits in Adult Alcoholics. *Archives of Psychiatry Research, 55*(1), 39-48. https://doi.org/10.20471/may.2019.55.01.03

Chirico, A., Shiota, M. N., & Gaggioli, A. (2021). Positive emotion dispositions and emotion regulation in the Italian population. *PloS one, 16*(3), e0245545. https://doi.org/10.1371/journal.pone.0245545

Crumbaugh, J. C., & Maholick, L. T. (1969). Manual of Instructions for the Purpose in Life Test. *Psychometric Publishers*.

Diener, E., Oishi, S., & Lucas, R. E. (2003). Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life. *Annual Review of Psychology, 54*, 403-425.

Dittmar, H., Bond, R., Hurst, M., & Kasser, T. (2014). The relationship between materialism and personal well-being: A meta-analysis. *Journal of Personality and Social Psychology, 107*(5), 879-924.

Folkman, S., & Moskowitz, J. T. (2000). Positive affect and the other side of coping. *American Psychologist, 55*(6), 647-654.

Fredrickson, B. L. (1998). What good are positive emotions? *Review of General Psychology, 2*(3), 300–319. https://doi.org/10.1037/1089-2680.2.3.300

Górnik-Durose, M. E. (2020). Materialism and well-being revisited: The impact of personality. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being, 21*(1), 305–326. https://doi.org/10.1007/s10902-019-00089-8

Górnik-Durose, M. E., & Boroń, K. (2018). Not Materialistic, Just Neurotic. The Mediating Effect of Neuroticism on the Relationship between Attitudes to Material Assets and Well-Being. *Personality and Individual Differences*, 123, 27-33. https://doi.org/10.1016/j.paid.2017.10.040

Górnik-Durose, M. E., & Pilch, I. (2016). The dual nature of materialism. How personality shapes materialistic value orientation. *Journal of Economic Psychology, 57,* 102–116. https://doi.org/10.1016/j.joep.2016.09.008

Górnik-Durose, M. E., & Pyszkowska, A. (2020). Personality matters—Explaining the link between materialism and well-being in young adults. *Personality and Individual Differences, 163*, Article 110075. https://doi.org/10.1016/j.paid.2020.110075

Halama, P. (2005). Relationship between meaning in life and the big five personality traits in young adults and the elderly. *Studia Psychologica, 47*(3), 167–178.

Hayes, A. F. (2017). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (2nd ed.). The Guilford Press.

Kasser, T. (2002). *The High Price of Materialism*. MIT Press.

Kasser, T., & Ryan, R. M. (1996). Further examining the American dream: Differential correlates of intrinsic and extrinsic goals. *Personality and Social Psychology Bulletin, 22*(3), 280-287.

Kasser, T., Ryan, R. M., Couchman, C., & Sheldon, K. M. (2004). *Materialistic values: Their causes and consequences*. In T. Kasser & A. Kanner (Eds.), Psychology and consumer culture: The struggle for a good life in a materialistic world (pp. 11–28). American Psychological Association. https://doi.org/10.1037/10658-002

Kim, Y., & Chen, H.-T. (2015). Discussion network heterogeneity matters: Examining a moderated mediation model of social media use and civic engagement. *International Journal of Communication, 9,* 2344–2365. http://ijoc.org

Kuiper, N. A., Martin, R. A., & Olinger, L. J. (1993). Humor and mental health. *Personality and Individual Differences, 14*(6), 785-797.

Kuiper, N. A. (2012). Humor and resiliency: Towards a process model of coping and growth. *Europe's Journal of Psychology, 8*(3), 475–491.

Lind, A., Rindfleisch, A., & Ahuvia, A. (2014). *Materialism.* In A. C. Michalos (Ed.), Encyclopedia of Quality of Life and Well-Being Research (pp. 4132–4135). Springer. https://doi.org/10.1007/978-94-007-0753-5_1743

Lo, O. T., Cheung, S. H., & Lai, V. K. W. (2024). Materialistic happiness is negatively associated with meaning in life: A serial double mediation model. *Scandinavian Journal of Psychology*, 10.1111/sjop.13063. Advance online publication. https://doi.org/10.1111/sjop.13063

Martin, R. A. (2007). *The Psychology of Humor: An Integrative Approach*. Academic Press.

Martin, *R. A.*, & *Ford*, *T*. (*2018*). *The psychology of humor: An integrative approach*. Academic Press.

Martin, R. A., & Lefcourt, H. M. (1983). Sense of humor as a moderator of the relation between stressors and moods. *Journal of Personality and Social Psychology, 45*(6), 1313-1324.

Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J., & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of Research in Personality, 37*(1), 48-75.

Marwa, S., & Xiaosong, G. (2023). Meaning in life as predicted by basic psychological needs: A study with Chinese undergraduate students. *Humanities and Social Sciences, 11*(6), 188–195. https://doi.org/10.11648/j.hss.20231106.12

Nezu, A. M., Nezu, C. M., & Blissett, S. E. (1988). Sense of humor as a moderator of the relation between stressful events and psychological distress: A prospective analysis. *Journal of Research in Personality, 22*(4), 505-517.

Ng, C. K. J., Kwan, L. Y. J., & Chan, W. (2024). A note on evaluating the moderated mediation effect. *Structural Equation Modeling: A Multidisciplinary Journal*, 31(2), 340–356. https://doi.org/10.1080/10705511.2023.2201396

Pearson, P. R., & Sheffield, B. F. (1974). Purpose in life and the Eysenck Personality Inventory. *Journal of Clin-*

ical Psychology, 30(4), 562-564.

Peterson, C., & Seligman, M. E. P. (2004). *Character Strengths and Virtues: A Handbook and Classification*. Oxford University Press.

Pinchot, A. (2020). *Humor as a moderator of neuroticism's effect on psychopathology and life satisfaction* (Master's thesis). St. John's University. St. John's Scholar.

Platt, T., Hofmann, J., Ruch, W., & Proyer, R. T. (2013). Duchenne display responses towards sixteen enjoyable emotions: individual differences between no and fear of being laughed at. *Motivation and Emotion, 37*, 776–786. https://doi.org/10.1007/s11031-013-9342-9

Reker, G. T., & Wong, P. T. P. (2012). Personal meaning in life and psychosocial adaptation in the later years. In P. T. P. Wong (Ed.), The human quest for meaning: *Theories, research, and applications* (2nd Ed., pp. 433-456). Routledge.

Richins, M. L. (2004). The Material Values Scale: Measurement properties and development of a short form. *Journal of Consumer Research, 31*(1), 209-219.

Richins, M. L., & Dawson, S. (1992). A consumer values orientation for materialism and its measurement: Scale development and validation. *Journal of Consumer Research, 19*(3), 303-316.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Roberts, J. A., & Clement, A. (2007). Materialism and its impact on satisfaction with life in a consumer society. *Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, 20*, 1-12.

Ruch, W. (1998). The sense of humor: Explorations of a personality characteristic. *Humor: International Journal of Humor Research, 11*(3), 213-245.

Ruch, W., & McGhee, P. E. (2014). Humor intervention programs. In A. C. Parks & S. M. Schueller (Eds.), *The Wiley Blackwell handbook of positive psychological interventions* (pp. 179–193). Wiley Blackwell. https://doi.org/10.1002/9781118315927.ch10

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Shiota, M. N., Keltner, D., & John, O. P. (2006). Positive emotion dispositions differentially associated with Big Five personality and attachment style. *Journal of Positive Psychology, 1*(2), 61-71.

Simione, L., & Gnagnarella, C. (2023). Humor Coping Reduces the Positive Relationship between Avoidance Coping Strategies and Perceived Stress: A Moderation Analysis. *Behavioral sciences (Basel, Switzerland), 13*(2), 179. https://doi.org/10.3390/bs13020179

Steger, M. F. (2009). Meaning in life. In S. J. Lopez & C. R. Snyder (Eds.), *Oxford handbook of positive psychology* (2nd ed., pp. 679–687). Oxford University Press.

Suls, J., & Martin, R. (2005). The daily life of the garden-variety neurotic: Reactivity, stressors, and distress. *Journal of Personality, 73*(6), 1485-1509.

Tackett, J. L., & Lahey, B. B. (2017). *Neuroticism*. In T. A. Widiger (Ed.), The Oxford handbook of the Five Factor Model (pp. 39–56). Oxford University Press.

Trzcińska, A., & Sekścińska, K. (2021). Financial status and materialism – The mediating role of self-esteem. *Australian Journal of Psychology, 73*(4), 557–568. https://doi.org/10.1080/00049530.2021.1944315

Vulić-Prtorić, A., & Bubalo, J. (2006). Skala smisla života - SSŽ. In V. Ćubela Adorić (Ed.), *Zbirka psihologijskih skala i upitnika - Svezak 3* (pp. 49-55). Sveučilište u Zadru.

Yang, Y., Wu, S., Duñabeitia, J. A., & Li, Y. (2021). The influence of L2 proficiency on bilinguals' creativity: The key role of adaptive emotion regulation strategies during the COVID-19 pandemic. *Frontiers in Psychology, 12,* 695014. https://doi.org/10.3389/fpsyg.2021.695014

5

Rasch Analysis of the Purdue Nonverbal Test: Any Use for Ancient Tools in the Modern Era?

Siniša Lakić
Department of Psychology, Faculty of Philosophy, University of Banja Luka
sinisa.lakic@ff.unibl.org

Biljana Mirković
Department of Psychology, Faculty of Philosophy, University of Banja Luka

Lana Vujaković
Department of Psychology, Faculty of Philosophy, University of Banja Luka

Abstract

Among the instruments used by psychologists in the Western Balkans, the Purdue Nonverbal Test (PNT) of visual perceptual abilities stands out as an iconic tool. Despite its venerable age of over 60 years and the clearly outdated norms questioning its validity, psychologists in Bosnia and Herzegovina continue to employ it, especially in driving ability assessments. With this in mind, we aimed to evaluate the test's functionality by establishing actual norms for new drivers via Rasch analysis. The study involved 721 participants (52.1% male), all of whom were final-year high school students. In addition to the PNT, participants completed the Short Sensation Seeking Scale (from SUPPS-P) and a specially designed Scale of Attitudes Towards Risky Driving Behaviours for this study. Analysis of the items' unidimensionality, local independence, and item fit exposed issues with three items, which were subsequently removed from further analysis. The corrected version showed no gender bias and no indication of items with significant differential functioning. Criterion validity analysis showed that participants with a driving license significantly outperformed those without one, together with expected differences appearing across high schools attended. We observed insignificant correlations with risky driving behaviours and sensation seeking, which aligns with theoretical expectations. In addition, the information value of the PNT peaked at lower levels of ability which is vital for assessing deficient driver candidates. We discuss the benefits of employing Rasch analysis for crafting similar selection instruments and adapting time-honoured tools in the contemporary era.

Keywords: *Rasch analysis, cognitive abilities, driving, psychometrics, Purdue Nonverbal Test*[1]

[1] The study was funded by the Traffic Safety Agency of the Republic of Srpska and represents a substantially updated English-language version of an internal technical report prepared for the Agency.

Introduction

After instances of near-miss traffic incidents, drivers frequently make snap judgments about other drivers' cognitive deficiencies. While such assessments may reflect bias, since non-cognitive personality attributes might better explain risky driving behaviours (see e.g., Bowen et al., 2020), there is substantial empirical evidence supporting lay theories of the association between cognitive abilities and hazardous traffic behaviours. There are early studies indicating a direct link between cognitive functioning and accident rates (e.g., Arthur et al., 1991; McKnight & McKnight, 1999), studies that show associations with driver training success and job performance (e.g., Bertua et al., 2005; Salgado et al., 2003), and studies conducted in specialised driving simulators linking cognitive abilities and risky behaviours in virtual traffic scenarios (e.g., Casutt et al., 2014; Kaber et al., 2016; and recently Collins, 2023 for train drivers). Given that cognitive decline is a natural consequence of the ageing process and of a number of medical conditions, including schizophrenia, dementia, and multiple sclerosis, one should not be surprised that a significant portion of this research has focused on determining the relationship between driving behaviour and cognitive functioning within these subpopulations (e.g., Mathias & Lucas, 2009; Reger et al., 2004; Schultheis et al., 2010).

The convergence of commonsense knowledge and scientific evidence has prompted many countries to introduce the evaluation of cognitive abilities when obtaining or renewing a driving license. However, legal regulations vary significantly among countries. In the majority of them psychological assessment which involves cognitive testing is mandated only for specific candidates: those with certain medical conditions, history of substance use, history of prior traffic violations, professional drivers, and those who have reached a certain age. Only a small number of countries mandate a psychological assessment for all cases of issuance and/or renewal of driving licenses (e.g., Spain, Brazil). Bosnia and Herzegovina is among them.

In particular, the Regulation on Health Requirements for Motor Vehicle Drivers in Bosnia and Herzegovina (Pravilnik, 2007) distinguishes among three categories of drivers, and accordingly specifies three different cognitive ability cut-offs. For non-professional drivers—who constitute the largest number of candidates—the disqualifying conditions are defined as "forms of intellectual insufficiency below borderline values, regardless of cause." For professional drivers whose primary occupation is operating motor vehicles, the disqualifying conditions are defined as "all forms of reduced intellectual ability classified as borderline or below, regardless of aetiology." Finally, for Category D drivers (commercial passenger vehicle operators), intellectual abilities "must not be below average." The consistency of such a classification is open to debate since one needs to define the cut-offs operationally. However, from the perspective of a psychometrician and applied practitioner, the choice of assessment instruments could be of even greater importance.

According to our informal inquiry of practising psychologists in the Republic of Srpska (a self-governing entity within Bosnia and Herzegovina), the Purdue Nonverbal Test (PNT; originally called Purdue Non-Language Test, Tiffin et al., 1958) appears to be the first choice among cognitive ability assessment tools. The PNT is a non-verbal test designed to evaluate visual-perceptual cognitive abilities while purportedly being "culture-free". The test comprises two parallel forms (A and B), each consisting of 48 items. Notably, our literature search revealed very few studies examining the instrument's psychometric properties, yet it seems to have gained considerable popularity in applied psychology in the former Yugoslavia already since the 1960s (according to Dautović & Borčić-Konjarek, 1998) and it has continued to be used in its successor states. For example, in one of two found studies from the region, Piri (2007) reports that at least until 2006, psychologists in Croatia—Bosnia and Herzegovina's neighbouring country—systematically employed the PNT to test driver candidates.

This is despite the fact that almost a decade earlier, Dautović and Borčić-Konjarek (1998), questioned the appropriateness of the then-used norms based on a study of 500 driving licence applicants. Their primary criticism centred on the asymmetric distribution of scores across all age groups, resulting from the test's easiness

for all except the oldest cohort (age 54 and above). These authors argued for revisions focusing on increasing the number of moderately difficult and difficult items to "enhance discrimination at higher score ranges." As will be seen in the remainder of this paper, we will demonstrate why we believe that a negatively skewed distribution is actually desirable for the intended purpose of the test in the context of drivers' evaluation. Furthermore, we will discuss how, rather than increasing the number of items, it would be more beneficial to make PNT-like tests more efficient through the use of modern technologies that were not available at the time.

The second important study on the PNT conducted in Croatia was carried out by Piri (2007). The author collected archival records of 413 drivers under 30 years of age and 237 drivers aged 65 and above. Rather than focusing on internal psychometric properties, this study examined the criterion validity of the scores registered at psychological assessment in 2002. Piri reported that, in the younger age group, PNT scores negatively correlated with the number of attempts required to pass the theoretical driving examination. However, no significant relationships were found with practical driving test performance, nor were there associations in either age group between PNT scores and legally documented traffic violations or accidents during the approximately four-year follow-up period. While these findings might appear to challenge the functionality of the PNT, the author convincingly argues that the reasons for the obtained null effect could have been the relative rarity of documented violations, the presence of a large number of external variables influencing both the passing of the practical test and traffic incidents, and the sample selection bias, as the study excluded individuals denied license due to intellectual dysfunction.

All in all, despite understandable reservations about examining the functioning of a test that is over 60 years old, we believe there are at least two reasons why this endeavour is valuable. Firstly, while its persistent use in practice likely stems from the limited availability of superior commercial alternatives in Bosnia and Herzegovina, we posit that practitioners would not continue to use the same instrument for decades unless it maintained some utility. Whether this is the case needs to be determined empirically.

Secondly, the aforementioned criticism directed at the PNT regarding the non-normal distribution of scores stems from reliance on the postulates of formulaic application of classical test theory. For this reason, this paper aims to present to interested readers the application of Rasch modelling, which is more appropriate for ability tests and adheres more consistently to the fundamental principles of developing psychological instruments—that one must always be aware of the purpose of testing, and that generic solutions are not always optimal.

With the above in mind, we investigated the psychometric properties of the PNT by exploring (1) its internal structure (unidimensionality, local independence, and internal consistency) and (2) its relationships with theoretically relevant external variables. As for criterion validity, we tested several hypotheses. To support its utility in drivers' testing, the PNT should show the absence of gender bias and higher performance among those who have already become drivers. As a cognitive ability measure, the PNT should discriminate among students from schools with varying academic demands. Finally, based on theoretical considerations and recent empirical findings (Anglim et al., 2022; Stanek & Ones, 2023), one should observe minimal associations with relevant non-cognitive attributes, specifically sensation seeking and its derivative, attitudes towards risky driving.

Method

Participants and Procedure

Given that the study was primarily aimed at testing the adequacy of the PNT for potential drivers, we selected high school seniors as the operational population. Almost all high school seniors in Bosnia and Herzegovina are 18 years old, meeting the age requirement for taking the driving test. Not only do they rep-

resent typical candidates for a Category B driving license, but many of them have also had the opportunity to obtain an A1 category license starting at the age of 16. Indeed, out of the total sample of 721 valid responses, nearly half (340, 49.6%) reported already holding a driving license (ranging from categories A1 to CE, with the majority holding category B).

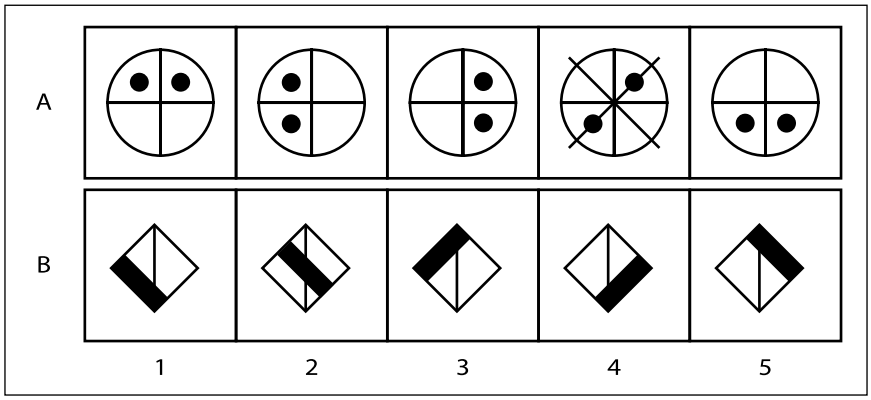
To better reflect the variability in cognitive abilities within the population, we included six high schools in the city of Banja Luka (the largest city in the Republic of Srpska), representing a range of academic achievement among students. Specifically, we included the Gymnasium (*n* = 231), Electrotechnical School (*n* = 94), and Music School (*n* = 13), which enrol students with higher academic performance in primary school, as well as the Polytechnic (n = 94), Technical (*n* = 172), and Agricultural School (*n* = 117), which tend to enrol students with lower educational achievement. This selection resulted in a slightly higher number of male respondents (376, 52.1%) compared to females (300, 41.5%), with 46 participants (6.4%) choosing not to disclose their gender. Although the sample was not perfectly balanced, this gender distribution more closely reflected the gender distribution of drivers in the Republic of Srpska, which is approximately 2:1 male to female (BIHAMK, 2018).

Data collection took place during 2021, after approvals were obtained from the Ministry of Education and Culture of the Republic of Srpska and from each individual school. A member of the research team, in collaboration with the school psychologist or pedagogue, administered the battery in a group-setting following a precisely defined protocol during a single 45-minute school class.

Measures

As should be clear from above, the primary focus of this study was the Purdue Nonverbal Test, specifically its Form A. The test consists of 48 nonverbal items. Each item presents five drawings composed of geometric shapes, where one drawing deviates from others by not following the pattern applied to the remaining combinations. Sample items, used for practice before independent work, are shown in Figure 1.

Figure 1
Initial practice items from the PNL test



The Short Sensation Seeking Scale was taken from the SUPPS-P instrument (see Cyders et al., 2014). SUPPS-P is a short version of a longer impulsivity assessment tool (UPPS-P; Lynam et al., 2006), which, in addition to sensation seeking, includes four other subscales: positive urgency, negative urgency, lack of perseverance, and lack of premeditation. The shortened versions comprise four items, to which participants respond using a 4-point Likert scale (ranging from *not at all* to *completely*). Example items for the sensation-seeking scale include: "*I quite enjoy taking risks*" and "*I welcome new and exciting experiences and sensations, even if they are a little frightening and unconventional*" Given the small number of items, the internal consistency coefficients were expectedly low, but reasonably satisfactory (α = .67). Other studies have shown adequate

psychometric characteristics of the scale, including its test-retest reliability (Dugre et al., 2019).

The Scale of Attitudes Towards Risky Driving Behaviours was developed for the purposes of this project, and the final version consisted of 20 items. The participants were asked to rate, on a 4-point scale (ranging from *never* to *often*), to what extent they believe certain, generally risky behaviours are justified for drivers in traffic. The statements were phrased in a way that tried to minimise socially desirable responding through positive framing. Example items include: "*Having a phone conversation while the vehicle is in motion, holding the phone in your hand, if the person is confident in their abilities*" and "*Driving aggressively with sudden accelerations and braking to demonstrate to less skilled drivers how to drive efficiently and alleviate traffic congestion*" The first unrotated factor accounted for 31% of the variance in the scores, with all items significantly loaded on it ($\lambda > .32$). The internal consistency of the summation score was high ($\alpha = .90$). The correlation with the Short Sensation Seeking Scale was moderate ($r = .31$, $BF_{10} > 1000$, $p < .001$), suggesting that they measure different constructs.

Data Analysis

After double data entry, the quality of the data was checked by analysing the theoretically possible range of values and using techniques to check for response patterns' plausibility (i.e., improbable runs of same responses, psychometric synonyms and antonyms scores). Clear indications of demotivated responding were observed for 14 participants, who were excluded from further analysis, resulting in a final sample of 721 participants. Given the advantages of the Rasch model in psychometric analysis of abilities tests, the techniques within this approach were the focus of the analysis. Specifically, the unidimensionality of the PNT was primarily tested using the Martin-Löf test, Ponocny's T_{1m} test, calculating the percentage of variance explained by the Rasch model, and conducting a principal component analysis on the residuals (see, e.g., Smith, 1996). Local independence was tested using Yen's Q3 measure, and item difficulty and discrimination were complemented by fit analyses of items and respondents using the Rasch model. Criterion validity was assessed by examining the relationship with variables such as gender, school type, possession of a driving license, sensation-seeking scales, and attitudes towards risky driving behaviours. Additionally, differential item functioning was analysed based on gender, and the informativeness of the test at different levels of ability was examined. Finally, score norming was performed, and optimisation was conducted for more efficient ability estimation. All analyses were conducted in R environment using the following packages: ltm (Rizopoulos, 2006), eRm (Mair & Hatzinger, 2007), perfit (Tendeiro & Meijer, 2014), sirt (Robitzsch, 2020), difR (Magis et al., 2010), TAM (Robitzsch et al., 2020), WrightMap (Torres Irribarra & Freund, 2014), eatATA (Becker et al., 2021), mirt (Chalmers, 2012), and ufs (Peters, 2021).

Results

After it had been established that each item exhibited at least minimal variability (with item difficulty ranging from .28 to .99), we performed analyses based on a factor model. Interestingly, none of the four frequently recommended methods for determining the number of common factors in exploratory factor analysis (Horn's parallel analysis, Horn, 1965; MAP, Velicer, 1976; Hull's method, Lorenzo-Seva et al., 2011; and Empirical Kaiser Criterion, Braeken & van Assen, 2017) suggested a unidimensional solution. The number of proposed factors ranged from two to seven, depending on the method and the type of correlation coefficient used for analysing the correlation matrix (phi, tetrachoric, or gamma).

However, the first unrotated factor was dominant, accounting for 34% of the total variability and saturating 46 out of 48 items with a value greater than 0.32 (which is frequently considered the recommended lower threshold, Costello & Osborne, 2005). Further analysis revealed a strong effect of item dif-

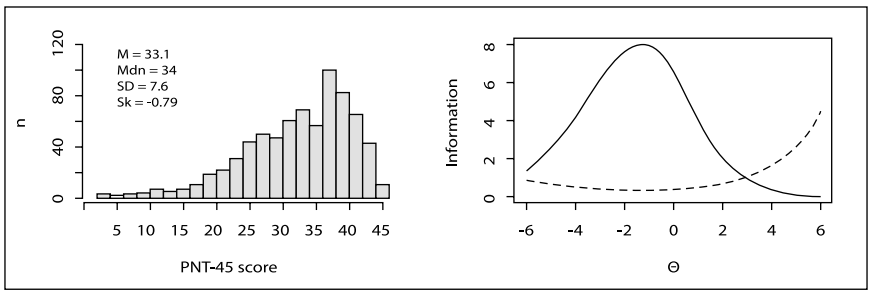
ficulty on the factor solution. Specifically, it was observed that the first two rotated components were relatively highly correlated (r = 0.47), but more importantly, both components were highly correlated with item difficulty (r = -0.66 and r = 0.86). In other words, the classical factorial approach on dichotomous data could not distinguish substantive effects from methodological artifacts, a long-recognised issue even when using corrective correlation matrices (e.g., Bock et al., 1988; Kubinger, 2003).

In contrast to the above, three of the four analyses typically used in Rasch modelling suggested good data fit with the unidimensionality assumption: the Martin-Löf test (LR = 388.1, df = 574, p = 1.00), the non-parametric Ponocny *T1m* test (p = 1.00), and the proportion of response variance attributable to the Rasch model (33.2%, which is above the recommended lower practical threshold of 20.0%, according to Linacre, 2006). The fourth analysis, a principal component analysis on the residuals remaining after extracting the first component, yielded a result of 2.14, which, since it exceeded the 2.0 threshold, was considered grounds for additional item grouping review and the search for potential patterns (see Boone & Staver, 2020). Further scrutiny of individual items was also prompted by the results of other analyses characteristic of Rasch modelling: the local independence analysis, corrected item-total correlations, and the fit of items to the Rasch model using infit MSQ and standardised t-scores.

A content analysis following these psychometric procedures suggested that three items were indeed problematic. Item #22 was identified as redundant since it shared an identical solution rule and similar graphical shapes with item #27, as was indicated by the highest residual correlation among 1128 item pairs (Q_3 = .37). Items #33 and #42 showed poor model fit (#33, t = 7.24; #42, t = 8.51), accompanied by item-total correlations below .20. These were also the two items with loadings below .32 on the first unrotated factor. A content analysis revealed that these items had two plausible solutions. Thus, we excluded mentioned items from further analysis, with subsequent analyses conducted on the remaining 45 items (from now on PNT-45).

Improved psychometric indicators were obtained for all the previously mentioned criteria on the corrected scale (e.g., all items had loadings on the first unrotated factor greater than 0.32, the eigenvalue of the first component of the residuals was reduced to 1.98, and the proportion of variance explained by the Rasch model increased to 34.3%). Moreover, the internal consistency reliability of the summation score was high for both the classical and factor models (α = *0.89*, ω = *0.90*), as well as for the reliability within the Rasch model, R = *0.87* (see Anselmi, Colledani, & Robusto, 2019). However, one of the distinctive advantages of the Rasch model is its capacity to describe the informativeness of the test depending on the level of ability of the respondents. As seen in Figure 2, the distribution of correct answers is markedly asymmetric, and the test can be characterised as being easy. The results from our sample are almost identical to those obtained by Dautović and Borčić-Konjarek (1998) on a subsample of respondents aged 16 to 19. However, Figure 2 also shows that the test's informativeness is greatest for assessing low achievement, specifically within the range from -3 to 0 theta. From a generic perspective, which assumes a symmetric, ideally normal distribution of

Figure 2
Distribution of PNT-45 total scores (n = 721) and test information curve with standard errors (dashed line)



abilities, such a distribution might appear disadvantageous. However, in practice, this type of distribution is ideal for the purpose for which it is used by psychologists in evaluating required cognitive ability. Namely, the objective of such testing is to reliably detect drivers who lack the requisite ability, and thus the reliability must be highest around the cut-off values.

Encouraging results were also obtained through criterion validity analysis. To begin with, the results were in favour of the hypothesis of no gender differences (Males: M = 32.9, SD = 8.06, Females: M = 33.4, SD = 7.12; $t_w(665) = -0.93$, d = -0.07, BF_{01} = 7.69, p = .355). Additionally, no item showed significant indications of differential item functioning (DIF). Specifically, only four items had one significant indicator each, even though each was tested with four DIF tests (Mantel-Haenszel, standardised proportion test, logistic regression, and Raju's test). In contrast to gender, expected differences were found based on the type of school attended by the student. The type of school accounted for more than a quarter of the variance ($F(5, 715) = 50.69$, $\eta^2 = 0.26$, $BF_{10} > 1000$, $p < 0.001$), with three school clusters identified. The students from the gymnasium, music school, and electrotechnical school had comparable results (pooled M = 37.0, SD = 5.7), significantly higher than the students from technical schools (M = 31.6, SD = 6.2), as well as the students from polytechnic and agricultural schools (pooled M = 28.1, SD = 7.9).

Of particular interest is the finding that the students holding a driving license (n = 340, M = 34.9, SD = 6.6) achieved moderately higher scores than those without a license (n = 345, M = 31.4, SD = 8.2), a difference that was statistically significant ($t_w(656) = 6.14$, d = 0.47, $BF_{10} > 1000$, $p < .001$). On the other hand, a null correlation was found between PNT scores and attitudes towards risky driving behaviours (r = .00, BF_{01} = 21.24, p = .992), while a statistically inconclusive—but practically negligible—low positive correlation was observed with sensation seeking (r = .08, BF_{01} = 2.05, p = .031). These correlations are fully in line with meta-analytic findings regarding the relationships between impulsivity, sensation seeking, and related personality constructs with intelligence (Anglim et al., 2022; Stanek & Ones, 2023). Such finding corroborates the rationale behind regulations requiring that psychological evaluation of candidate drivers includes both cognitive ability and personality testing. This makes even more sense in light of recent empirical evidence suggesting that lower cognitive ability increases the likelihood of making erroneous decisions, but that it does not systematically increase risk-taking preferences (Mechera-Ostrovsky et al., 2022).

The final stages of the analysis were focused on norming. Thanks to Rasch modelling, it was possible to nonlinearly transform the raw scores into person parameter estimates along with their 95% confidence intervals. Due to the PNT's higher informativeness for the lower ability, the confidence intervals were narrower for below-average performance (e.g., for a score of 20, the confidence interval ranged from -2.56 to -1.15 θ) than for above-average results (e.g., for a score of 43, the confidence interval was from +0.63 to +3.34 θ). As a result, it was possible to create a table that also includes percentile ranks of raw scores and an indication of whether a given lower confidence interval falls into any of the areas of insufficient intelligence (defined as IQ thresholds of 70, 80, and 90 points, depending on the driving license category). Finally, using an automatic item selection algorithm (Becker et al., 2021) by maximising test information at the average ability level (θ = 0), it was possible to create a scale of only five medium-difficulty items, where a respondent needs to correctly answer three out of five items to reasonably eliminate the suspicion that they have below average performance.

Discussion and Conclusion

So, what is our study telling us about the state of the ancient PNT in the modern age? At first glance, the PNT appears to be holding up well. It is content-valid for assessing visually saturated information processing, and the corrected version shows good psychometric properties in terms of internal structure.

Reasonable criterion correlations were obtained, without indications of gender bias. In addition—and as we have already emphasised—scores were distributed in such a way that they were highly informative for detecting intellectual insufficiency. When we also consider the possibility of administering an optimised version that might save precious time to practitioners in approximately 50% of testing cases, can we then unreservedly recommend it for further use when evaluating drivers' cognitive ability?

Unfortunately, the answer is no. The first and most important reason for such an answer is that the PNT can be found on the internet—albeit through a thorough search. This automatically weakens the arguments for its legal use, even though the target groups are unlikely to prepare themselves for testing. The second reason is more of a paradigmatic dilemma about the comparative validity of using relatively general tests such as the PNT versus modern testing options, which include domain-specific testing in the simulators. Thirdly, it remains questionable, both from an ethical and functional perspective, whether norms derived from young adults should be applied to licence renewal candidates, particularly older experienced drivers. Finally, the negatively skewed distribution of the scores—while appropriate for screening out unsuitable driver candidates—limits the PNT's utility in other contexts such as recruitment or educational selection where the goal is typically to discriminate among candidates across the full range of ability. Despite these limitations, we believe that our study contributes to the discussion on the use of appropriate psychometric techniques for the validation and adaptation of tests such as the PNT.

For example, our review of regional literature has led us to the impression that Rasch modelling—only slightly younger than the PNT—is still a neglected approach compared to factor or classical modelling of ability tests, despite obvious advantages. Not only is Rasch modelling more consistently adapted to the statistical analysis of dichotomous data, but it also directly enables advanced options such as optimisation of administration through direct test shortening or computerised adaptive testing and norming that is particularly focused on specific ability levels. Rasch modelling—like other techniques belonging to the IRT family—places greater focus on item content and more thorough item analysis. Finally, if in the past the obstacle to using Rasch analysis was that it could only be done in commercial and rather hermetic software solutions, today almost all analyses can be done in open-source code using R packages, as was the case in our study.

In the end, we should answer the question posed in the title: whether using ancient tools provides benefits in the modern era. We are convinced that it does. For example, technological advancements and the current AI revolution allow for a more efficient generation of tests that would have items analogous to the PNT (e.g., Choi & Zhang, 2019; Gierl et al., 2015; Sayin et al., 2023). Considering all the psychometric advantages we have demonstrated in this paper, we believe it would be worthwhile to embark on creating a new version that would be administered in a computerised adaptive manner. That would also be an opportunity to upgrade the PNT by increasing its informativeness for higher visual-perceptual cognitive abilities, which might be beneficial for other selective purposes. All in all, old is not dead, but a thorough modernisation would be needed for the PNT to maintain its fitness and relevance in contemporary assessment. Before embarking on such a process, it is essential to clearly argue for the usefulness of developing such general-type tests and to ensure their ethical safeguarding in an age when it is hard to hide anything, so that we do not end up with yet another zombie wandering the Internet.

References

Anglim, J., Dunlop, P. D., Wee, S., Horwood, S., Wood, J. K., & Marty, A. (2022). Personality and intelligence: A meta-analysis. *Psychological Bulletin, 148*(5-6), 301–336. https://doi.org/10.1037/bul0000373

Anselmi, P., Colledani, D., & Robusto, E. (2019). A comparison of classical and modern measures of internal consistency. *Frontiers in Psychology, 10*, 2714. https://doi.org/10.3389/fpsyg.2019.02714

Arthur, W., Barret, G. V., & Alexander, R. A. (1991). Prediction of vehicular accident involvement: A meta-analysis. *Human Performance, 4*(2), 89-105. https://doi.org/10.1207/s15327043hup0402_1

Becker, B., Debeer, D., Sachse, K. A., & Weirich, S. (2021). Automated test assembly in R: The eatATA package. *Psych, 3*(2), 96-112. https://doi.org/10.3390/psych3020010

Bertua, C., Anderson, N., & Salgado, J. F. (2005). The predictive validity of cognitive ability tests: A UK meta-analysis. *Journal of Occupational and Organizational Psychology, 78*(3), 387-409. https://doi.org/10.1348/096317905X26994

Bowen, L., Budden, S. L., & Smith, A. P. (2020). Factors underpinning unsafe driving: A systematic literature review of car drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 72*, 184-210. https://doi.org/10.1016/j.trf.2020.04.008

BIHAMK. (2018). Informacija o ukupnom broju aktivnih i proizvedenih/izrađenih vozačkih dozvola u BiH u periodu januar – decembar 2017 [Information on the total number of active and produced/issued driving licenses in BiH for the period January - December 2017]. https://bihamk.ba/assets/upload/informacija_o_vozackim_dozvolama%20-%20Copy%201.pdf

Bock, R. D., Gibbons, R., & Muraki, E. (1988). Full-information item factor analysis. *Applied Psychological Measurement, 12*(3), 261-280. https://doi.org/10.1177/014662168801200305

Braeken, J., & Van Assen, M. A. (2017). An empirical Kaiser criterion. *Psychological Methods, 22*(3), 450-466. https://doi.org/10.1037/met0000074

Casutt, G., Martin, M., Keller, M., & Jäncke, L. (2014). The relation between performance in on-road driving, cognitive screening and driving simulator in older healthy drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 22*, 232-244. https://doi.org/10.1016/j.trf.2013.12.007

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software, 48*(6), 1-29. https://doi.org/10.18637/jss.v048.i06

Choi, J., & Zhang, X. (2019). Computerized item modeling practices using computer adaptive formative assessment automatic item generation system: A tutorial. *The Quantitative Methods for Psychology, 15*(3), 214-225. https://doi.org/10.20982/tqmp.15.3.p214

Collins, M. D. (2024). Train driver selection: The impact of cognitive ability on train driving performance. *International Journal of Selection and Assessment, 32*(2), 249-260. https://doi.org/10.1111/ijsa.12425

Costello, A. B., & Osborne, J. (2019). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation, 10*(1), 7. https://doi.org/10.7275/jyj1-4868

Cyders, M. A., Littlefield, A. K., Coffey, S., & Karyadi, K. A. (2014). Examination of a short English version of the UPPS-P Impulsive Behavior Scale. *Addictive Behaviors, 39*(9), 1372-1376. https://doi.org/10.1016/j.addbeh.2014.02.013

Dautović, M., & Borcić-Konjarek, L. (1998). Use of the Purdue Non-verbal Test in a group of candidates for a driving license. *Arhiv za Higijenu Rada i Toksikologiju, 49*(2), 179-187.

Dugre, J. R., Giguere, C-É., Percie du Sert, O., Potvin, S., Dumais, A., & Consortium Signature (2019). The Psychometric Properties of a Short UPPS-P Impulsive Behavior Scale Among Psychiatric Patients Evaluated in an Emergency Setting. *Frontiers in Psychiatry, 10*, 1-9. https://doi.org/10.3389/fpsyt.2019.00139

Gierl, M. J., Ball, M. M., Vele, V., & Lai, H. (2015). A method for generating nonverbal reasoning items using n-layer modeling. In *Computer Assisted Assessment. Research into E-Assessment: 18th International Conference, CAA 2015* (pp. 12-21). https://doi.org/10.1007/978-3-319-27704-2_2

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185. https://doi.org/10.1007/BF02289447

Kaber, D., Jin, S., Zahabi, M., & Pankok Jr, C. (2016). The effect of driver cognitive abilities and distractions on situation awareness and performance under hazard conditions. *Transportation Research Part F: Traffic Psychology and Behaviour, 42*, 177-194. https://doi.org/10.1016/j.trf.2016.07.014

Kubinger, K. D. (2003). On artificial results due to using factor analysis for dichotomous variables. *Psychology Science, 45*(1), 106-110. https://doi.org/10.1002/0471264385.wei0205

Linacre, J. M. (2006). Data variance explained by Rasch measures. *Rasch Measurement Transactions, 20*(1), 1045. https://www.rasch.org/rmt/rmt201a.htm

Lorenzo-Seva, U., Timmerman, M. E., & Kiers, H. A. (2011). The Hull method for selecting the number of common factors. *Multivariate Behavioral Research, 46*(2), 340-364. https://doi.org/10.1080/00273171.2011.564527

Lynam, D. R., Whiteside, S. P., Smith, G. T., & Cyders, M. A. (2006). *The UPPS-P: Assessing five personality pathways to impulsive behavior*. West Lafayette, IN: Purdue University. [Unpublished report]

Magis, D., Béland, S., Tuerlinckx, F., & De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*(3), 847-862. https://doi.org/10.3758/BRM.42.3.847

Mair, P., & Hatzinger, R. (2007). Extended Rasch modeling: The eRm package for the application of IRT models in R. *Journal of Statistical Software, 20*(9), 1-20. https://doi.org/10.18637/jss.v020.i09

Mathias, J. L., & Lucas, L. K. (2009). Cognitive predictors of unsafe driving in older drivers: a meta-analysis. *International Psychogeriatrics, 21*(4), 637-653. https://doi.org/10.1017/S1041610209009119

McKnight, A. J., & McKnight, A. S. (1999). Multivariate analysis of age-related driver ability and performance deficits. *Accident Analysis & Prevention, 31*(5), 445-454. https://doi.org/10.1016/S0001-4575(98)00070-7

Mechera-Ostrovsky, T., Heinke, S., Andraszewicz, S., & Rieskamp, J. (2022). Cognitive abilities affect decision errors but not risk preferences: A meta-analysis. *Psychonomic Bulletin & Review, 29*(5), 1719-1750. https://doi.org/10.3758/s13423-022-02097-x

Peters, G. (2021). ufs: A collection of utilities for univariate frequency statistics and psychometrics [R package version 0.5.0]. CRAN. https://CRAN.R-project.org/package=ufs

Pravilnik o zdravstvenim uslovima koje mora ispunjavati vozač motornog vozila [Regulation on health requirements for motor vehicle drivers]. (2007). *Službeni glasnik Bosne i Hercegovine [Official Gazette of Bosnia and Herzegovina]*, 13/2007 & 89/2011.

Reger, M. A., Welsh, R. K., Watson, G., Cholerton, B., Baker, L. D., & Craft, S. (2004). The relationship between neuropsychological functioning and driving ability in dementia: a meta-analysis. *Neuropsychology, 18*(1), 85-93. https://doi.org/10.1037/0894-4105.18.1.85

Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response analysis. *Journal of Statistical Software, 17*(5), 1-25. https://doi.org/10.18637/jss.v017.i05

Robitzsch, A. (2020). sirt: Supplementary item response theory models [R package version 3.9-4]. CRAN. https://CRAN.R-project.org/package=sirt

Robitzsch, A., Kiefer, T., & Wu, M. (2020). TAM: Test analysis modules [R package version 3.5-19]. CRAN. https://CRAN.R-project.org/package=TAM

Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., De Fruyt, F., & Rolland, J. P. (2003). A meta-analytic study of general mental ability validity for different occupations in the European community. *Journal of*

Applied Psychology, 88(6), 1068-1081. https://doi.org/10.1037/0021-9010.88.6.1068

Sayın, A., Bozdağ, S., & Gıerl, M. J. (2023). Automatic item generation for non-verbal reasoning items. *International Journal of Assessment Tools in Education, 10*(Special Issue), 132-148. https://doi.org/10.21449/ijate.1279120

Schultheis, M. T., Weisser, V., Ang, J., Elovic, E., Nead, R., Sestito, N., … & Millis, S. R. (2010). Examining the relationship between cognition and driving performance in multiple sclerosis. *Archives of Physical Medicine and Rehabilitation, 91*(3), 465-473. https://doi.org/10.1016/j.apmr.2009.11.026

Smith, R. M. (1996). A comparison of methods for determining dimensionality in Rasch measurement. *Structural Equation Modeling: A Multidisciplinary Journal, 3*(1), 25-40. https://doi.org/10.1080/10705519609540027

Tendeiro, J. N., & Meijer, R. R. (2014). perfit: An R package for person-fit analysis in IRT. *Journal of Statistical Software, 74*(5), 1–27. https://doi.org/10.18637/jss.v074.i05

Torres Irribarra, D., & Freund, R. (2014). Wright Map: IRT item-person map with ConQuest integration. CRAN. https://CRAN.R-project.org/package=WrightMap

Triffin, J., Grubner, A., & Inaba, K. (1958). *Purdue non-language test: Preliminary Manual*. West Lafayette: Occupational Research Center, Purdue University.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41*(3), 321-327. https://doi.org/10.1007/BF02293557

6

Grit and Cognitive Load in Students with Different Achievement Goal Orientation Profiles

Barbara Rončević Zubković
Department of Psychology, Faculty of Humanities and Social Sciences, University of Rijeka
barbararz@uniri.ffri.hr

Rosanda Pahljina-Reinić
Department of Psychology, Faculty of Humanities and Social Sciences, University of Rijeka

Marta Batel
Teaching Institute of Public Health of Istarska County, Service for Mental Health and Addiction Prevention, Pula, Croatia

Dora Korać
Department of Psychology, Faculty of Humanities and Social Sciences, University of Rijeka
Department of Psychology, Faculty of Croatian Studies, University of Zagreb

Abstract

The aim of this study was to identify goal orientation profiles among students from different academic fields and examine differences among these profiles in terms of students' grit and perceived cognitive load during an authentic knowledge assessment situation. A total of 309 students completed a goal orientation questionnaire and a brief grit scale (assessing perseverance of effort and consistency of interest) one week prior to a midterm exam. Immediately after the exam, they reported on intrinsic, extraneous, and germane cognitive load experienced during the exam. Latent profile analysis identified three groups of students with different goal orientation profiles: success-oriented students (27.8%), indifferent students (53.9%), and work-avoidance students (18.3%). In the final step of the latent profile analysis, the Bolck-Croon-Hagenaars (BCH) approach was applied to identify differences in grit and cognitive load between the goal orientation profiles. Results indicated that success-oriented students reported higher grit, as well as greater germane cognitive load, reflecting higher mental effort investment in general, but also in the specific knowledge assessment situation. Indifferent and work-avoidant

students showed no significant differences. No significant differences were found among goal orientation profiles in perceptions of intrinsic and extraneous cognitive load, i.e., the perceived complexity and clarity of exam tasks. The findings are discussed in the context of linking motivational concepts with cognitive load theory.

***Keywords**: grit, cognitive load, performance, achievement goal orientations, profiles*

***Acknowledgment:** We would like to acknowledge Zita Jurić for her assistance with data collection and participation in the broader study. We are also deeply grateful to the professors and teaching assistants who kindly allowed us to collect data during their class time, facilitating the completion of this research.*

***Funding:** This research is supported by the University of Rijeka, under grant Learning through self-regulation in the context of dynamic changes in society and education (UKS; uniri-iskusni-drustv-23-118).*

Introduction

Achievement motivation is crucial to student success at all educational levels, including higher education. While the pressure to achieve high grades at the university programmes may be less apparent than in high school, successfully completing assignments and exams remains essential for academic advancement and degree attainment. Students deal with educational challenges differently, and the goals they set in achievement contexts shape their approach to and perception of grading tasks.

Tendencies to prefer certain types of goals and outcomes over some others in achievement-related settings are referred to as achievement goal orientations (Niemivirta at al., 2019). Two main types of goal orientations are commonly distinguished. The first is mastery orientation (or learning orientation), which reflects a student tendency to focus on the development of abilities and mastering the task. The second type is performance orientation, where the individual is focused on outcomes and demonstrating their abilities (Ames, 1992). Individuals with a mastery orientation are motivated by the intrinsic value of learning, such as acquiring new knowledge and skills, as well as by desire to understand the material (Dweck & Leggett, 1988). In contrast, when endorsing performance orientation, students' primary goal is to outperform others. In trichotomous achievement goal framework, performance orientation is separated into approach and avoidance tendencies. Whereas the goal of performance-approach orientation is to demonstrate competence, performance-avoidance orientation reflects the tendency to avoid demonstrating incompetence in comparison to others (Elliot & Church, 1997). In a 2x2 model of achievement goals (Elliot & McGregor, 2001), both mastery and performance orientations are divided regarding their valence dimension into positive ("approaching success") and negative ("avoiding failure") orientations. Mastery avoidance goals reflect the tendency to avoid failure in learning and the deterioration of skills. Mastery avoidance goals are rarely examined and have weaker empirical support compared to other orientations (Maehr & Zusho, 2009). Some conceptualisations of achievement goal orientations also identify work-avoidance as a distinct goal orientation, characterised by the tendency to avoid challenges and invest as little effort as possible in completing academic tasks (Nicholls et al., 1985).

Niemivirta (2002, Niemivirta et al., 2019) has differentiated between five goal orientations. In addition to performance-approach, performance-avoidance, and work-avoidance orientations, the model distinguishes between mastery intrinsic and extrinsic goal orientations. Mastery intrinsic goals correspond to mastery approach goals, while mastery extrinsic goals reflect the desire for achieving good grades and obtaining absolute success. Mastery-extrinsic oriented students evaluate their achievement using external criteria (like grades) to evaluate the attainment but without comparing themselves to others.

In the first few decades of research on goal orientations, the relationships between specific goal orientations and various educational outcomes were widely explored (Senko et al., 2011). However, as it became evident that individuals can pursue different goals at the same time (Pintrich, 2000), attention shifted to examining multiple goal orientations simultaneously (Niemivirta et al., 2019; Senko et al., 2011). This person-centred approach groups students into distinct, homogeneous profiles based on similar patterns of goal orientations and offers a deeper understanding of students' motivational tendencies (Tuominen-Soini et al., 2011, 2012). Although various achievement goal profiles were identified in numerous studies (Wormington & Linnenbrink-Garcia, 2017), several studies using Niemivirta's (Helsinki 5) model have consistently resulted in four-profile solutions (Niemivirta et al., 2019). The following groups are typically identified: a group with a dominant tendency towards mastery (mastery-oriented students), a group with a tendency towards performance or success (students oriented mainly on performance or both performance and mastery goals), a group with a tendency towards avoidance (avoidance-oriented students), and a group of students without a dominant tendency towards any specific achievement goal orientation (indifferent students). Earlier studies on Croatian university and high school students yielded similar four profile solutions (Pahlji-

na-Reinić, 2022; Pahljina-Reinić et al., 2024). Both mastery-oriented and success-oriented students concentrate on understanding and mastery. However, success-oriented students also tend to focus on comparison with other students. Indifferent students, on the other hand, are those who aim to meet basic expectations without investing a significant amount of effort. Compared to other groups, avoidance-oriented students set lower mastery goals and strive to put in the least possible effort in studying (Niemivirta et al., 2019).

Building on these findings, it is plausible to expect that students' achievement goals are also related to their grit. Grit is defined as perseverance and passion for long-term goals, and it comprises two dimensions: perseverance of effort, which refers to the ability to sustain effort in the face of adversity, and consistency of interests, which refers to the stability of interests over time (Duckworth et al., 2007). Gritty individuals pursue their goals with great effort and determination, overcoming challenges and progressing despite obstacles and failures. They also tend to attain higher levels of education (Duckworth et al., 2007).

Muenks et al. (2018) suggested that students with *approach* goal orientations, whether focused on mastery and skill development or on superior performance, are likely to show more grit when faced with challenges compared to those with *avoidance* goal orientations. In line with this, they found that both mastery and performance-approach orientations were positively related, while performance-avoidance was negatively related to perseverance of effort. However, consistency of interest was positively linked only to mastery goal orientation. Alhadabi and Karpinski (2020), as well as Sadoughi and Eskandari (2024) reported comparable findings, adding that performance-avoidance was also (weakly) negatively associated with consistency of interest. Han et al. (2023) confirmed that mastery orientation was positively related to perseverance of effort, while performance-avoidance had a negative effect, though they found no significant effect of performance-approach goal orientation. Zhao et al. (2024) found a positive correlation between mastery goal orientation and overall grit, though they did not examine grit dimensions separately.

To our knowledge, differences in grit across goal orientation profiles have not yet been explored. However, based on previous findings, we can hypothesise that mastery-oriented and success-oriented students would exhibit higher levels of grit compared to indifferent or work-avoidant students. While performance orientation has not been consistently linked to grit, we propose that students who combine a performance-approach orientation with mastery goal orientation are likely to demonstrate high levels of grit.

Whereas grit is related to long-time perseverance and effort expenditure, goal orientations can also affect perceptions of specific academic tasks and challenges. It is already acknowledged that both mastery- and success-oriented students invest considerable effort in academic tasks (Pintrich, 2000; Tuominen-Soini et al., 2011, 2012). However, since success-oriented students are more preoccupied with possible failures in school, they are at risk for certain adverse outcomes such as emotional exhaustion and higher sense of inadequacy as a student (Tuominen et al., 2020; Tuominen-Soini et al., 2012). Therefore, they might perceive academic assignments, especially the ones that are graded, as more threatening and difficult. Task difficulty and effort at the task specific level essentially overlap with cognitive load constructs (Plass & Kalyuga, 2019).

Cognitive load refers to the degree of strain on the cognitive system, specifically on an individual's working memory capacity, caused by information processing during learning or task-solving. According to widely recognised classifications (Klepsch & Seufert, 2021; Sweller et al., 1998), there are three types of cognitive load: intrinsic cognitive load (ICL), which arises from the complexity of the presented material and is related to the individual's prior knowledge; extraneous cognitive load (ECL), which is triggered by ambiguous task formulation or inadequate design of the material (such as irrelevant information or unclear instruction); and germane cognitive load (GCL), which reflects deliberate investment of additional internal resources and an individual's willingness to engage actively in task-solving and persevere. Scheiter et al. (2020) distinguish between mental load, which refers to the resources required to perform a task, and mental effort, which refers to the resources invested in performing the task. Feldon et al. (2019) argue that mental effort reflects a willingness to fully engage under the demands of the cognitive load imposed by a learn-

ing context. Because GCL entails the intentional investment of internal resources, it is often equated with mental effort. Seufert (2020) argues that if learners still have working memory resources available, they can allocate GCL to learning and self-regulation while engaging with the task, despite the task's intrinsic and extraneous demands. However, GCL is influenced not only by the available capacity of working memory and characteristics of the learning task, but also by learners' willingness to allocate their resources to specific learning activities and engage in them for extended period. Learners' willingness to actively engage in a specific learning activity involves investing additional working memory resources to process information related to the learning situation, as well as effectively regulating emotions that may arise during the learning or problem-solving process. This is influenced by the learner's motivational tendencies, which contribute significantly to persistence and success in the task (Feldon et al., 2019; Schnotz et al., 2009).

Findings from sparse previous research exploring relations between goal orientations and GCL mostly indicate positive relations between mastery orientation and GCL (Cook et al., 2017; Sunawan et al., 2021) with some exceptions (Xu et al., 2021). Performance orientation was rarely examined, indicating positive relation of GLC with its approach dimension (Cook et al., 2017) and no significant correlation with its avoidance dimension (Cook et al., 2017; Sunawan et al., 2021). Therefore, it can be expected that students with distinct goal orientation profiles would differ in reported GCL in specific knowledge assessment situation. We propose that mastery-oriented students will report highest effort investment (GCL) compared to other achievement goal profiles, while work-avoidant students will report minimal GCL, since they are inclined to avoid engagement and effort investment (King & McInerney, 2014; Niemivirta et al., 2019; Tuominen et al., 2020). Success-oriented students are expected to invest substantial effort, similar to mastery-oriented students. Previous findings show that in students with performance goal orientation, other beliefs such as their implicit theories of intelligence may influence their effort investment. Yu and McLellan (2020) found that performance-oriented students with a fixed mindset (i.e., the belief that abilities are unchangeable) reported low perseverance and high self-handicapping. This suggests that such students are less inclined to exert effort, likely to avoid implications of low ability. Conversely, when students view their abilities as malleable, performance goals shift towards demonstrating their developed skills. Recognising that effort supports their growth, these students sustain effort and perseverance. It should be noted that for this group of students, performance goals coexisted with mastery goals, aligning closely with success-oriented profiles obtained in other studies.

Since ICL and ECL depend on prior knowledge and task clarity, predicting their relationship with goal orientations is more complex. However, as mastery-oriented students tend to gain knowledge and understanding, they are likely to be more familiar with exam material, and therefore, perceive lower ICL. In contrast, indifferent, and especially work-avoidant students may perceive higher ICL, as well as elevated ECL since they are more prone to avoid effort and have higher cynicism towards the meaning of school (Tuominen-Soini et al., 2012). Success-oriented students could perceive tasks as more threatening and difficult as they are concerned with possible failures (Tuominen-Soini et al., 2011). Previous studies are not completely consistent in findings, mainly indicating no or weak correlations of goal orientations with ICL and ECL (Cook et al., 2017; Sunawan et al., 2021; Xu et al., 2021). In a study investigating relation of goal orientations to tenseness in achievement situation, no relation between goal orientations and perceived difficulty of the task was found, regardless of differences in tenseness (Wimmer et al., 2018).

Despite these initial findings, further research is needed to relate motivational concepts, such as goal orientation profiles and sustained perseverance in challenging contexts to the cognitive load, and more precisely its specific types. Although the need to integrate motivational constructs with cognitive load has been recognised for some time (e.g., Feldon et al., 2019; Schnotz et al., 2009), existing findings remain limited. This study represents a step in that direction, offering integrative examination of these constructs. More specifically, the study aimed to compare students' achievement goal orientation profiles in terms of grit, as a trait variable, and more situation-specific measures of cognitive load and performance on a midterm exam.

Methods

Participants

A total of 418 participants (63.4% female) were recruited for the broader study at various faculties of the University of Rijeka; 309 students participated in the first and 377 in the second assessment point. In the present study, we analysed the results of 309 participants that took part in the first assessment point (69.6% female). Forty-one out of them did not take part in the second time point. The participating students were mostly undergraduate students (4.9% first-year, 58.6% second-year, 23.9% third-year) and 12.6% were first-year graduate students. The sample was convenient as the selection was based on the willingness of the instructors to allocate a portion of their class time for the implementation of the study. The students were recruited from the Faculty of Economics and Business (n = 126; 41.9%), Faculty of Engineering (n = 44; 14.2%), Faculty of Law (n = 58; 18.8%), Faculty of Humanities and Social Sciences-Psychology (n = 32; 10.4%), and from different teacher education programmes studying the same course (n = 49; 15.9%).

Instruments

Achievement Goal Orientations were assessed by an instrument designed to assess five goal orientations with three items per scale (Niemivirta, 2002; Croatian adaptation Pahljina-Reinić, 2022). The scales are as follows: 1) Mastery-Intrinsic Orientation—assesses students' focus on learning, understanding content, and gaining competencies (e.g., *"To acquire new knowledge is an important goal for me"*); 2) Mastery-Extrinsic Orientation—assesses the students' aspirations to achieve high grades and success in courses (e.g., *"My goal is to succeed at the university"*); 3) Performance-Approach Orientation—assesses students' focus on relative abilities and judgments of competence (e.g., *"An important goal for me is to do better than other students"*); 4) Performance-Avoidance Orientation—assesses students' tendency to avoid situations that could demonstrate relative incompetence (e.g., *"I try to avoid situations in which I might fail or make a mistake"*); 5) Work-Avoidance Orientation—assesses students' tendency to avoid challenges and minimise effort in learning (e.g., *"I try to get away with as little effort as possible in my academic work"*). The participants were asked to rate the extent to which each statement applies to them using a 7-point Likert-type scale ranging from 1 (*not true at all for me*) to 7 (*very true to me*), where a higher score indicates a greater expression of each of the five goal orientations. The internal consistencies of the scales are shown in Table 1.

Grit was assessed with the Short Grit Scale (Duckworth & Quinn, 2009; Croatian adaptation Zrilić, 2018). The scale consists of two four-item subscales: Consistency of Interest (e.g., *"I often set a goal but later choose to pursue a different one"*) and Perseverance of Effort (e.g., *"Setbacks don't discourage me"*). The participants rated the extent to which each statement applies to them using a 5-point Likert-type scale ranging from 1 (*not true at all for me*) to 5 (*very true to me*), with a higher score indicating greater grit. The internal consistencies of the scales are shown in Table 1.

Perceived cognitive load was assessed with the Cognitive Load Questionnaire (Zu et al., 2021; Croatian adaptation Korać & Rončević Zubković, 2023). The scale comprises eight items measuring three types of cognitive load perceived by the participants in a test situation: the Intrinsic Cognitive Load (three items, e.g., *"The material covered by the exam was very complex"*); the Extraneous Cognitive Load (three items, e.g., *"The exam had confusing language that was not clear to me"*); the Germane Cognitive Load (two items, e.g., *"I devoted a lot of mental effort in finding and applying the relevant concepts needed to answer the exam questions"*). Each item was rated on a Likert scale from 1 (*not at all the case*) to 9 (*completely the case*), with

a higher score reflecting a greater tendency to assess exam as more difficult (ICL) and ambiguous (ECL), as well as a greater level of effort devoted to the exam (GCL).

The internal consistencies of the ICL and ECL scales are shown in Table 1.

Procedure

The study was carried out during regular class sessions at various faculties at the University of Rijeka. Each student group participated in two phases of the research. The first phase occurred approximately one week before a regularly scheduled midterm exam, in agreement with the course instructor. During this phase, the students completed scales assessing goal orientations and grit (among other scales not presented in this paper). The second phase was conducted immediately after the midterm exam when the participants filled out the cognitive load questionnaire. The participants were subsequently provided with a link to a digital table, where they were required to record the number of points they had achieved on the midterm exam and maximal possible score. A unique participant-generated code was used by the participants throughout all stages of the study, including during performance data entry. However, only a small percentage of students (n = 81) provided data about their performance, so we decided not to further analyse these results.

At the start of each phase of the study, the participants were informed, both in writing and orally, about the purpose and conditions of the research. These included voluntary participation, the option to withdraw at any time, guaranteed anonymity, and the possibility of contacting the researchers via email for further inquiries.

Data Analyses

Since the factor structure of the goal orientation scale had been validated in a previous study on a university student sample (Pahljina-Reinić, 2022)—and given that this variable assesses students' general tendencies regardless of a specific educational context—Confirmatory Factor Analysis (CFA) was not conducted in the present study. However, as cognitive load is context-specific variable, CFA and multigroup CFAs using maximum likelihood (ML) estimation were employed to assess the structural validity of the questionnaire, considering that the students assessed cognitive load for different midterm exams depending on their study programme. Adequacy of model fit was assessed by comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardised root mean square residual (SRMR) using the following cut-off values: CFI > .90, RMSEA < .06 and SRMR <.08 (Browne & Cudeck, 1992; Hu & Bentler, 1999).

To identify the students with similar patterns of achievement goal orientation, Latent Profile Analysis (LPA) was performed, using the composite scores of achievement goal orientation scales. Five groups were added stepwise to explore the most optimal data fit in terms of number of profiles. In line with the existing recommendations (Masyn, 2013), the following statistical criteria were used to select the optimal time-specific solution: Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Vuong-Lo-Mendell-Rubin (VLMR) likelihood ratio test, and Lo-Mendell-Rubin (LMR) adjusted likelihood ratio test. A better fit to the data is indicated by a model with lower AIC and BIC values, while p-values of the VLMR and LMR tests less than .05 indicate that the estimated model is preferable over the reduced model. The classification quality (entropy value > .70), meaningfulness and interpretability of the latent profiles were also considered for choosing the best-fitting model. In the last step of LPA, Bolck-Croon-Hagenaars (BCH) approach was applied to identify statistically significant differences between the mean scores of grit and cognitive load across profile groups. The models were estimated using the full-information maximum likelihood estimation with robust standard errors, as implemented in Mplus Statistics Software Version 8.10 (Muthén & Muthén, 1998–2023).

Results

CFA failed to identify the three-factor cognitive load model given convergence problems involving GCL items. Therefore, both GCL items were excluded from the model. The two-factor CFA on the remaining six items, hypothetically representing ICL and ECL, suggested a good fit to the data: χ^2 (8) = 9.53, p > .05, CFI = .99, RMSEA = .023, SRMR = .022. We used this solution in further analyses, while GCL was analysed using a single-item measure. Out of the two items applied, the item *"I devoted a lot of mental effort in finding and applying the relevant concepts needed to answer the exam questions"* was chosen. Its content aligns better with the concept of GCL that refers to the active investment of germane resources by the learner (Seufert, 2018) and "reflects the effort that contributes to the construction of schemas" (Sweller et al., 1998, p. 259).

The metric measurement invariance of ICL and ECL cognitive load measures across five study programmes was confirmed through multigroup CFAs. The configural invariance model suggested an acceptable fit to the data, χ^2 (40) = 65.50, p < .001, CFI = .98, RMSEA = .092, SRMR = .045, as did the metric invariance model, χ^2 (56) = 86.50, p < .001, CFI = .98, RMSEA = .085, SRMR = .086. The comparisons between these models yielded a nonsignificant χ^2-difference test, $\Delta\chi^2$ = 20.99, Δdf = 16, p = .18, and the change in CFI (ΔCFI = .004) was below the recommended threshold of ΔCFI < .01 (Chen, 2007), thus indicating that the metric invariance was met.

Descriptive statistics, internal consistencies, and correlations between goal orientations, grit, and cognitive load are presented in Table 1.

GCL showed a positive correlation with both ICL and ECL. However, GCL and grit dimensions were not significantly correlated. Both dimensions of grit correlated positively with mastery-intrinsic and mastery-extrinsic goal orientations, while GCL correlated positively with mastery-extrinsic goal orientation. Performance-avoidance goal orientation was negatively correlated with consistency of interest. As expected, work-avoidance orientation negatively correlated with grit but positively correlated with ECL.

Table 1

Descriptive Statistics, Internal Consistencies, and Bivariate Pearson Correlations for Goal Orientation, Grit, and Cognitive Load Scales

Variable	1	2	3	4	5	6	7	8	9	10
1. Mastery-intrinsic	-									
2. Mastery-extrinsic	.43**	-								
3. Performance-approach	.26**	.63**	-							
4. Performance-avoidance	.00	.31**	.41**	-						
5. Work-avoidance	-.42**	-.35**	-.12*	.19**	-					
6. Consistency of interest	.21**	.12*	-.06	-.17**	-.26**	-				
7. Perseverance of effort	.20**	.14*	-.03	-.06	-.30**	.42**	-			
8. Intrinsic CL	.03	.03	-.02	.04	.04	-.04	.00	-		
9. Extraneous CL	-.09	-.08	-.03	-.01	.18**	-.09	-.12	.48**	-	
10. Germane CL	.12	.23**	.10	.06	-.10	.04	.11	.36**	.22**	-
M	5.51	4.34	3.87	4.46	3.97	3.09	3.41	4.96	2.70	6.07
SD	1.19	1.45	1.38	1.46	1.48	0.77	0.86	1.96	1.57	2.15
Range	1 - 7	1 - 7	1 - 7	1 - 7	1 - 7	1 - 5	1 - 5	1 - 9	1 - 9	1 - 9
α	.84	.83	.72	.77	.74	.71	.77	.92	.81	-

Note. CL = cognitive load.

*p** < .05. *p*** < .01.

Achievement Goal Orientation Profiles

The results of the LPA are reported in Table 2. The results provided support for the three-profile solution. Although AIC and BIC value suggested better fit for four-profile solution, *p*VLMR and *p*LMR were not significant, indicating that adding a fourth group did not result in a meaningful improvement in model fit compared to the three-profile model. Also, entropy value was satisfactory for both models but higher for three-profile solution (.82). The three identified profiles were qualitatively informative and consistent with previous research and theory. Mean differences in achievement goal orientations between the latent goal orientation profile groups based on the Wald test are shown in Table 3.

Table 2

Information Criteria Values for Different Profile Solutions

k	AIC	BIC	p_{VLMR}	p_{LMR}	Entropy	Group sizes
1	4404.520	4441.853	-	-	-	309
2	4184.036	4262.436	.0000	.0000	.73	173, 136
3	4097.389	4216.856	.0066	.0071	.82	167, 57, 85
4	4037.851	4198.385	.0763	.0796	.77	72, 44, 88, 105
5	3995.653	4197.253	.3466	.3508	.78	54, 38, 99, 41, 77

Note. k = number of latent profiles in the model; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion, pVLMR = Vuong-Lo-Mendell-Rubin test likelihood ratio test, pLMR = Lo-Mendell -Rubin test adjusted likelihood ratio test.

The profiles were labelled as *success-oriented* (n = 85, 27.5%), *indifferent* (n = 167, 54.0%), and *work-avoidant* (n = 57, 18.5%) (Figure 1). The success-oriented students showed high emphasis on both mastery- and performance-related orientations, along with relatively low work-avoidance orientation in comparison with other goal orientation profiles. The indifferent students were characterised by goal orientations that were average relative to the other groups. Work-avoidant profile displayed a stronger emphasis on work-avoidance orientation compared to the remaining goal orientation profiles, along with relatively lower performance-approach, performance-avoidance, mastery-intrinsic, and particularly mastery-extrinsic orientation.

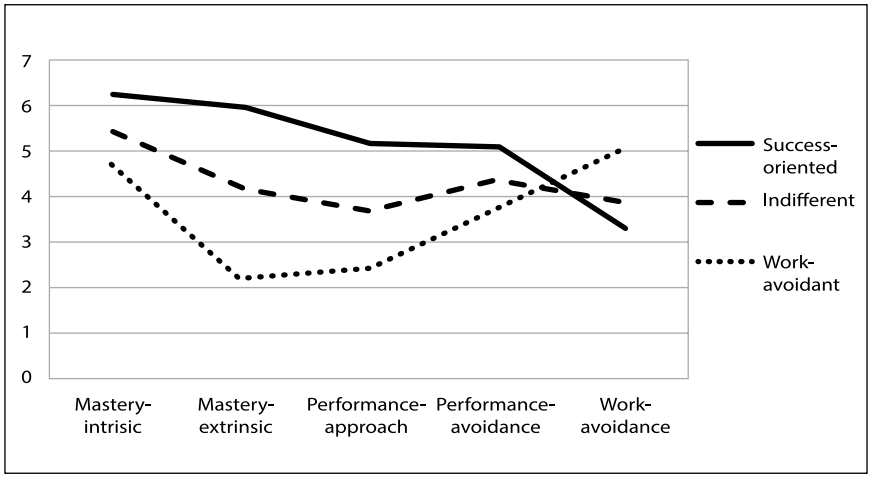
Table 3

Mean Differences in Achievement Goal Orientations Between the Goal Orientation Profiles

Measure	1. Success-oriented		2. Indifferent		3. Work-avoidant		χ^2	*p*	Significant differences[a]
	M	*SE*	*M*	*SE*	*M*	*SE*			
Mastery-intrinsic	6.30	0.07	5.42	0.07	4.60	0.21	107.27	.00	1 > 2 > 3
Mastery-extrinsic	6.05	0.07	4.21	0.05	2.17	0.09	1246.35	.00	1 > 2 > 3
Performance-approach	5.18	0.12	3.67	0.08	2.54	0.16	243.21	.00	1 > 2 > 3
Performance-avoidance	5.11	0.16	4.36	0.09	3.78	0.22	25.79	.00	1 > 2 > 3
Work-avoidance	3.38	0.15	3.89	0.10	5.10	0.20	47.83	.00	3 > 2 > 1

Note. [a] differences between achievement goal orientation profile groups based on the Wald test.

Figure 1
Achievement Goal Orientation Profiles



Comparison of achievement goal orientation profiles with BCH method revealed that the success-oriented students reported higher grit (both consistency of interest and perseverance of effort), as well as greater germane cognitive load compared to the indifferent and work-avoidant profiles. The indifferent and work-avoidant students showed no significant differences in grit, nor in perceived cognitive load. No significant differences were found among the goal orientation profiles in perceptions of intrinsic and extraneous cognitive load (Table 4).

Table 4
Mean Differences in Grit and Cognitive Load Between the Achievement Goal Orientation Profiles

Measure	1. Success-oriented		2. Indifferent		3.Work-avoidant		χ^2	*p*	Significant differences[a]
	M	*SE*	*M*	*SE*	*M*	*SE*			
CI	3.28	0.09	3.04	0.06	2.93	0.12	6.35	.04	1 > 2, 3
PE	3.67	0.11	3.32	0.07	3.27	0.13	8.11	.02	1 > 2, 3
ICL	4.98	0.28	5.09	0.16	4.53	0.35	1.95	.38	-
ECL	2.69	0.20	2.66	0.13	2.83	0.32	0.25	.88	-
GCL	6.72	0.24	5.94	0.18	5.36	0.43	9.89	.01	1 > 2, 3

Note. CI = consistency of interest; PE = perseverance of effort; ICL = intrinsic cognitive load; ECL = extraneous cognitive load; GCL = germane cognitive load.
[a] differences between achievement goal orientation profile groups based on the BCH method

Discussion

The aim of the present study was to examine the relationship between motivational concepts, more specifically, achievement goal orientation profiles and grit, and specific types of cognitive load. In doing so, we employed a person-centred approach to identify achievement goal orientation profiles among students and compare these profiles across grit and cognitive load.

Firstly, we hypothesised that a four-profile solution would best fit the data, as most earlier studies employing Helsinki 5 conceptualisation of achievement goal orientations have obtained four-profile solutions (for overview see Niemivirta at al., 2019). However, LPA indicated that three-profile solution was the most suitable. The largest group, comprising approximately half of the student sample, was identified as *in-*

different students (54%). Consistent with previous findings, these students represent a typical profile with goal orientations that are generally average compared to the other groups. However, their mastery intrinsic and extrinsic orientations, as well as performance-avoidance orientation were above the scale midpoint, while performance-approach and work-avoidance orientations were closer to midpoint. The *success-oriented* profile accounted for about 28% of the sample, making it about half the size of the indifferent group. The students in this profile displayed a strong emphasis on both mastery and performance orientations compared to other profiles, with absolute values on these scales exceeding the midpoints. Their work-avoidance orientation was around the scale midpoint and the lowest among the profiles. Notably, their emphasis on mastery-extrinsic goals was the most pronounced compared to the other groups. The smallest group, representing *work-avoidant* students, comprised about 18% of the sample. These students placed a stronger emphasis on work-avoidance goals than other profiles and reported the highest scores on this scale. In contrast, they demonstrated relatively low performance-approach, mastery-intrinsic, and particularly mastery-extrinsic orientations compared to the other groups. Their performance-approach scores were below the scale midpoint, performance-avoidance scores were around the midpoint, while their mastery-extrinsic orientation was very low. Interestingly, their mastery-intrinsic orientation was above the scale midpoint.

Obtained profiles align with previously identified profiles using LPA (Niemivirta et al., 2019; Pahljina-Reinić et al., 2024) but unexpectedly, mastery-oriented profile was not differentiated. Most of the earlier studies that identified mastery-oriented groups were conducted on younger (secondary schools) students (e.g., Mädamürk et al., 2020; Tuominen et al. 2020; Tuominen-Soini et al., 2012). Some findings suggest that mastery goals may be more beneficial, while performance-approach goals may be more detrimental, during elementary school. At the secondary level, there is evidence supporting the benefits of both mastery and performance-approach goals, while at the undergraduate level, some findings indicate that performance-approach goals may be more advantageous due to higher academic demands and a more competitive context (Linnenbrink-Garcia et al., 2008). However, benefits of mastery orientations have still been acknowledged, and studies show that mastery-oriented profile was also identified in university students, showing adaptive motivational patterns (Pahljina-Reinić, 2022; Pulkka & Niemivirta, 2013, 2015). In a study by Lee et al. (2017), mastery-oriented group was not identified as all groups of university students showed high levels of mastery orientation. In our study, the groups differed slightly in their levels of mastery-intrinsic orientation (which were above scales' midpoints), but the differences were not as pronounced as those in mastery-extrinsic orientation. Mastery-extrinsic orientation reflects orientation on absolute success and can be crucial for success at the university. Niemivirta et al. (2019) argued that such goals are related to adaptive coping strategies and behaviours, such as investment of effort, persistence, and academic success.

To explore whether the profiles identified in the present study will differ in their persistence and effort, more specifically, in their grit and in more situation specific appraisals of cognitive load on midterm exam, BCH approach was used in the last step of LPA. The results showed that, in line with our hypothesis, the success-oriented students were indeed grittier, reporting more consistent interests and higher perseverance of effort. In the present study, no differences were found between indifferent and work-avoidant profiles in terms of grit. Pulkka and Niemivirta (2013) found a similar dominance of success-oriented university students over work-avoidant students (but not over indifferent or mastery-oriented students) in terms of perceived effort and attainment during the course. However, no differences were found between the profiles in reported participation. The same authors partially replicated these findings in the subsequent study (Pulkka & Niemivirta, 2015), showing that the success-oriented group performed significantly better than the avoidance-oriented group in terms of reported effort and attainment. In addition, they found that the success-oriented students reported higher levels of participation in the course. To conclude, the findings indicate that the success-oriented students who strive to achieve both mastery but also academic success, especially in terms of absolute standards, demonstrate greater perseverance and more consistent interests,

which may facilitate the accomplishment of their goals. In contrast, the indifferent and work-avoidant students invest less effort when facing obstacles and exhibit less stable interests.

In the present study, we aimed to explore how these general motivational patterns are manifested in more situation-specific appraisals of students' perceived cognitive load during an exam. As elaborated earlier, GCL seems to be constrained not only by working memory capacity and the nature of the task, but also by the learner's motivation (Schnotz et al., 2009). Indeed, our findings indicate that the success-oriented students not only showed higher levels of grit in general, but they were more prone to invest effort during knowledge assessment situation. Although we expected the (mastery and) success-oriented students would also differ from indifferent and work-avoidant profiles in terms of intrinsic and extrinsic cognitive load, given their tendency to master the material more thoroughly and therefore perceive exam as less complex and more comprehensible, that result was not observed.

All three achievement goal profiles perceived moderate levels of ICL and low levels of ECL, suggesting they perceived exams as moderately complex and appropriately designed, mostly without irrelevant information. Since the students across different goal orientation groups reported similar levels of ICL and ECL, it can be concluded that differences in germane load do not stem primarily from task characteristics (i.e., ICL and ECL) but rather are driven by the students' willingness to utilise their available mental resources. That aligns with assumptions of Schnotz at al. (2009), which suggest that GCL depends on general learning orientations, as well as affective and motivational aspects, since learners determine how much effort they invest in learning or solving a task. Low levels of perceived ECL suggest that sufficient working memory capacity was available to invest into germane processing, yet students were differently inclined to do so.

While reducing ECL to a minimum is proposed as the most effective way to maximise students' germane effort (e.g., Wang & Lajoie, 2023), our findings indicate that motivational tendencies also play a significant role, as students' achievement goal orientations shape their decision to invest effort. A recent study (Wang et al., 2023) identified distinct cognitive load patterns based on the students' perceived intrinsic, extraneous, and germane load, using person-centred approach. Their findings indicate that different combinations of cognitive load types can be recognised, and that those patterns are linked to temporal dynamics of self-regulated learning. Future research should further investigate possible combinations of cognitive load types and their relationships with motivation and self-regulated learning, contributing to the ongoing research trend of integrating cognitive load theory with self-regulated learning (e.g., Seufert, 2020).

Limitations of the Study

Several limitations should be acknowledged that might affect the findings obtained in the present study. The most significant issue concerns the fact that the students were recruited from different study programmes and years of study, leading to variability in midterm exam content and difficulty. Although measures of ICL and ECL demonstrated metric invariance across the student groups, future studies should employ larger and more homogeneous samples. This would allow for a more rigorous assessment of the relations between cognitive load types and motivation. The second issue is related to participant attrition. Almost a quarter of the sample was not present at both measurement points, which impacts the robustness of the findings. Also, the present study brings attention to the need for more precise conceptualisation of cognitive load, as the construct is interpreted in various ways in the literature (see Orru & Longo, 2019; Seufert, 2020; Sweller et al., 2019). Our data could not replicate the original three-factor structure, and thus our findings concerning GCL rely on a single-item measure. Although it is not uncommon to use single-item measures in cognitive load assessment (Schuessler et al., 2025), conclusions drawn from them should be interpreted with caution. Additionally, it would be beneficial to include not only self-reported measures of cognitive load but more objective measures such physiological or eye-tracking techniques. However, such

measures are difficult to implement in natural educational settings, and additional research is needed to evaluate these techniques as valid measures of cognitive load (Sweller et al., 2019).

Despite the limitations, the present study underscores the importance of further investigating the relationships between motivational aspects, such as achievement goal orientations and persistence, and cognitive load, thereby integrating different research traditions.

References

Alhadabi, A., & Karpinski, A. C. (2020). Grit, self-efficacy, achievement orientation goals, and academic performance in university students. *International Journal of Adolescence and Youth*, *25*(1), 519–535. https://doi.org/10.1080/02673843.2019.1679202

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261–271. https://doi.org/10.1037/0022-0663.84.3.261

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*(2), 230-258. https://doi.org/10.1177/0049124192021002005

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464-504. https://doi.org/10.1080/10705510701301834

Cook, D. A., Castillo, R. M., Gas, B., & Artino Jr, A. R. (2017). Measuring achievement goal motivation, mindsets and cognitive load: Validation of three instruments' scores. *Medical Education*, *51*(10), 1061–1074. https://doi.org/10.1111/medu.13405

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087–1101. https://doi.org/10.1037/0022-3514.92.6.1087

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256–273. https://doi.org/10.1037/0033-295X.95.2.256

Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*(3), 501–519. https://doi.org/10.1037/0022-3514.80.3.501

Feldon, D. F., Callan, G., Juth, S., & Jeong, S. (2019) Cognitive load as motivational cost. *Educational Psychology Review, 31*, 319–337. https://doi.org/10.1007/s10648-019-09464-6

Han, C. (2023). Structural relations among achievement goals, perceptions of classroom goals, and grit. *Current Psychology, 42*, 16687–16697. https://doi.org/10.1007/s12144-022-02891-9

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

King, R. B., & McInerney, D. M. (2014). The work avoidance goal construct: Examining its structure, antecedents, and consequences. *Contemporary Educational Psychology*, *39*(1), 42–58. https://doi.org/10.1016/j.cedpsych.2013.12.002.

Klepsch, M., & Seufert, T. (2021). Making an effort versus experiencing load. *Frontiers in Education, 6*, Article 645284. https://doi.org/10.3389/feduc.2021.645284

Korać, D., & Rončević Zubković, B. (2023, November). *Odnos afektivnih i (meta)kognitivnih čimbenika prilikom rješavanja matematičkih problemskih zadataka*. [The relationship between affective and (meta)cognitive factors when solving mathematical problems]. [Conference presentation]. 26. Dani Ramina i Zorana Bujasa, Zagreb, Croatia.

Lee, Y. K., Wormington, S. V., Linnenbrink-Garcia, L., & Roseth, C. J. (2017). A short-term longitudinal study of stability and change in achievement goal profiles. *Learning and Individual Differences, 55*, 49-60. https://doi.org/10.1016/j.lindif.2017.02.002

Linnenbrink-Garcia, L., Tyson, D. F., & Patall, E. A. (2008). When are achievement goal orientations beneficial for academic achievement? A closer look at main effects and moderating factors. *Revue Internationale de Psychologie Sociale, 21*(1-2), 19–70. https://shs.cairn.info/revue-internationale-de-psychologie-sociale-2008-1-page-19?lang=fr

Mädamürk, K., Tuominen, H., Hietajärvi, L., & Salmela-Aro, K. (2021). Adolescent students' digital engage-

ment and achievement goal orientation profiles. *Computers & Education, 161*, Article 104058. https://doi.org/10.1016/j.compedu.2020.104058

Maehr, M. L., & Zusho, A. (2009). Achievement goal theory: The past, present, and future. In K. R. Wenzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 77–104). Routledge/Taylor & Francis Group.

Masyn, K. E. (2013). Latent class analysis and finite mixture modelling. In T. D. Little (Eds.), *The Oxford handbook of quantitative methods, Volume 2: Statistical analysis* (pp. 551–611). Oxford University Press.

McGloin, R., McGillicuddy, K. T., & Christensen, J. L. (2017). The impact of goal achievement orientation on student technology usage in the classroom. *Journal of Computing in Higher Education*, *29*, 240–266. https://doi.org/10.1007/s12528-017-9134-4

Muenks, K., Yang, J. S., & Wigfield, A. (2018). Associations between grit, motivation, and achievement in high school students. *Motivation Science, 4(*2), 158–176. https://doi.org/10.1037/mot0000076

Muthén, L. K., & Muthén, B. O. (1998 – 2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Orru G., & Longo L. (2019). The evolution of cognitive load theory and the measurement of its intrinsic, extraneous and germane loads: A review. In: L. Longo & M. Leva (Eds.) *Human mental workload: Models and applications*. *H-WORKLOAD 2018. Communications in Computer and Information Science*, *vol. 1012*. Springer, Cham. https://doi.org/10.1007/978-3-030-14273-5_3

Nicholls, J. G., Patashnick, M., & Nolen, S. B. (1985). Adolescents' theories of education. *Journal of Educational Psychology, 77*(6), 683–692. https://doi.org/10.1037/0022-0663.77.6.683

Niemivirta, M. (2002). Motivation and performance in context: The influence of goal orientations and instructional setting on situational appraisals and task performance. *Psychologia*, *45*, 250–270. https://doi.org/10.2117/psysoc.2002.250

Niemivirta, M., Pulkka, A.-T., Tapola, A., & Tuominen, H. (2019). Achievement goal orientations: A person-oriented approach. In K. A. Renninger & S. E. Hidi (Eds.), *The Cambridge handbook of motivation and learning* (pp. 566–616). Cambridge University Press. https://doi.org/10.1017/9781316823279.025

Pahljina-Reinić, R. (2022). Profili ciljnih orijentacija i strategije samoregulacije motivacije [Achievement goal orientation profiles and strategies for the self-regulation of motivation]. *Psihologijske teme / Psychological Topics, 31*(3), 721–742. https://doi.org/10.31820/pt.31.3.13

Pahljina-Reinić, R., Rončević Zubković, B., & Kolić-Vehovec, S. (2024). Achievement goal orientation profiles and the experiences of gymnasium students with digital technologies in education. *Psihologijske teme* / *Psychological Topics, 33*(1), 67–89. https://doi.org/10.31820/pt.33.1.4

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92*(3), 544–555. https://doi.org/10.1037/0022-0663.92.3.544

Plass, J. L., & Kalyuga, S. (2019). Four ways of considering emotion in cognitive load theory. *Educational Psychology Review, 31*(2), 339–359. https://doi.org/10.1007/s10648-019-09473-5

Pulkka, A.-T., & Niemivirta, M. (2013). Adult students' achievement goal orientations and evaluations of the learning environment: A person-centred longitudinal analysis. *Educational Research and Evaluation, 1*9(4), 297–322. https://doi.org/10.1080/13803611.2013.767741

Pulkka, A.-T., & Niemivirta, M. (2015). The relationships between adult students' achievement goal orientations, self-defined course goals, course evaluations, and performance. *Journal for Educational Research Online, 7*(3), 28–53. http://www.j-f-b.de/index.php/jero/article/view/609

Sadoughi, M., & Eskandari, N. (2024). The relationship between achievement goal orientations and academic burnout among medical students: The mediating role of academic grit. *Journal of Medical Education Development*, *17*(54), 11–20. https://edujournal.zums.ac.ir/article-1-1886-en.html

Scheiter, K., Ackerman, R., & Hoogerheide, V. (2020). Looking at mental effort appraisals through a metacognitive lens: Are they biased? *Educational Psychology Review, 32*(4), 1003–1027. https://doi.

org/10.1007/s10648-020-09555-9

Schnotz, W., Fries, S., & Horz, H. (2009). Some motivational aspects of cognitive load theory. In M. Wosnitza, S. A. Karabenick, A. Efklides, & P. Nenniger (Eds.), *Contemporary motivation research: From global to local perspectives* (pp. 69–96). Hogrefe & Huber Publishers.

Schuessler, K., Fischer, V., & Walpuski, M. (2025). Investigating construct validity of cognitive load measurement using single-item subjective rating scales. *Instructional Science, 53*, 71–97. https://doi.org/10.1007/s11251-024-09692-6

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions*. Educational Psychologist, 46*(1), 26–47. https://doi.org/10.1080/00461520.2011.538646

Seufert, T. (2020). Building bridges between self-regulation and cognitive load – An invitation for a broad and differentiated attempt. *Educational Psychology Review, 32*(4), 1151–1162. https://doi.org/10.1007/s10648-020-09574-6ž

Sunawan. Sugiyo., & Pranoto, Y. K. S. (2021). Achievement goals and extraneous load predict germane load: The mediating effects of achievement emotions. *Malaysian Journal of Learning and Instruction, 18*(2), 215–234. https://doi.org/10.32890/mjli2021.18.2.8

Sweller, J., Van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, Article 251296. https://doi.org/10.1023/A:1022193728205

Sweller, J., van Merriënboer, J.J.G. & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*, 261–292. https://doi.org/10.1007/s10648-019-09465-5

Tuominen, H., Niemivirta, M., Lonka, K., & Salmela-Aro, K. (2020). Motivation across a transition: Changes in achievement goal orientations and academic well-being from elementary to secondary school. *Learning and Individual Differences, 79,* Article 101854*.* https://doi.org/10.1016/j.lindif.2020.101854

Tuominen-Soini, H., Salmela-Aro, K., & Niemivirta, M. (2011). Stability and change in achievement goal orientations: A person-centered approach. *Contemporary Educational Psychology*, *36*, 82–100. https://doi.org/10.1016/j.cedpsych.2010.08.002

Tuominen-Soini, H., Salmela-Aro, K., & Niemivirta, M. (2012). Achievement goal orientations and academic well-being across the transition to upper secondary education. *Learning and Individual Differences*, *22*, 290–305. https://doi.org/10.1016/j.lindif.2012.01.002

Wang, T., & Lajoie, S. P. (2023). How does cognitive load interact with self-regulated learning? A dynamic and integrative model. *Educational Psychology Review, 35*, Article 69. https://doi.org/10.1007/s10648-023-09794-6

Wang, T., Li, S., Tan, C., Zhang, J., & Lajoie, S. P. (2023). Cognitive load patterns affect temporal dynamics of self-regulated learning behaviors, metacognitive judgments, and learning achievements. *Computers & Education, 207*, Article 104924. https://doi.org/10.1016/j.compedu.2023.104924

Wimmer, S., Lackner, H. K., Papousek, I., & Paechter, M. (2018). Goal orientations and activation of approach versus avoidance motivation while awaiting an achievement situation in the laboratory. *Frontiers in Psychology, 9*, Article 1552. https://doi.org/10.3389/fpsyg.2018.01552

Wormington, S. V., & Linnenbrink-Garcia, L. (2017). A new look at multiple goal pursuit: The promise of a person-centered approach. *Educational Psychology Review*, *29*, 407–445. https://doi.org/10.1007/s10648-016-9358-2

Xu, K. M., Koorn, P., De Koning, B., Skuballa, I. T., Lin, L., Henderikx, M., Marsh H. W., Sweller J., & Paas, F. (2021). A growth mindset lowers perceived cognitive load and improves learning: Integrating motivation to cognitive load. *Journal of Educational Psychology, 113*(6), 1177–1191. https://doi.org/10.1037/edu0000631

Yu, J., & McLellan, R. (2020). Same mindset, different goals and motivational frameworks: Profiles of mindset-based meaning systems. *Contemporary Educational Psychology*, *62*, Article 101901. https://doi.org/10.1016/j.cedpsych.2020.101901

Zrilić, I. (2018). *Kako izvori smisla i životni ciljevi pridonose subjektivnoj i psihološkoj dobrobiti?* [How do sources of meaning and life aspirations contribute to subjective and psychological well-being?] [Diplomski rad, Sveučilište u Rijeci, Filozofski fakultet, Odsjek za psihologiju]. https://urn.nsk.hr/urn:nbn:hr:186:572483

Zu, T., Munsell, J., & Rebello, N. S. (2021). Subjective measure of cognitive load depends on participants' content knowledge level. *Frontiers in Education, 6*, Article 647097. https://doi.org/10.3389/feduc.2021.647097

7

Different Effects of Word Concreteness in a Recall and Recognition Task

Dunja Mićunović
FMK Lab—Laboratory for Psychological Research,
Department of Psychology, Faculty of Media and Communications
dunja.micunovic.20211057@fmk.edu.rs

Anja Lujić
Department of Psychology, FMK Lab—Laboratory for Psychological Research,
Faculty of Media and Communications, Serbia

Senka Vasović
Department of Psychology, FMK Lab—Laboratory for Psychological Research,
Faculty of Media and Communications, Serbia

Marko Janković
Department of Psychology, FMK Lab—Laboratory for Psychological Research,
Faculty of Media and Communications, Serbia

Milica Popović Stijačić
Department of Psychology, FMK Lab—Laboratory for Psychological Research,
Faculty of Media and Communications, Serbia
Laboratory for Experimental Psychology, Faculty of Philosophy,
University of Novi Sad, Serbia

Abstract

Based on Dual Coding Theory, Depth of Processing Theory, and Perceptual Symbols Systems Theory, we examined the relationship between the concreteness effect and depth of processing in memory tasks. We also manipulated the number of sensory modalities associated with concrete words to test the peg hypothesis, which suggests that additional sensory cues enhance recall. The participants were randomly assigned to one of three semantic rating tasks—concreteness, context availability, or vividness—and rated words accordingly. The words were categorised as abstract, few-modalities, or many-modalities based on the number of sensory experiences they evoke. After completing the rating task, all the participants proceeded to a mental rotation task, followed by random

assignment to one of three previously unannounced memory tasks: free recall, cued recall, or recognition. As expected, the vividness group achieved the highest retrieval across all tasks, while the other groups showed no clear pattern. Notably, abstract words were recalled most accurately in the cued recall task, while many-modality concrete words led to the highest recognition accuracy. These findings are discussed using the Dynamic Visual Noise Paradigm, as the mental rotation task may have interfered with the visualisation processes involved in consolidating concrete word representations. Our results do not support the peg hypothesis or the assumptions of Perceptual Symbols Systems Theory since no recognition difference was found between few- and many-modality concrete words.

Keywords: *Dual Coding Theory, Depth of Processing Theory, Embodiment Theory, incidental learning, concreteness effect, memory tasks*

Introduction

The concreteness effect is one of the most extensively studied phenomena in cognitive psychology, yet a definitive explanation for this effect remains elusive. It refers to the advantage that concrete words have over abstract words in word-processing tasks—namely, shorter processing latencies for concrete words and higher accuracy in recall and recognition (Doest & Semin, 2005; Hamilton & Rajaram, 2001; Kounios & Holcomb, 1994; Paivio, 1991; Paivio et al., 1994; Popović Stijačić & Filipović Đurđević, 2015, 2022; Romani et al., 2008; Schwanenflugel & Stowe, 1989; Schwanenflugel et al., 1992; Vandendaele & Grainger, 2022). The two most prominent theoretical accounts are the *Dual Coding Theory* (DCT; Paivio, 1991) and the *Context Availability Theory* (Schwanenflugel et al., 1992).

Theoretical Explanations of Concreteness Effect

The DCT (Paivio, 1991) postulates that concrete words have a memory advantage because they are encoded through additional sensory information besides the symbolic code they share with abstract words. These analogue codes serve as extra cues (i.e., pegs) during retrieval. This assumption is known as the *peg hypothesis* (Paivio, 1991, 1994). The theory emphasises the role of visual and other sensory components in word processing and recall, suggesting that concrete words are more easily remembered than abstract words due to their specificity and ease of visualisation. Over the past several decades, the importance of sensory-motor information related to word meaning has been highlighted by the Perceptual Symbol Systems Theory—one of the main approaches within embodiment theories—which postulates that our knowledge is grounded in the sensory-motor system (Barsalou, 1999, 2004). The Dual Coding Theory and the Perceptual Symbol Systems Theory are complementary, as the latter decomposes sensory-motor experience and distinguishes between words not based on general concreteness but through fine-grained measures such as the number of sensory modalities (NoM) through which a word's meaning can be experienced (Filipović Đurđević et al., 2016; Popović Stijačić & Filipović Đurđević, 2015, 2022).

Previous research on the Serbian language (Popović Stijačić & Filipović Đurđević, 2015, 2022) has tested the extension of the peg hypothesis in relation to the Dual Coding Theory (DCT) and Perceptual Symbol Systems Theory, examining both free and cued recall (Paivio, 1991; Paivio et al., 1994). According to this hypothesis, analogue codes related to concrete words serve as additional cues during retrieval. Merging the ideas of both DCT and Perceptual Symbol Systems Theory, the authors hypothesised that words whose meanings can be experienced through more sensory modalities would show better recall accuracy. As predicted, reproduction accuracy increased with the number of sensory modalities through which the meaning of a word could be experienced. More specifically, the lowest reproduction accuracy was observed for abstract words (e.g., "truth"), which cannot be experienced through the senses, followed by concrete words whose meanings can be experienced through one or two senses (e.g., "moon"). The highest reproduction accuracy was recorded for concrete words whose meanings can be experienced through more than two senses (e.g., "fire"). In other words, a greater number of sensory modalities through which a word's meaning can be experienced serves as a greater number of "pegs" or additional cues during retrieval. Still, this effect was present only in the cued recall task, not in the free recall task.

Contrary to DCT and Perceptual Symbol Systems, the Context Availability Theory (Schwanenflugel & Stowe, 1989) posits that the advantage of concrete words stems from their richer contextual information, which is more readily available than that of abstract words during word processing and recall. Namely, Schwanenfluegel et al. (1992) found that the concreteness effect is the consequence of strategic imagery usage during encoding and retrieval, and that depends on individual differences. Thus,

the individuals who reported using imagery during encoding and retrieval had greater recall accuracy of the concrete words.

Depth of Processing and Recall

Concerning recall accuracy in general, the encoding phase of words plays a crucial role. With this in mind, we draw on the assumptions of the Depth of Processing (DoP) Theory (Craik & Lockhart, 1972), according to which words are processed at varying levels of depth, from shallow sensory processing to deeper semantic processing. According to DoP, deeper processing involves attentional mechanisms oriented towards the meaning of stimuli, resulting in more durable memory traces. The deeper the processing, the better the recall or reproduction of the word.

By integrating multiple theoretical perspectives, including Dual Coding Theory (DCT), Perceptual Symbol Systems, Context Availability, and the Depth of Processing (DoP) framework, we aim to understand better the role of incidental learning in retrieving concrete and abstract words. Specifically, by directing attention to different semantic properties of words, we hypothesised that it is possible to manipulate the depth of processing during incidental encoding, a method similar to that used by West and Holcomb (2000).

For example, we examined whether the concreteness effect could be enhanced if participants had previously rated words based on concreteness or vividness of mental imagery rather than context availability. Although all three types of ratings are semantic, both concreteness and vividness of mental imagery are more likely to engage additional sensory memory cues. Consequently, these assessments may require greater cognitive effort and time, leading to more thorough processing and improved recall.

West and Holcomb (2000) applied a similar procedure in a sentence verification task. They found that the concreteness effect was strongest when participants made semantic judgments about the sentence, somewhat reduced when they performed imagery-based judgments, and absent in a shallow task (e.g., "Is the letter *n* present in the target word?").

Study Aims and Hypotheses

Our study had two primary aims. (1) To investigate the concreteness effect across different memory tasks (free recall, cued recall, and recognition), using a definition of concreteness based on the number of sensory modalities through which a concept can be experienced. This variation allowed us to test the peg hypothesis more directly by comparing two groups of concrete words differing in their number of potential sensory "pegs": words associated with one or two modalities ("few-modality concrete") versus those linked to three or more modalities ("many-modality concrete").

(2) To examine how the different aspects of semantic processing influence memory performance as an extension of the Depth of Processing (DoP) Theory. Specifically, we investigated whether directing attention to various semantic properties of words during encoding, namely concreteness, context availability, and vividness of mental imagery, would enhance the recall of abstract and concrete words differently. Depth of processing was additionally explored by the time the participants spent rating words during the encoding, assuming that more time reflects deeper processing.

We hypothesised that the main effect of the memory task would be such that recognition would yield the highest retrieval accuracy, followed by cued recall, with the lowest accuracy expected in free recall.

Additionally, based on previous research, we expected the main effect of the number of sensory modalities associated with word meanings. Specifically, we predicted the highest accuracy for concrete words representing concepts that can be experienced through many modalities, somewhat lower accuracy

for concrete words representing concepts experienced through fewer modalities, and the lowest accuracy for abstract words.

We also examined how different types of semantic rating tasks—used during the encoding phase—might influence memory performance by manipulating the depth of processing. We expected vividness ratings to be the most cognitively demanding, as they require participants to generate mental images of word meanings and evaluate their vividness. The concreteness ratings were assumed to be less demanding, as they involve evaluating whether the concept can be experienced through the senses without necessarily generating a mental image. The context availability ratings, while similarly effortful, were thought to involve a different attentional focus—directed more towards referential and situational knowledge than sensory experience.

Based on this, we predicted that the concreteness effect would be more substantial when participants rated words for vividness or concreteness and weaker when they rated context availability due to differences in the nature and depth of semantic processing.

To test these assumptions, we conducted an experiment in which the participants were randomly assigned to one of three semantic assessment groups: concreteness, context availability, or vividness. The participants rated a list of words in each group according to the assigned dimension. The words varied by the number of sensory modalities through which their meanings can be experienced and were categorised as abstract, few-modalities, or many-modalities words. After completing the rating task, the participants were again randomly assigned to one of three unannounced memory tasks: free recall, cued recall, or recognition. This design allowed us to examine the interaction between the type of semantic processing during encoding and the type of memory task during retrieval.

Method

Participants

A total of 273 psychology students from the Faculty of Media and Communications at Singidunum University in Belgrade participated in the study (85.7% female), with a mean age of 22.78 **(**SD **=** 5.75). The participants were randomly assigned to one of three incidental learning groups, with group sizes as follows: n (G1) = 47, n (G2) = 107, and n (G3) = 46. Table 1 presents the distribution of the participants across the assessment and retrieval groups.

Table 1
Number of participants per assessment and retrieval groups

		Retrieval task			Total
		Free recall	Cued recall	Recognition	
Assessment group	Concreteness	23	23	26	72
	Context availability	83	24	24	131
	Vividness	23	24	23	70
Total		129	71	73	273

Stimuli

The list comprised 45 target Serbian nouns selected from a normative study (Popović Stijačić, 2021), including 15 abstract nouns (e.g., *custom*) and 30 concrete nouns. The concrete nouns were divided

into two groups: 15 nouns referred to concepts that can be experienced through one or two sensory modalities (*few-modality words*, e.g., *jaw*), and 15 referred to concepts that can be experienced through more than two modalities (*many-modality words*, e.g., *bell*), thereby operationalising the Number of Modalities (NoM) factor. All three noun groups were matched on a wide range of psycholinguistic variables, including objective word frequency (Kostić, 1999), familiarity (subjective frequency), context availability, emotional valence, arousal, and age of acquisition (Popović Stijačić, 2021). The two concrete noun groups were also matched for concreteness and imageability. This careful matching ensured that any observed differences in retrieval accuracy could be specifically attributed to the perceptual richness of the words, that is, the number of sensory modalities (NoM) through which a concept can be experienced. Four filler words were added at both the beginning and end of the list to control for primacy and recency effects (Murdock, 1962; Glanzer & Cunitz, 1966). Descriptive statistics for the psycholinguistic variables used to match the three NoM word groups are presented in Table 2.

Table 2
Descriptive statistics and F test for the target words by number of modalities factor

		N	Mean	SD	*F*-test	*p*-value
logF	Abstract	15	4.113	1.363		
	Few-modalities	15	3.476	1.445	2.094	0.136
	Many-modalities	15	2.957	1.806		
Length	Abstract	15	6.067	1.335		
	Few-modalities	15	6.400	1.404	0.213	0.809
	Many-modalities	15	6.333	1.676		
Imageability	Abstract	15	3.866	1.211		
	Few-modalities	15	6.317	0.452	50.512	0.000
	Many-modalities	15	6.362	0.381		
Context availability	Abstract	15	5.508	0.527		
	Few-modalities	15	5.448	0.422	0.258	0.774
	Many-modalities	15	5.591	0.668		
EV	Abstract	15	4.675	1.086		
	Few-modalities	15	4.016	0.754	2.301	0.113
	Many-modalities	15	4.217	0.692		
Familiarity	Abstract	15	5.405	0.994		
	Few-modalities	15	5.167	0.629	0.454	0.638
	Many-modalities	15	5.403	0.695		
Concreteness	Abstract	15	3.077	0.844		
	Few-modalities	15	6.079	0.374	124.408	0.000
	Many-modalities	15	6.044	0.471		

Note. logF = logarithm of objective frequency; length = number of letters; EV = emotional valence; N = number of words per group; SD = standard deviation.

In the cued recall task, an additional set of 45 cue words was selected to accompany the target words during retrieval. These cues were either associatively related (based on experiential relatedness, e.g., *weight—scale*) or semantically related (belonging to the same conceptual category, e.g., *lie—truth*) to the target words. The complete list of target—cue pairs is provided in the Appendix (Table A). The same cue words were also used as distractors in the subsequent recognition task.

Procedure

The entire experiment was administered online using the SoSci Survey (Leiner, 2023). Figure 1 illustrates the organisation and timeline of the experiment. Initially, the participants provided informed consent. Since the experiment included unannounced memory tasks, the whole purpose of the study was disclosed to the participants during the debriefing, which took place after they had completed the experiment.

Incidental learning phase:

The participants were randomly assigned to one of three groups based on the type of incidental learning tasks. These tasks involved estimating either word concreteness, context availability, or the vividness of mental images evoked by the words. The words were presented simultaneously as a list, with rating scales placed next to each word. Above the list, the participants saw instructions corresponding to their assigned task. For concreteness, the participants rated the extent to which a word denotes something that can be experienced through the senses. For context availability, the participants rated how easily a meaningful context could be evoked for the word. To assess the vividness of mental images, the participants were asked to imagine the word and rate the vividness of the mental image that was produced. No mention was made of the upcoming recall tasks.

Mental rotation task:

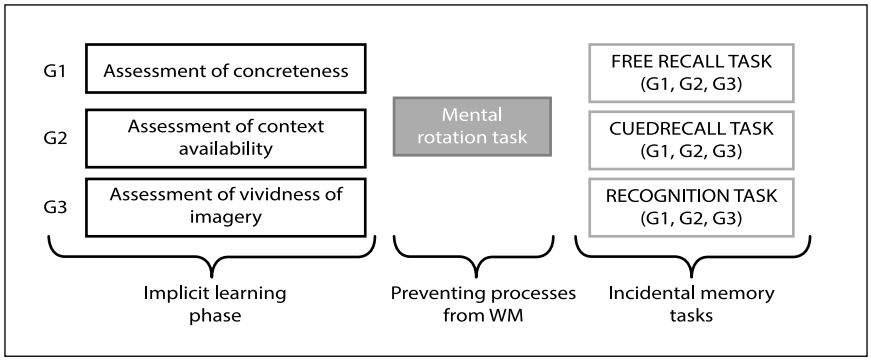
After completing the incidental learning phase, the participants performed a mental rotation task. This task was designed to increase retrieval difficulty and minimise the contribution of working memory. It included 96 experimental trials, preceded by eight practice trials. Two geometric figures were presented on the screen, with the figure on the right rotated at one of four angles (0°, 50°, 100°, or 150°). The participants had to mentally rotate the figure and judge whether it matched the figure on the left. If the figures matched, they pressed the "V" key; if not, they pressed "M". The average time for this task was 6.1 minutes (SD = 3.23 minutes).

Incidental retrieval phase:

The participants were then randomly assigned to one of three retrieval tasks. In free recall, the participants were asked to recall as many words as possible from the incidental learning phase (i.e., the rating task) within a five-minute time limit. In cued recall, the participants were given associatively or semantically related cues to the target words to help them recall the previously estimated words. Reproduction in this task was also limited to five minutes.

Figure 1

Research flow of the experiment



In the recognition task, the participants were presented with a random sequence of previously seen ("old") and unseen ("new") words. Each trial began with a fixation cross (500ms), and words remained on the screen until the participant responded by pressing "V" for old words and "M" for new words.

Variables

Although all three memory tasks had the same stimuli and design, we analysed the answers from free and cued recall separately from the recognition answers.

We employed a 3x3x3 mixed-factor design, with the incidental learning group and memory task as between-participants and within-stimuli factors and the Number of Modalities (NoM) as a between-stimuli and within-participants factor. The incidental learning group consisted of three levels: assessment of concreteness, context availability, and vividness of mental imagery. The memory task consisted of two levels: free recall and cued recall. Finally, NoM had three levels: abstract, few-modality, and many-modality nouns.

The dependent variable was the proportion of correct reproductions in free and cued recall, averaged across the participants. In the recognition task analysis, we used the proportion of correctly recognised targets (hits) corrected for false alarms.

Additionally, we measured the time spent in the encoding phase (i.e., during the assessment of the words) to test the assumptions of the DoP theory. We assumed that more time spent on word assessment reflected deeper processing of the stimuli.

Statistical Analysis

Since we manipulated three factors, we conducted a three-way mixed-effects ANOVA for the by-participant analysis of the recall data, and a two-way mixed-effects ANOVA was applied for the recognition data. We performed a one-way ANOVA for independent samples to test the DoP assumptions regarding the time spent in the incidental learning phase. Data were analysed using SPSS for Windows, version 20 (IBM Corp., 2011).

Results and Discussion

Two participants from the recognition task were excluded from further analysis, as their accuracy, after correcting hits for false alarms, was below zero, indicating random responses.

First, we examined the relationship between encoding time and assessment type to investigate the assumption that vividness ratings would necessitate the most extensive word processing. We conducted a one-way ANOVA with assessment type as the independent variable and encoding time (i.e., the time spent on ratings during the incidental learning phase) as the dependent variable. A significant main effect of assessment type was found: $F(2, 268) = 13.70$, $p < .001$, $\eta^2 = .093$. Table 3 presents the mean encoding times across different assessment and memory tasks. Although vividness ratings required the most time compared to the other assessment groups, post hoc tests with Bonferroni correction revealed significant differences between vividness and context availability ratings (Mdiff = 126.89, $p < .001$), as well as between concreteness and context availability ratings (Mdiff = 95.81, $p < .001$). These results are not entirely consistent with the hypothesis that vividness ratings would be the most demanding compared to both concreteness and context availability ratings.

Table 3
Mean encoding time across assessment group and memory tasks

Assessment group		Mean	SD	N
Concreteness	Free recall	297.35	160.49	23
	Cued recall	292.43	136.68	23
	Recognition	339.24	191.51	25
	Total	310.51	164.32	71
Context availability	Free recall	188.75	76.90	83
	Cued recall	252.08	165.37	24
	Recognition	269.35	138.97	23
	Total	214.70	114.69	130
Vividness	Free recall	350.13	275.17	23
	Cued recall	272.17	131.05	24
	Recognition	405.48	347.59	23
	Total	341.59	267.30	70

Analysis of Recall Data

The three-way mixed-effects ANOVA revealed a significant main effect of the recall task: $F(1, 194) = 16.053$, $p < .001$, $\eta^2 = .076$. As hypothesised, retrieval accuracy was higher in the cued recall condition. Table 4 presents retrieval accuracy across the different memory tasks.

Table 4
Mean retrieval accuracy in recall tasks

	Mean	SE	N
Free recall	.193	.12	129
Cued recall	.264	.13	71

Note. SE = standard error.

Furthermore, a main effect of the assessment group was observed: $F(2, 194) = 4.813$, $p < .01$, $\eta^2 = .047$. Table 5 presents the proportion of retrieval accuracy across the assessment groups. Consistent with our assumptions, the highest recall was achieved in the vividness group, while the concreteness and context availability groups showed similar levels of accuracy (Mdiff = –.01, $p = 1.00$). Post hoc analysis with the Bonferroni correction revealed that the vividness group had significantly better accuracy than both the concreteness group (Mdiff = .066, $p < .01$) and the context availability group (Mdiff = .055, $p < .01$).

Table 5
Mean retrieval accuracy across the assessment groups

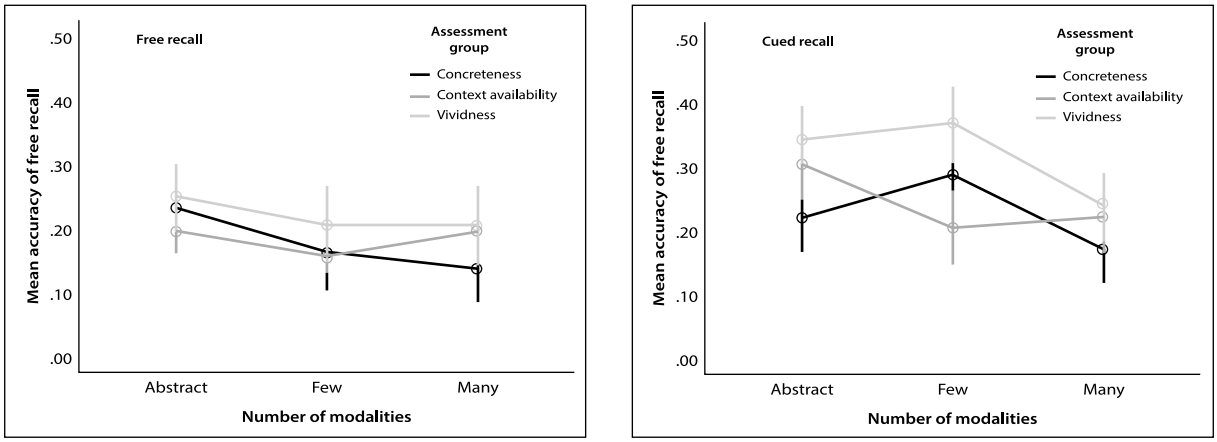
	Mean	SE	N
Concreteness	.203	.017	46
Context availability	.214	.013	107
Vividness	.269	.016	47

Note. SE = standard error.

In contrast to previous research (Paivio et al., 1994; Popović Stijačić & Filipović Đurđević, 2015, 2022), the highest recall accuracy was obtained for abstract nouns (M = .259 ± .14), then for few-modality nouns (M = 20.93 ± .16), and eventually for many-modalities (M = 19.6 ± .14), meaning reversed NoM effect: $F(2, 388) = 16.13$, $p < .001$, $\eta^2 = .076$. The Post Hoc analysis with Bonferroni correction revealed that abstract words significantly differed from few- (Mdiff = .282, $p = .044$) and many-modalities nouns (Mdiff = .626, $p < .001$), where few-modalities also significantly differ from many-modalities nouns (Mdiff = .345, $p = .005$). We also recorded the NoM and assessment task: $F(4, 388) = 4.85$, $p < .001$, $\eta^2 = .048$, where a similar pattern of recall, a reversed NoM effect, was observed for the vividness and concreteness assessment groups, and a different pattern for the context availability group. The interaction of NoM and recall task reached statistical significance: $F(4, 388) = 6.35$, $p = .002$, $\eta^2 = .032$: a reversed NoM effect was present in the cued but not in the free recall task. A significant triple interaction was also obtained: $F(4,388) = 4.495$, $p = .001$, $\eta^2 = .044$ (Figure 2). Post Hoc tests with Bonferroni correction showed that none of the differences, regardless of the assessment group, were significant in free recall. In cued recall, the abstract words in the vividness assessment group had a higher recall than the many-modalities words (Mdiff = .108, $p = .020$). Furthermore, in the same task, two groups of concrete words differed, resulting in the lowest recall accuracy for many-modalities groups (Mdiff = .131, $p < .001$). In the concreteness assessment group, only two groups of concrete words differed, again in the same direction as in the vividness group (Mdiff = .11, $p = .003$), and the abstract words did not differ from either of the concrete word groups. There were no differences in the cued recall accuracy between word groups in the context availability group.

Figure 2

Interaction of NoM and assessment group in free recall (left panel) and cued recall (right panel)



Note. The vertical lines denote 95% confidence intervals of the group means.

The depth of processing assumptions concerning recall accuracy was partially confirmed, as the vividness assessment group achieved better reproduction than the other two assessment groups, and the concreteness and context availability groups had a similar recall rate. However, comparing the assessment time, the vividness and concreteness groups did not significantly differ, indicating that the processing time cannot entirely explain the depth of processing and that some other processes are involved, such as intentional mental visualisation during vividness assessment (the visualisation was part of the instruction).

The reversed NoM effect implies that the participants retrieved abstract words with less effort, although by definition, they are relying only on symbolic code and associative semantic networks. Further, the fine grain division of concrete words showed that, contrary to previous findings (Popović Stijačić & Filipović Đurđević, 2015; 2022), words denoting concepts with richer perceptual experience (many-modalities nouns) had the lowest recall accuracy. These results do not support the peg hypothesis (Paivio, 1991; Paivio et al., 1994) and Perceptual Symbol Systems Theory (Barsalou, 1999), as it was expected that a greater number of sensory modality codes would enhance recall.

Analysis of Recognition Data

In the recognition task, the effect of the assessment group was not statistically significant: $F(2,70) = 2.75$, $p = .07$, $\eta^2 = .075$. Consistent with recall tasks, the most accurate recognition was recorded for the vividness group, while the other two groups achieved similar accuracy (Table 7). Contrary to recall tasks, we recorded the NoM effect: $F(2,67) = 9.92$, $p < .001$, $\eta^2 = .228$, where the abstract words had significantly lower hit rates (corrected for false alarms) compared to many-modalities (*Mdiff* = -.058, $p = .008$) and compared to few-modalities nouns (*Mdiff* = -.074, $p < .001$). Table 7 presents recognition accuracy across the NoM factor. There was no significant difference between the two groups of concrete nouns. No significant interaction was found between the assessment group and the NoM factor: $F(4,136) = .655$, $p = .624$, $\eta^2 = .019$. Figure 3 depicts the recognition accuracy of different word categories and assessment groups.

Table 6
Recognition accuracy (hits corrected for false alarms) across the assessment groups

	Mean	SE	N
Concreteness	.709	.034	25
Context availability	.719	.035	23
Vividness	.814	.035	23

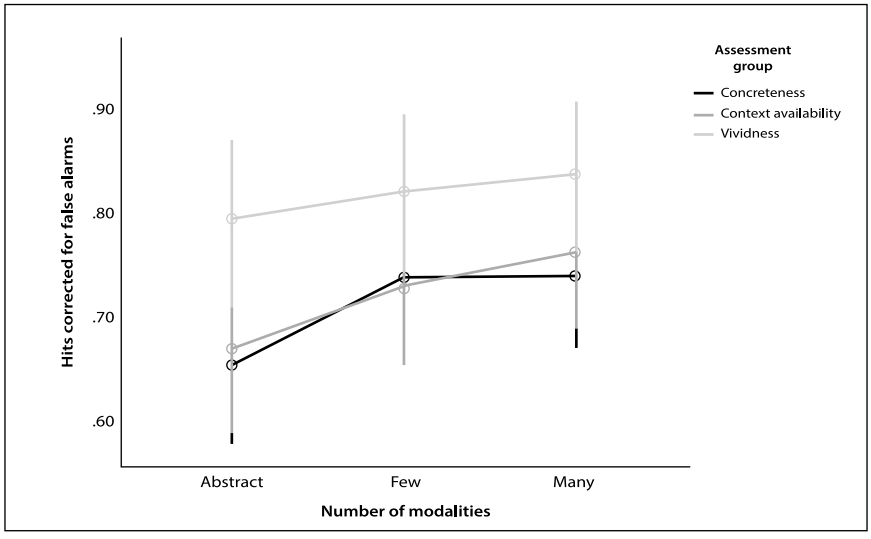
Note. SE = standard error.

Table 7
Recognition accuracy (hits corrected for false alarms) by the NoM word groups

	Mean	SE	N
Concreteness	.709	.034	25
Context availability	.719	.035	23
Vividness	.814	.035	23

Note. SE = standard error.

Figure 3
The interaction of assessment group and NoM in recognition task.



Note. The vertical lines denote 95% confidence intervals of the group means.

Discussion and Conclusion

Our study took a unique approach to explore the concreteness effect (e.g., Kounios & Holcomb, 1994; Paivio, 1991; Paivio et al., 1994; Popović Stijačić & Filipović Đurđević, 2015, 2022) in different memory tasks: free recall, cued recall, and recognition tasks. By varying the incidental encoding phase, we aimed to investigate the Depth of Processing Theory (Craik & Lockhart, 1972) and its influence on the retrieval of abstract and concrete words divided into those whose meaning can be experienced with few sensory modalities and those that can be experienced with many sensory modalities.

We partially confirmed the assumptions concerning the DoP Theory (Craik & Lockhart, 1972). As hypothesised, the participants spent most of the encoding time estimating the vividness of mental images of target words, which was followed by the highest recall and recognition accuracy. However, the encoding time could not explain the better reproduction and recognition for the concreteness assessment group compared to the context availability group since there was no difference in encoding time between these two groups. This result implies that the depth of processing cannot be solely operationalised via encoding time. Still, as the recall was consistently higher for the vividness group, we may conclude that assessing vividness requires additional mental effort compared to assessing concreteness and context availability since the participants were instructed to create mental images of a given word.

The most prominent finding was the opposite direction of the concreteness effect in the dependency of different memory tasks. In this study, we extended the traditional definition of concreteness and operationalised perceptual richness as the number of sensory modalities through which a concept can be experienced. We divided concrete words into two groups: the first group contained nouns that refer to concepts that can be experienced with few senses, and the other contained concepts experienced with many senses. Such manipulation enabled us to test the peg hypothesis (Paivio, 1991; Paivio et al., 1994), extended with the assumptions of the Perceptual Symbol Systems Theory (Barsalou, 1999, 2002). Our results contrast with the peg hypothesis, considering that both groups of concrete words were averaged for imageability and concreteness; thus, the groups differed only in perceptual richness, which did not enhance the reproduction of many-modalities words.

Some researchers (e.g., Anderson & Peterson, 2022; Parker & Dagnall, 2009) have hypothesised that Visual Working Memory (VWM) is engaged during the processing of concrete words due to the mental visualisation processes typically associated with these words. To test this assumption, they applied a Dynamic Visual Noise (DVN) paradigm while participants learned lists of concrete and abstract words during the encoding phase. DVN involves a continuously changing visual pattern that participants view while simultaneously listening to the words they are asked to remember (Anderson & Peterson, 2022; Parker & Dagnall, 2009).

In studies employing this paradigm, the authors reported a reversed concreteness effect, that is, inconsistent performance across memory tasks (e.g., free recall and recognition), attributed to the presence of DVN. They concluded that visual interference impaired participants' ability to utilise visual working memory (utiliseVWM), thereby disrupting the encoding of concrete words. In contrast, the encoding of abstract words, which rely less on visual imagery, remained largely unaffected by DVN (Anderson & Peterson, 2022; Parker & Dagnall, 2009).

The DVN paradigm can partially explain our results, as abstract words showed the highest recall accuracy but only in free recall. In cued recall, equal recall accuracy was achieved for the abstract and few-modalities words, compared to many-modalities nouns. On the other hand, in the recognition task, the hit rate (corrected for false alarms) was highest for the many-modalities word group. Therefore, the concreteness effect remained intact. One possible explanation is that concrete words are more distinct than abstract ones, leading to better recognition accuracy. This explanation is grounded in a recent study which found that con-

crete words exhibit properties of *typical word form*, making them distinct at the word form level (*concreteness form typicality*) and easier to process in various word processing tasks (Kearney et al., 2024). However, we did not manipulate this variable; thus, future studies should explore this explanation. Furthermore, as we did not include a control group (i.e., a group without the mental rotation task), we cannot unequivocally conclude that the mental rotation task prevented mental visualisation during the consolidation of memory traces.

The novelty of this study lies in the fine grain of the word concreteness. Namely, none of the previously mentioned studies varied the perceptual richness of words or investigated the possible influence of additional perceptual cues related to word meanings as additional signs during recall or recognition. Furthermore, the advantages of our research included incidental learning, which allowed us to control the strategic imagery during the encoding. Further, we loaded VWM during the memory consolidation phase, while, in previous studies, such manipulation was applied during the encoding or retrieval phase.

To summarise, our study demonstrated the reversed concreteness effect in incidental recall tasks under the load of Visual Working Memory during memory consolidation, as well as the concreteness effect in recognition tasks under the same conditions. By manipulating the perceptual richness of concrete nouns, we can reject the peg hypothesis, as additional perceptual cues do not enhance the retrieval of concepts that can be experienced through more than two senses. Future studies should include a control group in the design, match the word form typicality of abstract and concrete words, and examine different working memory loads.

Acknowledgements

We would like to express our sincere gratitude to the Faculty of Media and Communications for providing us with the necessary support and to the participants for taking part in this research.

References

Anderson, B. A., & Peterson, S. A. (2022). Concreteness and levels of processing: a test of the dual-coding hypothesis using dynamic visual noise. *Memory*, *30*(10), 1405–1420. https://doi.org/10.1080/09658211.2022.2120622

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671-684. https://doi.org/10.1016/S0022-5371(72)80001-X

Doest, L. T., & Semin, G. (2005). Retrieval contexts and the concreteness effect: Dissociations in memory for concrete and abstract words. *European Journal of Cognitive Psychology*, *17*(6), 859–881. https://doi.org/10.1080/09541440540000031

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5*(4), 351-360. https://doi.org/10.1016/S0022-5371(66)80044-0

Hamilton, M., & Rajaram, S. (2001). The concreteness effect in implicit and explicit memory tests. *Journal of Memory and Language, 44*(1), 96–117. https://doi.org/10.1006/jmla.2000.2749

IBM Corp. Released 2011. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.

Kearney, E., McMahon, K. L., Guenther, F., Arciuli, J., & De Zubicaray, G. I. (2024). Revisiting the concreteness effect: Non-arbitrary mappings between form and concreteness of English words influence lexical processing. *Cognition*, *254*, 105972. https://doi.org/10.1016/j.cognition.2024.105972

Kostić, Đ. (1999). Frekvencijski rečnik savremenog srpskog jezika. Tom I – VII. Institut za eksperimentalnu fonetiku I patologiju govora, Beograd i Laboratorija za eksperimentalnu psihologiju Filozofskog fakulteta u Beogradu.

Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 804. https://doi.org/10.1037/0278-7393.20.4.804

Leiner, D. J. (2023). SoSci Survey [Computer software]. https://www.soscisurvey.de

Marks D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*(1), 17–24. https://doi.org/10.1111/j.2044-8295.1973.tb01322.x

Murdock, B. B., Jr. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64*(5), 482–488. https://doi.org/10.1037/h0045106

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *45*(3), 255. https://doi.org/10.1037/h0084295

Paivio, A., Walsh, M., & Bons, T.A. (1994). Concreteness effects on memory: when and why? *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 1196–1204. https://psycnet.apa.org/doi/10.1037/0278-7393.20.5.1196

Parker, A., & Dagnall, N. (2009). Concreteness effects revisited: the influence of dynamic visual noise on memory for concrete and abstract words. *Memory*, *17*(4), 397–410. https://doi.org/10.1080/09658210902802967

Popović-Stijačić, M., & Filipović Đurđević, D. (2015). Number of sensory modalities through which a concept can be experienced: effect on recall. *Primenjena psihologija*, *8*(3), 335-352. https://doi.org/10.19090/pp.2015.3.335-352

Popović Stijačić, M. (2021). The relevance of perceptual information in word processing – the perspective of embodied cognition [University of Novi Sad]. University of Novi Sad. https://www.cris.uns.ac.rs/DownloadFileServlet/Disertacija161762163851549.pdf?controlNumber=(BISIS)117683&fileName=161762163851549.pdf&id=17699&licenseAccepted=true

Popović Stijačić, M., & Filipović Đurđević, D. (2022). Perceptual richness and its role in free and cued recall. *Primenjena psihologija*, *15*, 255-381. http://dx.doi.org/10.19090/pp.v15i3.2400

Romani, C., Mcalpine, S., & Martin, R. C. (2008). Concreteness effects in different tasks: Implications for models of short-term memory. *Quarterly Journal of Experimental Psychology*, *61*(2), 292–323. https://doi.org/10.1080/17470210601147747

Schwanenflugel, P. J. (1991). Contextual constraint and lexical processing. *Advances in psychology*, *77*, 23-45. https://doi.org/10.1016/S0166-4115(08)61528-9

Schwanenflugel, P.J., Akin, C. & Luh, WM (1992). Context availability and the recall of abstract and concrete words. *Memory & Cognition, 20*, 96–104. https://doi.org/10.3758/BF03208259

Schwanenflugel, P. J. & Stowe, R. W. (1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly*, *24*(1), 114-126.https://psycnet.apa.org/doi/10.2307/748013

Vandendaele, A., & Grainger, J. (2022). Now you see it, now you don't: Flanker presence induces the word concreteness effect. *Cognition*, *218*, 104945. https://doi.org/10.1016/j.cognition.2021.104945

West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *Journal of Cognitive Neuroscience, 12*(6), 1024–1037. https://doi.org/10.1162/08989290051137558

APPENDIX

Table A

The list of target and cue stimuli

Serbian		English translation	
Target	**Cue**	**Target**	**Cue**
DOSADA	avantura	boredom	adventure
PEDALA	bicikl	pedal	bicycle
SLIČNOST	bliskost	similarity	closeness
KLJUČ	brava	key	lock
PEPELJARA	dim	ashtray	smoke
LIPA	drvo	linden (tree)	tree
MRAVINJAK	gužva	anthill	crowd/jam
MUVA	insekt	fly	insect
LAŽ	istina	lie	truth
DIVOTA	krasota	delight	beauty/splendour
DAVANJE	krv	giving	blood
ANĐEO	lepota	angel	beauty
KARMIN	lice	lipstick	face
VOĐA	lider	leader	leader
POKLOPAC	lonac	lid	pot
MODRICA	masnica	bruise	bruise
POŠTENJE	moral	honesty	morality
OSTRVO	more	island	sea
LOPATA	motika	shovel	hoe (garden tool)
GENIJE	naučnik	genius	scientist
KOŠNICA	pčela	beehive	bee
MEDUZA	pipak	jellyfish	tentacle
BUNDEVA	pita	pumpkin	pie
PESAK	plaža	sand	beach
JELKA	poklon	fir tree / Christmas tree	gift
PONUDA	potražnja	offer	demand
PAUČINA	prašina	cobweb	dust
OBIČAJ	praznik	custom	holiday
VILICA	proteza	jaw/fork (depending on context)	denture/prosthesis
VERNIK	religija	believer	religion
LUSTER	sijalica	chandelier	light bulb
SLAMČICA	sok	straw (drinking)	juice
BANDERA	stub	pole	pillar/pole
SVETLOST	sunce	light	sun
SUMRAK	suton	twilight	dusk/twilight
TEG	vaga	weight	scale/balance
PLJUVAČKA	varenje	saliva	digestion

PROLEĆE	vesnik	spring	herald/messenger
MAHOVINA	vlaga	moss	moisture/humidity
ARMIJA	vojska	army	army
SVEĆA	vosak	candle	wax
ZVONCE	vrata	bell	door(s)
DOBROTA	vrlina	kindness	virtue
ROLETNA	zavesa	blind / roller blind	curtain
KOPNO	zemlja	land	earth/land/soil

8

The Tripartite Model of Students' Basic Psychological Needs—Relationship with Teachers' (De)Motivating Styles

Helena Tomšek
Ruđer Bošković Techincal School, Vinkovci, Croatia
htomsek@gmail.com

Aleksandra Huić
Department of Psychology, Faculty of Humanities and Social Sciences, University of Zagreb, Croatia

Abstract

According to self-determination theory, students with satisfied psychological needs for autonomy, competence, and relatedness have higher levels of motivation and achieve better results at school. On the other hand, need frustration is related to negative outcomes. Recently, Reeve et al. (2023) have expanded this dual model of need states into a tripartite model, adding the dormant state of psychological needs within the school context. This study aimed to evaluate the Croatian translation of the Three States Questionnaire (TSQ), focusing on psychological needs and their relationship with students' perceptions of teachers' (de)motivating styles. A total of 274 high school students from Croatia participated in an online study conducted during regular school hours. Results generally support the validity of the Croatian version of the TSQ. Its internal structure and reliability are in line with the original, while correlational patterns between need states and correlations with other need states operationaliations speak in favour of concurrent and discriminant validity. Consistent with previous research, teachers' motivating styles of autonomy support and structure were related to need satisfaction, while demotivating styles of control and chaos were related to need frustration. A significant association was found between chaotic teaching styles and the dormant need state. Strong associations of chaotic teaching with both the dormant and frustrated need states are not only in line with theoretical expectations but highlight practical interventions for teachers.

Keywords: *basic psychological needs, tripartite model of needs, teachers' (de)motivating styles, high school students, self-determination theory*

Introduction

A recent synthesis of meta-analytic studies based on self-determination theory (SDT) showed the importance of basic psychological needs satisfaction for higher well-being and better performance in numerous life areas and various cultures (Ryan et al., 2022). Basic needs for autonomy (the desire for choice, volition, and self-direction in one's actions), competence (the desire to feel effective, experience mastery, and undergo challenging tasks), and relatedness (the desire to feel connected to others, receive care, and care for others, fostering a sense of belonging within a community) represent "psychological nutrients" that are crucial for individuals' ongoing psychological development, integrity, and overall well-being (Ryan & Deci, 2017).

According to SDT, our needs can either be satisfied or frustrated/thwarted (Ryan & Deci, 2020). A satisfied need state can be described as an energy-mobilising experience that occurs when one's current activity aligns with one's desires, such as fostering good relationships with friends (relatedness). On the other hand, a frustrated state occurs when one's activity is not aligned with one's desire (e.g., working alone in the context of the need for relatedness) or in cases where the environment imposes restrictions (Reeve et al., 2023). Another important postulate from the basic psychological need theory is that a stimulating environment enables the satisfaction of psychological needs, which consequently lead to increased engagement and motivation (Deci et al., 2001; Jang et al., 2016), while a controlling environment encourages the "dark side" of the frustration process that is associated with inefficient functioning and many negative outcomes (Bartholomew et al., 2011).

However, several authors have recently proposed expanding this view to a third (dormant) need state which describes an unsatisfied or diminished need state (see Vansteenkiste et al., 2020 for review). The dormant needs state arises when a person's actions are not aligned with their fundamental psychological needs for autonomy, competence, or relatedness (Reeve et al., 2023). In this state, individuals may not be consciously aware of their unmet needs, yet the lack of fulfilment gradually depletes their motivation and engagement. Unlike frustration, which is an energy-blocking experience that occurs when one's current activity directly contradicts psychological needs. Reeve et al. (2023) further emphasised that this third need state of diminished or dormant needs might be especially useful in the school/classroom context to explain how and why students become disengaged from their lessons, ultimately leading to poorer academic outcomes. In this study, we seek to validate these ideas in a different cultural and educational context. In addition, we investigate which teacher classroom behaviours create a context for students' needs to be satisfied, frustrated, or dormant.

Students' Basic Psychological Need States in the Classroom

Satisfaction and frustration of basic psychological needs in the school context significantly impacts students' emotional and academic well-being. When teachers support basic psychological needs of their students, students are more likely to engage actively in class, achieve better learning outcomes, and show greater motivation to participate in school activities (Haerens et al., 2015; Kanat-Maymon et al., 2015). Conversely, when teachers do not support students' needs, that is when students experience need frustration, they have a pronounced feeling of obligation, perceptions of failure and inadequacy, and feelings of loneliness and exclusion, leading to disobedience and class disruptions (Bartholomew et al., 2011; Reeve, 2009; Vansteenkiste & Ryan, 2013). In addition, this has been linked to higher levels of students'-controlled motivation (Behzadnia et al., 2018).

Recently, Reeve et al. (2023) proposed the existence of dormant needs as a third state. Several au-

thors had already hypothesised the existence of this third state (Bhavsar et al., 2019; Cheon et al. 2019; Costa et al., 2015). According to them, the third, so-called, indifferent need style may result in a state of dormancy regarding one's psychological needs, characterised as a condition between satisfaction and frustration. The anticipated outcomes of such a state include amotivation, maladaptive behaviour, boredom, and disengagement. Reeve et al. (2023) noticed that students are sometimes passive, unengaged, and seem to have no energy to engage in learning activities, and they tried to apply this third need state to the school context. They call this third state the dormant need state and define it as an experience of the absence of energy that occurs when a person's current action or activity is not congruent with the person's desire to realise their need for autonomy, competence, or relatedness (e.g., "Desiring relatedness, but having no opportunity for warm social interaction.", pp. 695, Reeve et al., 2023).

They proposed a new questionnaire, the *Three States Questionnaire* (TSQ, Reeve et al., 2023), to operationalise the newly added need state. They contrast their questionnaire to other major ones designed to assess self-reported psychological needs (e.g., BPNSFS; Chen et al., 2015; BMPN, Sheldon & Hilpert, 2012), and advocate not to confound the need state with its presumed cause. The TSQ focuses on personal experiences and feelings rather than relying only on contextual items (e.g., is my need for relatedness hurt, fulfilled, or irrelevant, regardless of whether that happens in the context of a group of people I want to belong to or not). They believe that this provides a clearer and more unbiased assessment of the need state.

In two other studies conducted as part of the research by Reeve et al. (2023), the impact of need states on classroom functioning was examined using the TSQ. Their results showed that the tripartite model outperformed the existing dual-process model, particularly in predicting various indicators of diminished classroom functioning (amotivation, behaviour and agentic disengagement, boredom, deactivated affect). More precisely, the dormant state was an individually significant predictor of diminished functioning. The satisfied state was a predictor of effective functioning (intrinsic motivation, behavioural engagement, agentic engagement, vitality, academic progress, self-concept, positively activated affect) and a low level of the dormant state was a supplemental predictor that explained additional independent variance across all seven outcomes. The frustrated state was an individually significant predictor of the most indices of defiant functioning (anti-internalisation, resentment, antisocial behaviour, problematic relationships). The last indicator of defiant functioning—negatively activated affect—was solely predicted by the dormant state. The most important result from their research was the fact that the dormant state explained additional variance in both low engagement and high defiance. Overall, the study emphasises the importance of assessing the dormant psychological need state to improve adolescents' educational experiences and outcomes.

Motivating and Demotivating Teaching Styles

As we previously mentioned, whether our needs are going to be satisfied, frustrated, or dormant is mostly contingent on the context around us. Given that adolescents spend most of their time at school, the question arises as to how various teacher behaviours influence need states. So far, studies based on self-determination theory have mostly focused on the styles that teachers use to motivate their students. When teachers are autonomy supportive, students' basic psychological needs for autonomy, competence, and relatedness are being satisfied. In other words, when teachers provide meaningful choices to students, give them opportunities to voice their opinions, incorporate student interests and struggles in their lessons, students show higher levels of satisfaction for all three needs, ultimately resulting in higher intrinsic motivation, higher engagement, and superior academic outcomes (Reeve & Jang, 2006; Cheon et al., 2023; Van-

steenkiste et al., 2020). On the other hand, when teachers are controlling, impose their way of doing things, leave little room for students' feelings and interests, raise their voice, cause shame and guilt, students report frustrated levels of all three needs and inferior academic outcomes (Reeve, 2009; Soenens et al., 2012).

Recently, Aelterman et al. (2019) have formulated a more comprehensive model of teachers' (de) motivating styles. They distinguished between two dimensions which describe four distinct teaching styles. Alongside the dimension of need satisfaction/frustration which describe the already mentioned autonomy supportive and controlling teaching styles, Aelterman et al. (2019) added a dimension of directiveness in teaching which describes two additional styles of structure and chaos. Providing informative feedback, clarifying tasks and assignments, scaffolding learning activities are all teacher behaviours which direct students in their learning and at the same time are associated with satisfied needs. Hence, a teachers' structuring style should also be viewed as a motivating style (Aelterman & Vansteenkiste, 2023). A chaotic style is associated with abandoning students to their own devices, not planning teaching activities, and leaving the initiative for learning solely to students. Low effort and direction do not enable students to satisfy their psychological needs, making chaos a demotivating teaching style (Aelterman & Vansteenkiste, 2023).

So far, studies focusing on the circumplex model have mainly tried to investigate which individual teacher characteristics are associated with them using more (de)motivating styles (e.g. Huić et al., 2024a; Moè & Katz, 2021, 2022; Vermote et al., 2020, 2022), or on the consequences of using these styles for teachers' professional well-being (Huić et al., 2024b; Katz & Moè, 2024; Moè & Katz, 2020). A few available studies focusing on students found that student-perceived autonomy-supportive and structuring styles exhibit a positive correlation with student need satisfaction (Cheon et al., 2020). On the contrary, students' perception of controlling and chaotic styles leads to them experiencing higher levels of need frustration (Haerens et al., 2015). Additionally, Aelterman et al. (2019) showed that students who report having autonomy supportive and structuring teachers have higher levels of autonomous motivation and self-regulated learning, and lower levels of amotivation. These students also rated their teachers' quality of teaching with higher marks and showed lower levels of defiance in their classrooms. In accordance with the dual process model of school functioning (Bartholomew et al., 2011), demotivating styles of control and chaos were associated with disengagement, controlled motivation, and higher levels of oppositional defiance in the classrooms. Similarly, but with observational data, Cents-Boonstra et al. (2021) found motivating styles of autonomy support and structure to be associated with student engagement while they observed disengagement especially when teachers exhibited chaotic behaviours.

In their seminal paper, Reeve et al. (2023) confirmed the expectations of the dual process model and found that autonomy supportive teaching style was associated with higher levels of student classroom engagement, specifically through students' need satisfaction. The teacher's controlling behaviours were associated with higher levels of classroom defiance, through students' need frustration. However, their expanded tripartite model, which assumes that a need-indifferent environment will be associated with the third dormant need state and ultimately with student disengagement was also confirmed. Moreover, a neglecting teaching style (but not other styles) was specifically associated with higher levels of student disengagement, specifically through the new dormant need state (but not other need states). Their results are similar to a study by Bhavsar et al. (2019) in which the third state of psychological needs (which they referred to as the indifferent state) was connected to a chaotic style of interpersonal communication. A chaotic style is characterised by unpredictability and uncertainty, while a controlling style is characterised by outward control and pressure. Therefore, it is highly plausible to expect that students will react by disengaging and showing a lack of care in the first case, and by reactivity and opposition in the second case. As Reeve et al. (2023) also point out, the dormant need state and its antecedents might be the missing link needed to explain so many quiet and unfulfilling classrooms in which students fail to reach their full potential.

Current Study

In order to contribute to theoretical claims, set forth by Reeve et al. (2023), this study had two research aims. First, we wanted to examine the validity of the *Three States Questionnaire* (TSQ) and the dormant need state in a culturally different educational system.

Reeve et al. (2023) conducted their study on high school students in the USA, whose average age was slightly lower than that of the students included in this research. Additionally, our study includes students from both general education (gymnasium) and vocational programmes, whereas no such distinction was made in the original research.

Second, to expand on the original study, we incorporate the full circumplex model of teachers' (de) motivating styles in order to obtain a more nuanced picture of how different (de)motivating styles are associated with all three basic psychological need states.

We expected to confirm the proposed internal structure and reliability of the Croatian translation of the TSQ. In addition, we expected to extend data on the concurrent validity by contrasting the TSQ with another, previously highly validated instrument operationalising need satisfaction and frustration, something not done in the original study. Furthermore, we expected to establish discriminant validity of the new dormant need state, by showing low associations between the dormant need state and the frustrated/satisfied need state through different operationalisations.

In line with the dual-process and the tripartite model we expected teachers' motivating styles of autonomy support and structure to be associated with students' need satisfaction, teachers' demotivating style of control to be associated with students' need frustration, and teachers' demotivating style of chaos to be associated with students' dormant need state. According to our knowledge, this is the first research in which the new TSQ was translated and used, which enables a comparison with the original results. Furthermore, this study extends previous literature by examining the concurrent validity of the TSQ and investigating all motivating and demotivating teaching styles.

Method

Participants

A total of 274 high school students (52% boys, 44% girls, 4% others) from two high schools in Croatia participated in the study (48% from a gymnasium programme, and the other half from a vocational programme). They were 14 to 19 years of age (M = 17 years, SD = 15.015). Most students attended the third grade (33%), and the fewest the fourth grade (16%). Their average final grade was 4.05 (SD = 0.758) on a grading scale ranging from 1 (lowest) to 5 (highest).

Procedure

The students filled out an online questionnaire during regular school hours under supervision of a school psychologist (first author). The sample was convenient. When selecting classes, we aimed to ensure a roughly equal number of participants from each grade (first to fourth) in both schools and achieve equal representation of students from three-year and four-year programmes in the vocational school. The students were informed about the aim of the study and their right to withdraw at any time and were advised to ask for clarification if they had any questions. The questions that the students answered did not evoke negative emotions, nor did they involve the collection of sensitive information. The questionnaire

was anonymous, and the students were not asked about any particular teacher or class (we further discuss this decision in the description of the Three States instrument in the following paragraph, as well as in the Discussion section). The study was approved by the Ethics Committee of the Department of Psychology, Faculty of Humanities and Social Sciences, University of Zagreb. Additional consent to conduct the study was secured by the principals of the aforementioned schools. In line with the Ethical Code of Research with Children (Ajduković & Keresteš, 2020), the students provided informed consent to participate in the study, while their parents/caregivers were informed of the research study taking place.

Measures

Basic Psychological Needs Satisfaction and Frustration Scale (BPNSF; school context; Chen et al., 2015) consists of 24 items rated on a 7-point scale (1—*completely untrue*; 7—*completely true*). The questionnaire has six subscales: autonomy satisfaction (e.g., *"I feel that my choices express who I really am"*), autonomy frustration (e.g., *"I felt forced to do things I would not choose to do"*), competence satisfaction (e.g., *"I am capable of learning new knowledge at school"*), competence frustration (e.g., *"I feel like a failure because of the mistakes I make"*), relatedness satisfaction (e.g., *"Teachers and classmates care about me at school"*), and relatedness frustration (e.g., *"I feel that colleagues dislike me"*). Following Reeve et al.'s (2023) analytical strategy, we used aggregated scores in our analyses (see also Chen et al., 2015). The reliability coefficients are $\alpha = .87$ for need satisfaction and $\alpha = .84$ for need frustration.

The Three States Questionnaire (TSQ; Reeve et al., 2023) was translated by the authors of this study using the standard "back-translation" method. Specifically, the items were first translated from English into Croatian by the researchers. Then, an expert in the English language, who was neutral to the study, conducted a back-translation from Croatian into English to verify accuracy. The TSQ consists of a three-page format with each page featuring a vignette related to a specific psychological need. Reeve et al. (2023) started from the fact that in order to understand the motivational nature of psychological needs, it is necessary to exclude the contextual precondition of frustration or satisfaction of needs. To evaluate the satisfied need state, the authors developed a 4-item scale comprising the following indicators—*satisfied, fulfilled, fully realised,* and *deeply satisfied*. For the assessment of the dormant need state, a 5-item scale was created, which includes the items *unengaged, inactive, irrelevant, disconnected from what I was doing,* and *switched off*. Lastly, the frustrated state is measured using a 4-item scale that encompasses the terms *thwarted, destroyed, hurt,* and *injured*. The participants' task is to determine how much the given adjective/phrase describes their psychological need described in the vignette on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*). In sum, each psychological need consists of 13 items, with a total of 39 items in the questionnaire.

The original instruction is: *"During my Wednesday, 2nd period class, my desire for choice and personal freedom to do interesting and important activities typically feels…"*. For the purpose of this research, the questionnaire instructions were generalised. The students were asked to assess their needs without referring to a specific subject or teacher. For example: *"In class, my desire to establish a good and close relationship with the teacher and classmates is mostly..."*). This was done to ensure anonymity and honesty, and to avoid any possible back-lash against the students. Psychometric characteristics are shown in the results section.

The Situations in School Questionnaire—Student Version (SIS-S; Aelterman et al., 2019) was used to measure the students' perception of the teachers' (de)motivating styles. This instrument contains 15 scenarios that occur in the classroom (e.g., *"At a difficult moment in the lesson, students start complaining"*) together with four possible responses that can be given by the teacher to handle the situation (autonomy support: *"Accept our negative feelings as OK. Assure us that he/she is open to our input and suggestions"*; structure: *"Show and teach us a helpful strategy for how to break down the problem to solve it step-by-step"*; control: *"Insist for us to pay attention. We must learn this material for our own good"*; chaotic: *"He/she just ignores the whining and complaining. We need*

to learn to get over the obstacles by ourselves") Students report their general perception of their teacher's behaviour for each individual response (a total of 60 ratings) on a scale from 1 (*does not describe my teachers at all*) to 7 (*describes my teachers very well*). The questionnaire demonstrates good internal consistency across various styles—autonomy style (α = .88), structuring style (α = .90), chaotic style (α = .85), and controlling style (α = .83).

Results and Discussion

Internal Structure

Firstly, we examined the internal structure of the TSQ by conducting exploratory factor analysis (promax rotation) like in the original study by Reeve et al. (2023). Following their analytical strategy, we first aggregated the scores for each experienced need state into a single variable. The EFA identified three factors with eigenvalues higher than 1.0, which together explained 74.58% of the total variance (Table 1).

Table 1
Descriptive statistics and standardised parameter estimates for the EFA 3-factor solution for the Three States Questionnaire

Items	Descriptive statistics				Factor scores		
	M	*SD*	m_3	m_4	Frustrated (λ)	Satisfied (λ)	Dormant (λ)
Dormant state items							
1. Inactive	2.72	1.250	0.496	-0.122	.418	-.064	.450
2. Unengaged	2.89	1.129	0.434	-0.176	.026	-.019	.828
3. Irrelevant	2.71	1.517	0.719	-0.280	.049	.033	.855
4. Disconnected from…	2.92	1.442	0.634	-0.087	.284	-.103	.573
5. Switched to "Off"	3.44	1.538	0.250	-0.637	-.282	.056	.968
Satisfied state items							
1. Satisfied	4.63	1.148	- 0.334	0.152	.048	.796	-.163
2. Fulfilled	4.12	1.162	-0.108	0.135	.077	.932	.011
3. Fully realised	3.94	1.184	0.059	0.017	-.049	.904	.047
4. Deep satisfaction	3.40	1.206	-0.137	-0.004	-.023	.923	.096
Frustrated state items							
1. Thwarted	2.93	1.261	0.356	-0.236	.881	-.072	-.195
2. Destroyed	2.21	1.242	1.147	0.946	.836	.031	.152
3. Hurt	2.31	1.187	0.832	0.279	.944	.055	-.077
4. Injured	2.19	1.209	1.047	0.592	.891	.059	.017
Intercorrelations							
Factor 1					-	-.348	.443
Factor 2						-	-.220
Factor 3							-

Note. *N* = 274 Possible range for each variable, 1–7.
M = mean, *SD* = standard deviation, m_3 = skewness, m_4 = kurtosis.

A clear distribution of items across three need state factors aligns with the theoretical framework and is comparable to Reeve et al.'s study (2023), in which the same three factors explained 76% of the vari-

ance. The range of lambda values on the frustrated scale varies from .836 to .944 (average .89), on the satisfied scale from .796 to .932 (average .89), and on the dormant scale from .450 to .968 (average .73), and the values are again comparable with Reeve et al.'s results (2023) where the average lambdas for frustration, satisfaction, and dormancy were .85; .85 and .81, respectively.

Individual item saturations as well as factor intercorrelations provide important information about the discriminant validity of the three need states. For the most part, items load onto their respective factors, which is in accordance with the tripartite model and our expectations. Furthermore, the dormant need state is clearly distinguishable from both the frustrated and satisfied need states.

What is also comparable to the original research are the reliability coefficients, which are presented in Table 2. It is evident that the reliability coefficients for the three states of psychological needs are of almost equal values and are notably high, indicating a strong level of reliability. High Cronbach alphas are an additional indicator of the expected internal consistency of the TSQ. A recommendation for future research is to verify this reliability using the test-retest method. This approach would provide valuable information regarding the consistency and repeatability of the results. In conclusion, factor and reliability analysis of our data provide evidence for the expected internal structure of the Croatian version of the TSQ, and for the discriminant validity of the third, dormant need state proposed by the tripartite model.

Table 2
The reliability coefficients of the subscales from the TSQ (from this research study) and comparison with the original research (Reeve et al., 2023)

	This study	Reeve et al. (2023)
Satisfied need state (TSQ)	.90	.87
Frustrated need state (TSQ)	.91	.92
Dormant need state (TSQ)	.89	.90

Descriptives

By comparing the descriptive statistics obtained in our study with those from Reeve et al. (2023), we observe that the means and standard deviations for items across all three factors are almost equal (see Table 3). We found some differences in skewness and kurtosis estimates, indicating that the results for the dormant and frustrated need state in our study are more concentrated around higher values (Table 1). In addition, the kurtosis indices indicate fewer extreme values in our findings compared to Reeve et al. (2023).

Overall, the students in our study reported on experiencing the satisfied need state the most, then the dormant need state, while experiencing need frustration the least (F (1,671) = 161.74, $p < .01$; post hoc testing using the Bonferroni method indicated significant differences among all arithmetic means analysed at $p < .01$ level). As can be seen in Table 3, our results are again similar to those of Reeve et al. (2023). It is encouraging that need frustration—identified as the motivational state associated with the dark pathway and poor academic outcomes (Bartholomew et al., 2011; Buzzai et al., 2021; Haerens et al., 2015)—is not highly pronounced among our students. However, at the same time our students on average neither agree nor disagree that their needs are satisfied. Given that previous studies found important links between need satisfaction and positive academic outcomes (e. g. Buzzai et al., 2021; Jang et al., 2016), this is a somewhat deflating result. The fact that the students reported on experiencing the dormant need state to a significantly higher degree than need frustration gives further credence to our efforts to further investigate this special need state.

Table 3
Average results for the experienced need states in this study and Reeve et al. (2023)

Need state	*M*	*SD*
Satisfied state (TSQ)	4.17 (3.87)	1.041 (1.32)
Frustrated state (TSQ)	2.41 (2.38)	1.073 (1.22)
Dormant state (TSQ)	2.94 (3.03)	1.120 (1.23)

Note. (xy) = the values as shown by Reeve et al.'s (2023) study.

Construct Validity

All three need states are significantly associated with each other. The students who reported higher levels of need satisfaction reported lower levels of need frustration and dormant needs at the same time. Frustrated and dormant needs, since they are both undesirable need states, were positively correlated (see Table 4). Overall correlations are low to moderate in size, which points to good discriminant validity of the three need states.

Table 4
Intercorrelations among the three subscales of the TSQ

Variable	1	2	3
1. Satisfied	1	-.26** (-.24**)	-.32** (-.01)
2. Dormant		1	.49** (.67**)
3. Frustrated			1

Note. (xy) = the intercorrelation from Reeve et al.'s (2023); ** = p < .01.

In relation to the original study by Reeve et al. (2023), the correlation between satisfied and frustrated states in this study is higher, which aligns more closely with the theoretical framework and other studies in the field (Vansteenkiste & Ryan, 2013; Zhang & Jiang, 2023). A smaller association between the dormant and frustrated scales has been observed in this study. Reeve et al. (2023) assumed that the possible multicollinearity complicates the ability of the two states to account for independent variance in the outcomes and that this correlation may indicate more than just a measurement issue; it suggests a potential reciprocal relationship between the two states. For example, experiencing a dormant state during a lesson today could lead to the development of a frustrated state over time. On the other hand, encountering a frustrated state during today's instruction might eventually result in a dormant state. Investigating this potential reciprocal relationship will necessitate future longitudinal studies.

We examined the concurrent validity of the TSQ for the satisfied and frustrated need states by investigating correlations between the subscales of the BPNSF questionnaire by Chen et al. (2015) and the TSQ (see Table 5). The correlation between the satisfied state from the BPNSF and TSQ is positive, statistically significant, and moderate to high in value. This finding supports concurrent validity for the need satisfaction subscale from the TSQ. The correlation between the need frustration scale from the BPNSF and TSQ is only of moderate value. In order to establish good concurrent validity, we would expect a moderate to high correlation. In addition, correlations between need satisfaction and need frustration in both instruments are of comparable value giving further credence to the concurrent validity of the TSQ. Considering that the TSQ does not have the confounding effect of experienced state and context (in relation to the BPNSF), a statistically significant moderate negative correlation between the subscales measuring satisfaction and frustration of needs indicates that these are two dimensions. That supports the assumptions of self-determination theory that the satisfaction and frustration of needs are two separate, moderately associated, dimensions (Bartholomew et al., 2011).

Table 5
Intercorrelations among subscales from BPNSF and TSQ

Variable	1	2	3	4	5
1. Frustrated state (TSQ)	-	.49**	-.32**	.35**	-.32**
2. Dormant state (TSQ)		-	-.26**	.27**	-.25**
3. Satisfied state (TSQ)			-	-.29**	.52**
4. Frustrated (BPNSF)				-	-.39**
5. Satisfied (BPNSF)					-

Note. **. Correlation is significant at the 0.01 level (2-tailed).

Assessing the concurrent validity of the dormant state was not feasible due to the absence of a comparable questionnaire. It is important to note that our findings cannot be directly compared to previous studies, as concurrent validity was not examined in the original research. Future studies should continue to examine the concurrent validity of the TSQ and possibly find comparable instruments to examine the concurrent validity of the dormant need state in particular.

Relationships to Teachers' (De)motivating styles

Our second research aim was to examine the relationships between the three states of students' psychological needs and their perception of teachers' (de)motivational styles according to the circumplex model (Aelterman et al., 2019). As expected, (see Table 6), there were statistically significant positive correlations between teaching styles—autonomy and structure with the satisfaction of psychological needs, as well as statistically significant positive correlations between teaching styles—chaos and control with frustration of psychological needs.

Table 6
Intercorrelations among students' psychological need state according TSQ and their perception of teachers' (de)motivating styles

Variable	1	2	3	4	5	6	7
1. Frustrated state (TSQ)	-	.49*	-.32*	.03	-.05	.20*	.47*
2. Dormant state (TSQ)		-	-.26*	-.14*	-.24*	.12	.45*
3. Satisfied state (TSQ)			-	.41*	.39*	.00	-.10
4. Autonomy				-	.87*	.10	-.09
5. Structure					-	.20*	-.18*
6. Control						-	.46*
7. Chaos							-

*Note.**. Correlation is significant at the 0.05 level (2-tailed).

The obtained results are in line with previous studies (Aelterman et al., 2019; Cheon et al., 2012; Reeve & Cheon, 2021) which indicate that teaching behaviours that foster student autonomy are linked to the fulfilment of their needs.

Specifically, such teacher actions provide students with the opportunity to make their own decisions regarding learning and participate in the classroom. Studies evaluating the effectiveness of teaching autonomy support to teachers further show that such endeavours lead to increases in student need satisfaction (Cheon et al., 2023). Similarly, just like our results show, providing structure in teaching helps meet students' psychological needs (Aelterman et al., 2019; De Meyer et al., 2014). This teaching style aligns

students' expectations with their abilities and offers support, which helps students feel competent (Ryan & Deci, 2017).

The obtained association between demotivating styles and need frustration is also consistent with previous research (Aelterman et al., 2019; Bartholomew et al., 2011). The imposition of restrictions, pressure on students, a commanding tone, or leaving students to their own devices in the learning process negatively affects the fulfilment of psychological needs and is, moreover, associated with their frustration. Such an environment, for example, prevents students from being autonomous and from establishing a quality relationship with the teacher. In the study by Aelterman et al. (2019), need frustration had a greater correlation with chaotic rather than with controlling (de)motivating teaching style, which is in line with our results.

The most similar study for comparison of results is that of Reeve et al. (2023), but it is important to keep in mind that Reeve et al. (2023) did not apply the complete circumplex model of teaching styles. Ultimately, they analysed the association between the need states and three distinct teaching styles, which were named: supporting, neglecting, and controlling. For this reason, the full comparison between their results and ours is limited; however, findings from both studies are rather similar.

Reeve et al. (2023) reported a significantly positive correlation between need satisfaction and supportive style, while neglecting and controlling styles did not show such a correlation, which aligns with our results. Although it was expected that there would be a stronger or more equal association between need frustration and the controlling, rather than chaotic style, our research reveals a high positive correlation between the need frustration scale and chaos, and a somewhat lower correlation with control. These results align with those from the study by Aelterman et al. (2019), with the exception that they observed results in a sample of teachers. Reeve et al. (2023) reported almost equal moderate correlation between the frustrated state and controlling and abandoning style, which is not completely aligned with our findings. Our results may indicate that adolescents are more affected by an uncertain environment and a lack of guidance than by the application of a controlling style. Given that this is a correlational study, definitive conclusions cannot be drawn, but this certainly deserves more future attention.

What is interesting is the association of these teaching styles with the new dormant state. As we can see, the dormant state is statistically significantly negatively associated with autonomy and structure and positively associated with chaos. We did not obtain a significant correlation with the controlling style, but Reeve et al. (2023) did. Bhavsar et al. (2019) emphasised that chaos is a specific communication style because it does not lead to the satisfaction of psychological needs but also does not exclusively lead to frustration of needs. According to Aelterman et al. (2019), chaos style consists of two subareas—abandoning and awaiting. Compared to the controlling style, teacher behaviours that fall under these two subareas do not necessarily lead to the frustration of psychological needs, but they also do not lead to satisfaction, which is actually the definition of a dormant state of psychological needs. Therefore, it can be concluded that the obtained relationship between dormant state and chaos style is expected.

General Discussion

This study provides valuable insights into the validity of the Croatian translation of the Three States Questionnaire (TSQ) and its relationship with students' perceptions of teachers' motivating and demotivating styles. The findings confirm assumed hypotheses regarding the questionnaire's internal structure, along with its concurrent and discriminant validity. Positive associations between motivating teaching styles and psychological need satisfaction, as well as between demotivating styles and need frustration, and a statistically significant association of chaotic style with the dormant need state are further in line with our hypotheses, theoretical proposition of the SDT, as well as other studies in the field. It seems that, next to the

bright and dark students' motivational pathways (Haarens et al., 2015; Jang et al., 2016), a third dormant path might help explain why so many students are disengaged in their classrooms, assuming passive roles and precluding the best academic outcomes. Clearly, before such conclusions can be made, future studies are needed which will circumvent some of the limitations of this study.

One such limitation may be the generalisation of the instructions in the TSQ to all the classes rather than focusing on a specific lesson like in Reeve et al. (2023). Such a change in the instructions may have resulted in less precise estimates and a reduction in overall variance. Conversely, the TSQ's format differs from most questionnaires, as it does not include contextual items, only experiential states. During testing, the students had an opportunity to ask for clarifications on less familiar terms, which helped ensure the validity of the research implementation—an opportunity not afforded in study by Reeve et al. (2023). Additional research on construct validity and the possibility of applying the TSQ in different samples (e.g., primary school students, employees, teachers, college students) and with different instructions is needed. For example, future studies should examine predictive validity to see whether the TSQ can determine students' academic success, mental health, college enrolment, and alike.

In addition, our participants were from two schools in the same city, limiting the generalisability of the results. However, our results do not deviate significantly from Reeve et al. (2023) giving further credence to them. Moreover, the similarity of our results to those from the USA supports one of the main assumptions of self-determination theory, which is that the need must be pervasive, meaning that the outcomes related to its satisfaction and frustration must be the same across all social levels and cultures (Ryan and Deci, 2017). However, future studies which will help examine cultural invariability of the TSQ are needed.

Additionally, employing more objective measures for the included variables and conducting longitudinal research could clarify relationships between need states and teacher (de)motivating styles, and even examine both between- and within-person variability, and/or reciprocal relationships between the teaching context and student needs.

By focusing only on need frustration and satisfaction, there is a danger of neglecting the third need state. At the same time, this dormant state seems to be equally represented, compared to the frustration of needs, and it would be advisable to examine to which short- and long-term consequences this state leads. Additionally, our findings suggest that the application of a chaotic style is more related to need frustration than is the controlling style. This might mean that structure is of great importance to adolescents, even in cases where it is set in a controlling manner. These findings emphasise the necessity for educators to adopt motivating teaching strategies to promote student engagement and psychological well-being through need satisfaction. Given that average results of the dormant state are close to the average result on need satisfaction, there is a need to raise awareness among teachers about how their actions in the classroom impact student motivation and engagement.

This study confirmed the validity of the questionnaire and provided valuable insights into the psychological needs that influence students' functioning. Furthermore, we recommend future research to validate the factor structure using Confirmatory Factor Analysis (CFA) across different samples (e.g., primary school, high school, and university students). Additionally, relationships between teaching styles and students' psychological needs were reaffirmed, highlighting the importance of educating teachers about the significance of teaching styles, as well as the associated consequences and benefits.

The implication for future research is to identify the predictors of teachers' styles to create practical conditions that facilitate the application of the most effective teaching styles in the classroom.

References

Aelterman, N., Vansteenkiste, M., Haerens, L., Soenens, B., Fontaine, J., & Reeve, J. (2019). Toward an integrative and fine-grained insight in motivating and demotivating teaching styles: The merits of a circumplex approach. *Journal of Educational Psychology, 111*(3), 497–521. https://doi.org/10.1037/edu0000293

Aelterman, N., & Vansteenkiste, M. (2023). Need-supportive and need-thwarting socialization: A circumplex approach. In R. M. Ryan (Ed.), *The Oxford handbook of self-determination theory* (pp. 236–257). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197600047.013.21

Ajduković, M. & Keresteš, G. (2020), *Etički kodeks istraživanja s djecom*, 2. revidirano izdanje. Nacionalno etičko povjerenstvo za istraživanja s djecom.

Bartholomew, K. J., Ntoumanis, N., Ryan, R. M., Bosch, J. A., & Thøgersen-Ntoumani, C. (2011). Self-determination theory and diminished functioning. *Personality and Social Psychology Bulletin, 37*(11), 1459–1473. https://doi.org/10.1177/0146167211413125

Behzadnia, B., Adachi, P. J. C., Deci, E. L., & Mohammadzadeh, H. (2018). Associations between students' perceptions of physical education teachers' interpersonal styles and students' wellness, knowledge, performance, and intentions to persist at physical activity: A self-determination theory approach. *Psychology of Sport and Exercise, 39*, 10–19. https://doi.org/10.1016/j.psychsport.2018.07.003

Bhavsar, N., Ntoumanis, N., Quested, E., Gucciardi, D. F., Thøgersen-Ntoumani, C., Ryan, R. M., Reeve, J., Sarrazin, P., & Bartholomew, K. J. (2019). Conceptualizing and testing a new tripartite measure of coach interpersonal behaviors. *Psychology of Sport and Exercise, 44*, 107-120. https://doi.org/10.1016/j.psychsport.2019.05.006

Buzzai, C., Sorrenti, L., Costa, S., Toffle, M. E., & Filippello, P. (2021). The relationship between school-basic psychological need satisfaction and frustration, academic engagement and academic achievement. *School Psychology International, 42*(5), 497-519. https://doi.org/10.1177/01430343211017170

Cents-Boonstra, M., Lichtwarck-Aschoff, A., Denessen, E., Aelterman, N., & Haerens, L. (2021). Fostering student engagement with motivating teaching: an observation study of teacher and student behaviours. *Research Papers in Education,* 1–26. https://doi.org/10.1080/02671522.2020.1767184

Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., Duriez, B., Lens, W., Matos, L., Mouratidis, A., Ryan, R. M., Sheldon, K. M., Soenens, B., Van Petegem, S., Verstuyf, J., & Ryan, R. M. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion, 39*(2), 216–236. https://doi.org/10.1007/s11031-014-9450-1

Cheon, S. H., Reeve, J., Lee, Y., Ntoumanis, N., Gillet, N., Kim, B. R., & Song, Y. G. (2019). Expanding autonomy psychological need states from two (satisfaction, frustration) to three (dissatisfaction): A classroom-based intervention study. *Journal of Educational Psychology, 111*(4), 685-702. https://doi.org/10.1037/edu0000306

Cheon, S. H., Reeve, J., & Marsh, H. W. (2023). Autonomy-Supportive Teaching Enhances Prosocial and Reduces Antisocial Behavior via Classroom Climate and Psychological Needs: A Multilevel Randomized Control Intervention. *Journal of Sport & Exercise Psychology, 45*(1), 26–40. https://doi.org/10.1123/jsep.2021-0337

Cheon, S. H., Reeve, J., & Moon, I. S. (2012). Experimentally based, longitudinally designed, teacher-focused intervention to help physical education teachers be more autonomy supportive toward their students. *Journal of Sport & Exercise Psychology, 34*(3), 365–396. https://doi.org/10.1123/jsep.34.3.365

Cheon, S. H., Reeve, J., & Vansteenkiste, M. (2020). When teachers learn how to provide classroom structure in an autonomy-supportive way: Benefits to teachers and their students. *Teaching and Teacher Education, 90*, 1-12. https://doi.org/10.1016/j.tate.2019.103004

Costa, S., Ntoumanis, N., & Bartholomew, K. J. (2015). Predicting the brighter and darker sides of interpersonal relationships: Does psychological need thwarting matter? *Motivation and Emotion, 39*(1), 11–24. https://doi.org/10.1007/s11031-014-9427-0

De Meyer, J., Tallir, I., Soenens, B. Vansteenkiste, M., Aelterman, N., Van den Berghe, L., Speleers, L., & Haerens,L. (2013). Does Observed Controlling Teaching Behavior Relate to Students' Motivation in Physical Education?. *Journal of Educational Psychology, 106*(2), 541-554. https://doi.org/10.1037/a0034399

Deci, E. L., Ryan, R. M., Gagné, M., Leone, D. R., Usunov, J., & Kornazheva, B. P. (2001). Need satisfaction, motivation, and well-being in the work organizations of a former Eastern bloc country: A cross-cultural study of self-determination. *Personality and Social Psychology Bulletin, 27*(8), 930–942. https://doi.org/10.1177/0146167201278002

Haerens, L., Aelterman, N., Vansteenkiste, M., Soenens, B., & Van Petegem, S. (2015). Do perceived autonomy-supportive and controlling teaching relate to physical education students' motivational experiences through unique pathways? Distinguishing between the bright and dark side of motivation. *Psychology of Sport and Exercise, 16*, 26–36. https://doi.org/10.1016/j.psychsport.2014.08.013

Huić, A., Pavlin-Bernardić, N., & Čižić, N. (2024a). Teachers' basic psychological needs, (de)motivating styles and professional well-being. *Primenjena psihologija, 17*(3), 407–432. https://doi.org/10.19090/pp.v17i3.2521

Huić, A., Pavlin-Bernardić, N., & Vlahović-Štetić, V. (2024b). Investigating the Circumplex Model of (De) Motivating Teaching Styles in Higher Education. *Psihologijske teme, 33* (2), 417-437. https://doi.org/10.31820/pt.33.2.9

Jang, H., Kim, E. J., & Reeve, J. (2016). Why students become more engaged or more disengaged during the semester: A self-determination theory dual-process model. *Learning and Instruction, 43*, 27–38. https://doi.org/10.1016/j.learninstruc.2016.01.002

Kanat-Maymon, Y., Benjamin, M., Stavsky, A., Shoshani, A., & Roth, G. (2015). The role of basic need fulfillment in academic dishonesty: A self-determination theory perspective. *Contemporary Educational Psychology, 43*, 1–9. https://doi.org/10.1016/j.cedpsych.2015.08.002

Katz, I., & Moè, A. (2024). Exploring teachers' psychological needs, motivating styles, emotion regulation and self-compassion: A comparative study before and during the COVID-19 lockdown. *Teaching and Teacher Education, 148*, 104706. https://doi.org/10.1016/j.tate.2024.104706

Moè, A., & Katz, I. (2020). Self-compassionate teachers are more autonomy supportive and structuring whereas self-derogating teachers are more controlling and chaotic: The mediating role of need satisfaction and burnout. *Teaching and Teacher Education, 96*, 1-9. https://doi.org/10.1016/j.tate.2020.103173

Moè, A., & Katz, I. (2021). Emotion regulation and need satisfaction shape a motivating teaching style. *Teachers and Teaching, 27*(5), 370–387. https://doi.org./10.1080/13540602.2020.1777960

Moè, A., & Katz, I. (2022). Need satisfied teachers adopt a motivating style: The mediation of teacher enthusiasm. *Learning and Individual Differences, 99*. https://doi.org/10.1016/j.lindif.2022.102203

Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist, 44*, 159–175. https://doi.org/10.1080/00461520903028990

Reeve, J., & Cheon, S. H. (2021). Autonomy-supportive teaching: Its malleability, benefits, and potential to improve educational practice. *Educational Psychologist, 56*(1), 54–77. https://doi.org/10.1080/00461520.2020.1862657

Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology, 98*, 209-218. https://doi.org/10.1037/0022-0663.98.1.209

Reeve, J., Jang, H. R., Cheon, S. H., Moss, J. D., Ko, H., & Jang, H. (2023). Extending self-determination theory's dual-process model to a new tripartite model to explain diminished functioning. *Motivation and Emotion, 47*, 691–710. https://doi.org/10.1007/s11031-023-10019-0

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford.

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*, 1-31. https://doi.org/10.1016/j.cedpsych.2020.101860

Ryan, R. M., Duineveld, J. J., Di Domenico, S. I., Ryan, W. S., Steward, B. A., & Bradshaw, E. L. (2022). We know this much is (meta-analytically) true: A meta-review of meta-analytic findings evaluating self-determination theory. *Psychological Bulletin, 148*(11-12), 813-842. https://doi.org/10.1037/bul0000385

Soenens, B., Park, S. Y., Vansteenkiste, M., & Mouratidis, A. (2012). Perceived parental psychological control and adolescent depressive experiences: a cross-cultural study with Belgian and South-Korean adolescents. *Journal of Adolescence, 35*(2), 261–272. https://doi.org/10.1016/j.adolescence.2011.05.001

Sheldon, K. M., & Hilpert, J. C. (2012). The Balanced Measure of Psychological Needs (BMPN) scale: An alternative domain general measure of need satisfaction. *Motivation and Emotion, 36*(4), 439–451. https://doi.org/10.1007/s11031-012-9279-4

Vansteenkiste, M., & Ryan, R. M. (2013). On psychological growth and vulnerability: Basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration, 23*, 263-280. https://doi.org/10.1037/a0032359

Vansteenkiste, M., Ryan, R.M. & Soenens, B. (2020). Basic psychological need theory: Advancements, critical themes, and future directions. *Motivation and Emotion, 44*, 1–31. https://doi.org/10.1007/s11031-019-09818-1

Vermote, B., Aelterman, N., Beyers, W., Aper, L., Buysschaert, F., & Vansteenkiste, M. (2020). The role of teachers' motivation and mindsets in predicting a (de)motivating teaching style in higher education: A circumplex approach. *Motivation and Emotion, 44*(2), 270–294. https://doi.org/10.1007/s11031-020-09827-5

Vermote, B., Vansteenkiste, M., Aelterman, N., Van der Kaap-Deeder, J., & Beyers, W. (2022). Teachers' psychological needslink social pressure with personal adjustment and motivating teaching style. *The Journal of Experimental Education, 91*(4), 696–717. https://doi.org/10.1080/00220973.2022.2039584

Zhang, L., & Jiang, Y. (2023). Patterns of the Satisfaction and Frustration of Psychological Needs and Their Associations with Adolescent Students' School Affect, Burnout, and Achievement. *Journal of Intelligence 11(*111). https://doi.org/ 10.3390/jintelligence11060111

9

Actively Open-Minded Thinking and Science Trust: Exploring Croatian Citizens' Views on Anthropogenic Climate Change and Scientific Consensus on the Issue

Marina Maglić
The Institute of Social Sciences Ivo Pilar, Zagreb, Croatia
Marina.Maglic@pilar.hr

Abstract

Despite the undeniable scientific consensus regarding the human impact on climate change (CC), research shows that public opinion varies, while digital media abound with information often contradicting scientific findings. In this study, I explore Croatian citizens' acceptance of anthropogenic CC and their perception of the scientific consensus from the dual-process perspective. Specifically, the primary focus of the study is on the predictive role of actively open-minded thinking (AOT) and science trust as its potential mediator. Additionally, the predictive effects of cognitive reflection and the mediating roles of political and general media trust are explored. Data were collected on a nationally representative sample based on age and gender quotas (N = 1528). The findings reveal that while a majority of citizens accept the reality of anthropogenic CC, many still perceive significant disagreement among scientists on the issue. In the tested structural equation models, AOT emerged as a significant positive predictor of acceptance of anthropogenic CC and a rather weak positive predictor of the scientific consensus on the issue. AOT also positively predicted trust in science and scientists, which partially mediated its effect on acceptance of anthropogenic CC, while there was little evidence for its mediation effect on the perception of scientific consensus. Additionally, media and science trust exhibited positive relationships, while political trust showed a negative relationship with the acceptance of anthropogenic CC. Overall, the findings highlight the importance of a tendency to think with an actively open mind and trust in science and scientists for the acceptance of anthropogenic CC.

Keywords: *climate change, actively open-minded thinking (AOT), cognitive reflection, science trust, Croatia*

Introduction

Climate change (CC) is a critical and urgent global challenge, posing significant risks to ecosystems, economies, and human societies. The overwhelming scientific consensus is that human activities have been the primary cause of changes in the Earth's climate since the mid-19th century, often referred to as anthropogenic climate change (Lynas et al., 2021; Myers et al., 2021). Despite the robust scientific evidence, public opinion on CC varies, and widespread misinformation in media has contributed to a gap between scientific understanding and public perception (e.g., see Lewandowsky et al., 2017). The persistence of CC scepticism and varying levels of public trust in science highlights the complexity of communicating climate science to the general public (e.g., Cologna et al., 2024). Namely, public trust in science, scientists, and other institutions responsible for communicating climate-related information, such as political institutions and the media, play a relevant role when considering public attitudes, beliefs, and risk perceptions regarding CC and support for environmental policies (Bogert et al., 2024; Fairbrother, 2017; Fawzi et al., 2021; Kulin & Johansson Sevä, 2021; Strömbäck et al., 2020; Tsfati et al., 2022).

However, even in areas where scientific consensus is unequivocal, such as CC, factors like scientific misconduct, alongside various conspiratorial narratives readily spread online, may erode confidence in scientific findings and foster scepticism. This is particularly concerning in regions like Croatia, which, as part of southern Europe and the Mediterranean region, is projected to be severely affected by CC, facing increased temperatures, more frequent and intense heatwaves, and a higher risk of droughts (European Environment Agency, 2017; Vitali Čepo, 2021).

Analytic Thinking, Trust, and Climate Change Beliefs from the Dual-Process Perspective

Understanding the beliefs of Croatian citizens regarding anthropogenic CC is critical for developing effective policies and interventions to mitigate and adapt to climate impacts. However, climate change beliefs are influenced by many psychological and societal factors and can be approached from many different perspectives and tested within different theoretical models. In this study, I set out to examine the beliefs of Croatian citizens towards CC, specifically their beliefs about the anthropogenic causes of CC and the scientific consensus on the issue from the dual-process perspective. I primarily focus on the propensity for actively open-minded thinking but also explore effects of cognitive reflection—the two indicators of analytic thinking that have been shown relevant when investigating rational judgments and decisions. Specifically, research has shown that individuals prone to analytic thinking, indexed by the Actively Open-Minded Thinking (AOT) Scale and Cognitive Reflection Test (CRT) are less prone to a wide range of epistemologically unfounded beliefs, such as paranormal and superstitious beliefs, various unscientific beliefs (e.g., rejection of evolutionary theory and climate change, acceptance of alternative and complementary medicine), so-called pseudo-profound bullshit, conspiratorial beliefs, and believing fake news (e.g., Binnendyk & Pennycook, 2022; Bronstein et al., 2019; Bowes et al., 2023; Mirhoseini et al., 2023; Pennycook et al., 2015c; Pennycook et al., 2020; for a review see Pennycook et al., 2015a; Stanovich et al., 2016).

Within the dual-process framework, thinking and reasoning critically about a complex issue such as CC is an example of judgment under uncertainty (Pennycook**,** 2023). Within this framework, the AOT and CRT as indicators of individual differences in analytical thinking have proven particularly valuable in investigating variations in rationality and the quality of thinking and decision-making. According to the *three-stage dual-process model* (Pennycook, 2023; Pennycook et al., 2015b; see also Stanovich, 2011; Stano-

vich et al., 2016), reasoning errors can occur at any stage of reasoning, either due to faulty intuitions or unsuccessful engagement of analytic reasoning. When it comes to the failure to engage in analytical reasoning, there are two pathways of cognitive miserliness. The first one is failing to recognise the need to engage in analytical reasoning. The second one is engaging in analytical thinking upon detecting a conflict, but not to an extent that facilitates accuracy (cognitive decoupling). Instead, it is applied merely to rationalise the intuitive response (see also, e.g., Cushman, 2020; Mercier & Sperber, 2017, for a discussion on the rational function of rationalisation). Theoretically, the CRT and AOT could tap tendencies towards cognitive miserliness at different levels of processing.

Actively open-minded thinking, measured by the AOT scale, is a disposition pertaining not only to the amount of thinking but also to its quality and direction (Baron, 2019; Baron et al., 2023), and it may serve a protective function against various biases and errors in judgment, particularly overconfidence and biases favouring one's own side. It represents a broad thinking disposition, encompassing reflectiveness, an active inclination to seek out and weigh new evidence (including the perspectives of others) against own preferred or prevailing opinions, and a willingness to update beliefs when faced with new or contradictory evidence (Baron, 2019; Baron et al., 2015; Haran et al., 2013; Stanovich et al., 2016). Hence, AOT as a cognitive style is open because it involves treating information and evidence fairly, as well as active because it requires an intentional search for alternatives (Baron, 2019).

Furthermore, AOT serves as a latent prescriptive norm for "good thinking", indicating how to approach and process information and evidence, evaluate their quality, and assess the credibility of sources (Baron, 2019; Baron et al., 2023). In this way, AOT provides standards for evaluating the reasoning of others, particularly the trustworthiness and reliability of sources that present themselves as authorities on the issue (Baron, 2019; Baron et al., 2023; see also Cohen et al., 2022). Thus, AOT is not just sceptical thinking; it includes recognising situations where trust in others (e.g., experts) is justified by seeking out cues that individuals or organisations who are able to provide information also adhere to principles of open-minded thinking. This means individuals do not always have to investigate every piece of information themselves, especially in areas where they lack knowledge or expertise, if they know that credible experts have already done so (Baron, 2019).

In other words, AOT pertains to the principles of sound reasoning and decision making (Baron, 2019; Baron et al., 2023; Pennycook et al., 2020). According to Stanovich and Toplak (2023), AOT should capture the analytical processes of cognitive decoupling and decontextualisation required in hypothetical reasoning. Consequently, individuals prone to AOT are more adapt at fairly evaluating evidence, situations, and source trustworthiness. Thus, in addition to self-directed, evidence-based reasoning and belief formation, AOT encompasses evaluating the reasoning of others against the standards of "good thinking", including the credibility and expertise of different sources.

Thus, in the context of reasoning about climate change, exposed to vast amounts of information and misinformation from different sources, individuals high in AOT would be more likely to critically assess evidence and recognise more credible sources of information.

On the other hand, cognitive reflection, measured with the Cognitive Reflection Test (CRT), is the ability and disposition to override a predominant intuitive but incorrect response and engage in further reflection, leading to the correct response (Frederick, 2005). The CRT is thus often viewed as a promising measure of overcoming cognitive miserliness (Stanovich et al., 2016). A key characteristic of the CRT, according to Baron (Baron, 2019; Baron et al., 2023), is a stable disposition towards allocating time to careful deliberation (i.e., the amount of thinking), to which he refers as the reflectiveness/impulsivity tendency (but see De Neys, 2014; Bago & De Neys, 2020, for a discussion of logical intuitions and Stanovich, 2018 for a discussion on mindware). In the context of reasoning about climate change, ability and tendency to engage in deliberative reasoning could be particularly relevant when misinformation is present.

Current Study

The objective of this study was to explore Croatian citizens' views on anthropogenic climate change and the scientific consensus on the issue from a dual-process perspective. Specifically, the study focuses on the predictive role of analytic thinking—primarily AOT—and trust in science and scientists as its potential mediator.

Drawing from the presented theoretical and empirical foundations, several hypotheses can be proposed.

Given that the CRT and AOT have been linked to a wide range of beliefs and behaviours in everyday life, individuals who exhibit higher levels of analytic thinking would be more likely to accept the anthropogenic causes of CC and the scientific consensus on the issue, as they may be better equipped to navigate (mis)information and arrive at evidence-based conclusions. Thus, the first hypothesis was that the AOT would positively predict acceptance of anthropogenic CC (H1a) and scientific consensus on the issue (H1b). Also, the CRT would positively predict acceptance of anthropogenic CC (H2a) and scientific consensus on the issue (H2b).

More importantly, since scientific research can provide most reliable systematic evidence on the causes of CC, individuals higher in AOT would more easily identify credible sources regarding CC and, in turn, place greater trust in them (H3).

Furthermore, since the scientific consensus on the issue of anthropogenic CC is unequivocal, which scientists have been consistently messaging, trust in science and scientists would be positively associated with the acceptance of anthropogenic CC (H4a) and of the scientific consensus (H4b).

And finally, in addition to its direct effect, the hypothesis was that the AOT would indirectly—via science trust—predict acceptance of anthropogenic CC (H5a) and scientific consensus (H5b). Additionally, I explore the indirect effect of the CRT, via science trust.

Regarding the other two dimensions of institutional trust, I adopted an exploratory approach. General media trust, trust in science and scientists, and political trust were entered simultaneously in the SEM with freed covariances, as these three dimensions of institutional trust are intertwined yet distinct. While they collectively shape individuals' perceptions of institutional credibility and the flow of information in society, each dimension represents a unique source of trust with potentially different implications for belief formation and information processing (see, e.g., de Zúñiga et al., 2019; Fawzi et al., 2021; Ladd, 2012; Nisbet et al., 2015; Pechar et al., 2018; Shehata & Strömbäck, 2021; Tsfati & Cappella, 2003; Tsfati et al., 2022; Verboord et al., 2023).

The issue of anthropogenic CC is not politicised in Croatian public discourse, and the vast majority of political actors do not abuse it in their political agendas. Instead, they typically nominally acknowledge it while prioritising many other issues as the most pressing national challenges. Furthermore, political trust in Croatia has been at low levels (e.g., Franc et al., 2020; Maglić, 2023b). Thus, I explore whether a higher level of political trust would predict acceptance of anthropogenic CC and scientific consensus.

Most (mainstream) media in Croatia is, for the most part, reliably reporting the anthropogenic CC evidence and scientific consensus, although various fringe media outlets abound with misinformation, including CC (Broz, 2024). Also, Croatian citizens perceive that they are exposed to disinformation and fake news close to the average European citizen (European Commission, Secretariat-General, 2023, p. 35). By using a general media trust measure, I explore whether a higher level of media trust would predict acceptance of anthropogenic CC and scientific consensus.

Additionally, I explore the indirect effects of the AOT and CRT, via political trust and media trust, on acceptance of anthropogenic CC and scientific consensus.

Method

Participants and Procedure

The participants were recruited into the study through Talk Online Panel with the goal of obtaining a quota representative sample regarding sex, age, and place of residence of adult citizens of the Republic of Croatia (> 18 years; for more details see Maglić, 2023a). The data were collected in December 2021 and January 2022 with the approval of ethics committees of the Department of Psychology at the Faculty of Humanities and Social Sciences, University of Zagreb, and the Institute of Social Sciences Ivo Pilar.

Of 1584 participants whose data were available in the initial cleaned dataset (for details see Maglić, 2023a), I used the data from 1528 (*M*age = 47, median = 48, *SD*age = 15.23, range: 18-85; 52% females) individuals who responded to all the relevant items.

Measures

Climate Change Beliefs

Climate change beliefs were measured with five items—four items relate to the *causes of CC*, and the final one relates to the *perception of scientific consensus on anthropogenic CC* (see Figure 1 and https://osf.io/2yndf/). The response scale ranged from 1 (*do not agree at all*) to 6 (*completely agree*), and the items were recoded such that higher scores indicate a more proscience position. A unidimensional CFA model of the beliefs regarding the causes of CC (four items) was tested, and it displayed a very good fit (robust RMSEA = .039, 90% CI [0, .080], robust CFI = .997, SRMR = .013) and acceptable reliability (ω = .75; for further details, see https://osf.io/2yndf/).

Measures of Analytic Thinking

The dispositional tendency to think with an actively open mind about evidence was measured with an eight-item version of the Actively Open-Minded Thinking (AOT) Scale (items taken from Baron, 2019; Baron et al., 2015; Haran et al., 2013; Bronstein et al., 2019). The response scale ranged from 1 (*do not agree at all*) to 6 (*completely agree*), and the items were recoded such that higher scores indicate higher levels of the measured concept. In this study, a unidimensional CFA model of the open-minded thinking disposition was tested, with freed covariances for the items reversed in the direction of less openness (i.e., closed-mindedness). The overall CFA model displayed a very good fit (robust RMSEA = .057, 90% CI [.045, .069], robust CFI = .964, SRMR = .035), although the scale exhibited a lower level of reliability (ω = .65; for further details, see https://osf.io/2yndf/).

As a performance-based measure of the ability and disposition to engage in analytic thinking, an eight-item version of the Cognitive Reflection Test (CRT) was used, based on the three original items (Frederick, 2005) and items taken from the research by Primi et al. (2016), Toplak et al. (2014), Thomson and Oppenheimer (2016), and Oldrati et al. (2016). The correct answers were coded as 1, while incorrect (both intuitive and all other) responses were coded as 0. In this study, a unitary latent factor of cognitive reflection was extracted with the overall CFA model displaying a very good fit (scaled RMSEA = .045, 90% CI [.035, .055], scaled CFI = .989, SRMR = .045) and reliability (ω = .81; for further details, see https://osf.io/2yndf/).

Institutional Trust

Trust in political institutions was measured by asking the participants whether they trusted: the Croatian government; politicians; the Croatian parliament (ω = .90; for further details, see https://osf.io/2yndf/).

Trust in science and scientists was measured by simply asking the participants whether they trusted science and scientists.

General media trust was measured by simply asking the participants whether they trusted the media.

The response scale for all the institutional trust items ranged from 0 (*do not trust at all*) to 6 (*trust completely*), and all the responses were coded such that higher scores indicate higher levels of the measured concept.

Control Variables

Sociopolitical orientation was measured with a single item—the participant's self-placement on a continuum ranging from (1) *very left/liberal* to (7) *very right/conservative*.

Religiosity was also measured using a single-item measure taken from the European Social Survey (2018): *"Regardless of whether you belong to a particular religion, how religious would you say you are?"* (0 = *not at all religious*; 10 = *very religious*).

Sociodemographic variables, specifically age, sex, educational level, income, and size of place of residence, were also entered as control factors in the analyses.

Analytic Strategy

All analyses were performed in *R* (available at https://osf.io/2yndf/), dominantly using packages lavaan (Rosseel, 2012), semTools (Jorgensen et al., 2022), psych (Revelle, 2023), ggplot2 (Wickham, 2016), and semPlot (Epskamp, 2022).

Hypotheses testing was based on structural equation modelling. In the first, base SEM model, the CRT and AOT were tested as predictors of anthropogenic CC beliefs and the perception of scientific consensus on anthropogenic CC, with sociodemographic variables, religiosity, and sociopolitical orientation included as control variables. Next, to test the indirect effects of the AOT via science trust and explore other mediation effects, the SEM mediation model was constructed (with sociodemographic variables, religiosity, and sociopolitical orientation included as controls).

Results

Firstly, basic descriptive statics and the (model-implied) correlations of all variables are presented, followed by results from testing the two SEM models.

Descriptive statistics and intercorrelations of all variables are shown in Table 1, and participant agreement with each CC statement in Figure 1. Citizens mostly agreed that the average global temperature has been continuously rising since the Industrial Revolution, that human activity is the primary cause of global warming, and that human-caused CC will increase the frequency of hurricanes, floods, and droughts, and they somewhat disagree that all recent CC is primarily caused by the Sun. Yet, they, on average, somewhat agreed that there is significant disagreement among scientists about whether CC is primarily caused by human activities (for the frequency distribution of the responses to each item, see Figure 1). Additionally,

Croatian citizens, on average, expressed relatively low levels of trust in political institutions and general media trust, and relatively high in science and scientists (Table 1).

Figure 1

Participant agreement with CC statements—four regarding anthropogenic CC beliefs and one regarding the perception of scientific consensus on the issue

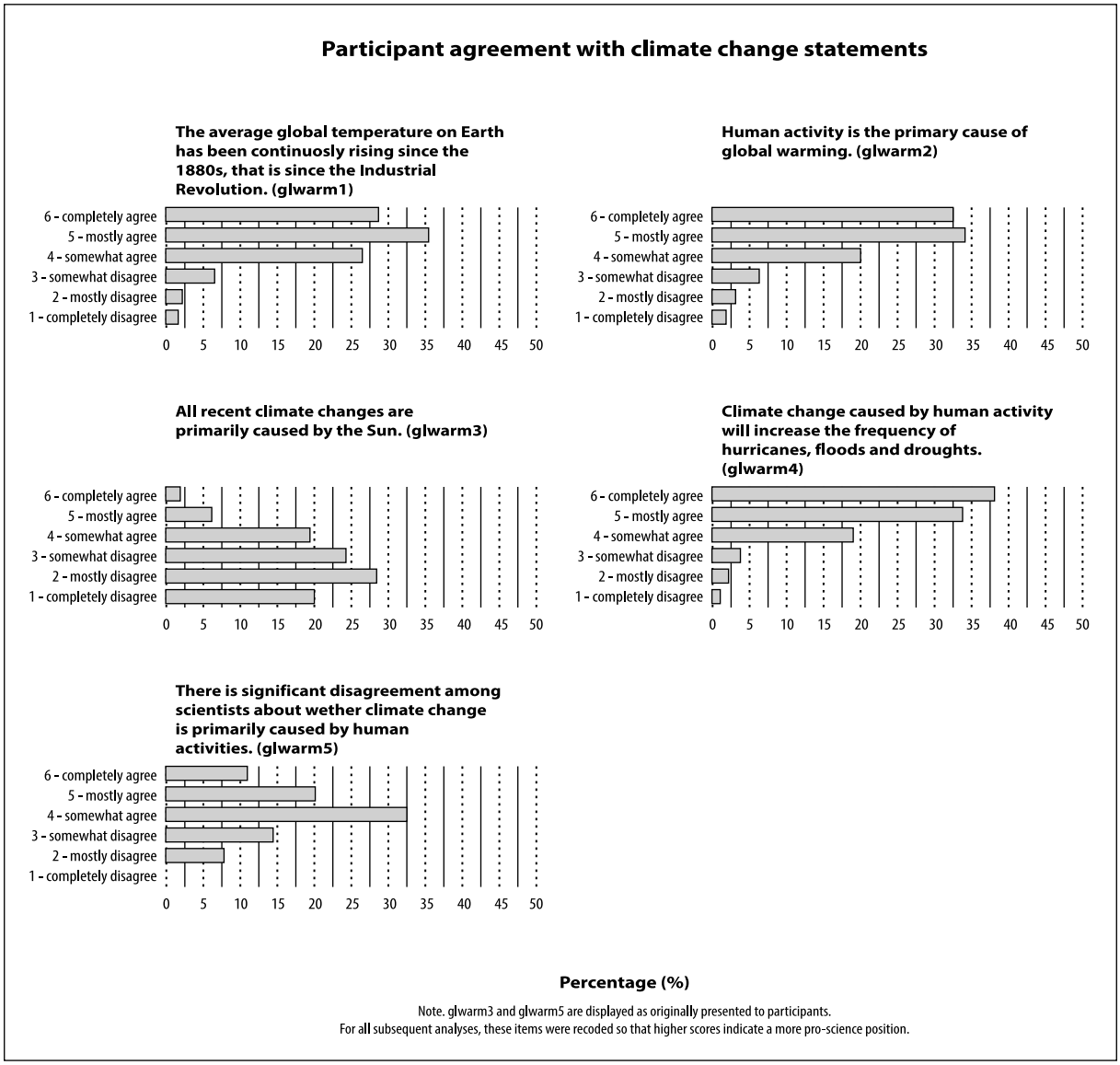


Table 1

Descriptive data and model-implied correlations between all variables

	M	*SD*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
(1) GL_warm	-	-	1																				
(2) glwarm1	4.78	1.08	.72	1																			
(3) glwarm2	4.82	1.18	.72	.52	1																		
(4) glwarm3 (R)	4.32	1.25	.80	.58	.58	1																	
(5) glwarm4	4.98	1.09	.36	.26	.26	.29	1																
(6) glwarm5 (R)	3.25	1.41	.19*	.14	.14	.16	.07	1															
(7) AOT	-	-	.26*	.19	.19	.21	.09	.14*	1														
(8) CRT	-	-	.02	.01	.01	.02	.01	.13*	.38*	1													
(9) science and scientists	4.2	1.42	.30*	.22	.22	.24	.11	.13*	.32*	.19*	1												
(10) polit_TR	-	-	-.05	-.04	-.04	-.04	-.02	.02	-.02	.02	.31*	1											
(11) politicians	0.75	1.08	-.04	-.03	-.03	-.03	-.01	.01	-.02	.02	.23	.76	1										
(12) government	1.28	1.48	-.04	-.03	-.03	-.04	-.02	.01	-.02	.02	.27	.88	.67	1									
(13) parliament	1.13	1.31	-.05	-.03	-.03	-.04	-.02	.01	-.02	.02	.28	.91	.69	.80	1								
(14) media	1.7	1.39	.15*	.11	.11	.12	.05	.08*	.01	.01	.34*	.47*	.35	.41	.42	1							
(15) PolitID	3.45	1.3	-.19*	-.14	-.14	-.15	-.07	-.08*	-.22*	-.08*	-.15*	.19*	.15	.17	.18	-.16*	1						
(16) religiosity	4.67	3.02	-.10*	-.07	-.07	-.08	-.04	-.13*	-.27*	-.19*	-.14*	.18*	.13	.15	.16	-.03	.43*	1					
(17) sex	1.48	0.5	-.08*	-.06	-.06	-.07	-.03	.01	.05	.30*	<.01	.05	.04	.05	.05	.01	.09*	-.13*	1				
(18) age	47.39	15.23	-.02	-.02	-.02	-.02	-.01	-.03	.02	-.05	.12*	.08*	.06	.07	.07	.13*	-.10*	-.07*	.12*	1			
(19) education	4.78	1.04	.08*	.05	.05	.06	.03	.07*	.19*	.22*	.17*	.02	.02	.02	.02	.02	-.08*	-.01*	-.06	-.01	1		
(20) income	3.86	1.69	.06*	.04	.04	.05	.02	.06*	.15*	.17*	.09*	.02	.02	.02	.02	-.03	.02	-.01	.22*	-.02	.38*	1	
(21) citysize	3.47	1.47	.03	.02	.02	.03	.01	.07*	.09*	.05	.06*	-.03	-.02	-.03	-.03	.03	-.12*	-.10*	-.02	.03	.16*	.12*	1

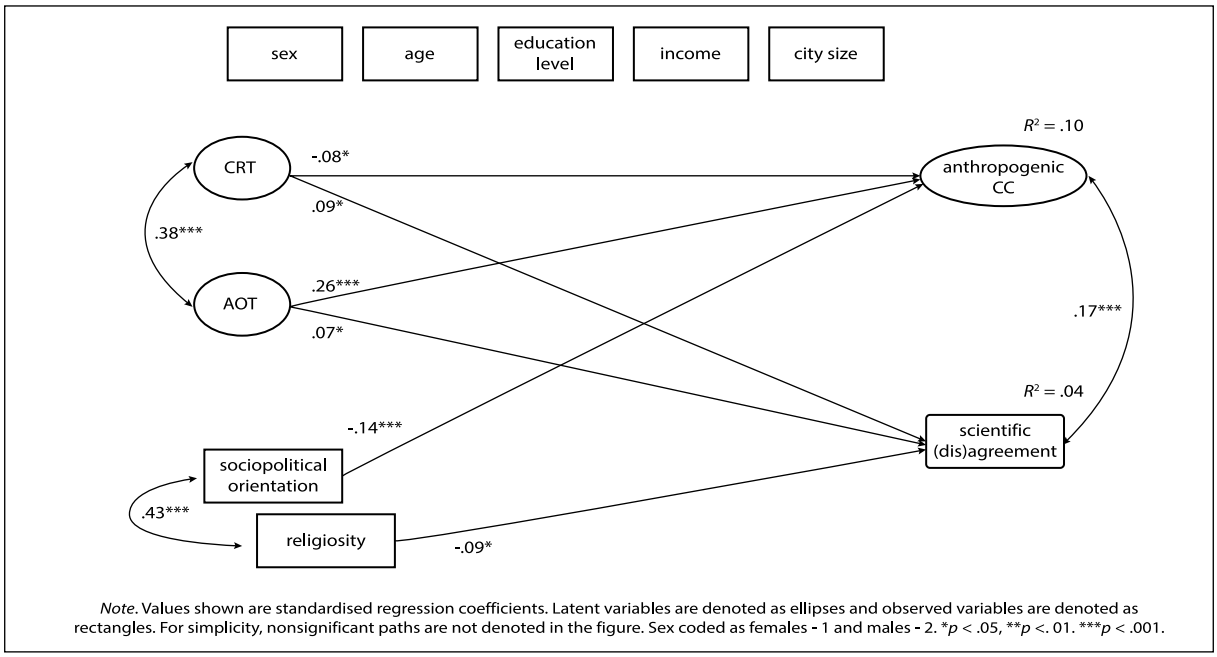
Note. Note. glwarm1-glwarm4 = anthropogenic causes of CC items; glwarm5 = perception of scientific consensus; polit_TR = trust in political institutions; PolitID = sociopolitical orientation, sex coded as females—1 and males—2, citysize = size of place of residence. (R) denotes a reverse-coded item. All |r| ≥ .07 significant at $p < .05$ and denoted with *. For observed indicators (glwarm1—glwarm4, politicians, government, parliament), correlations were constrained by factor loadings, preventing standard error estimation using the robust WLSMV estimator. Consequently, p-values for these correlations were not computed.

In the next step, using structural equation modelling, the CRT and AOT were tested as predictors of anthropogenic CC beliefs and the perception of scientific consensus on anthropogenic CC, with sociodemographic variables, religiosity, and sociopolitical orientation included as control variables. This base model[1] exhibited an acceptable fit (scaled RMSEA = .042, 90% CI [.040, .045], scaled CFI = .913, SRMR = .044). The AOT emerged as the strongest positive predictor of the acceptance of anthropogenic CC, yet a rather weak positive predictor of the perception of the scientific consensus on CC (Figure 2). The effects of the CRT, over and above the effects of the AOT, were generally rather weak and interestingly in the opposite direction of what was expected in the case of the perception of the scientific consensus (Figure 2).

Among the control variables, sociopolitical orientation exhibited a statistically significant negative relationship with the acceptance of anthropogenic CC, with left/liberal-leaning individuals being more inclined to accept anthropogenic CC. Religiosity exhibited a statistically significant (though weak) negative relationship with the perception of the scientific (dis)agreement, with more religious individuals being somewhat more likely to believe in the lack of scientific consensus on anthropogenic CC. Basic sociodemographic characteristics did not exhibit statistically significant relationships with either of the outcome variables.

Figure 2

Structural equation model of the CRT and AOT predicting anthropogenic CC beliefs and the perception of scientific consensus on anthropogenic CC, with sociodemographic variables, religiosity, and sociopolitical orientation as control variables



Note. Values shown are standardised regression coefficients. Latent variables are denoted as ellipses and observed variables are denoted as rectangles. For simplicity, nonsignificant paths are not denoted in the figure. Sex coded as females - 1 and males - 2. *p < .05, **p <. 01. ***p < .001.

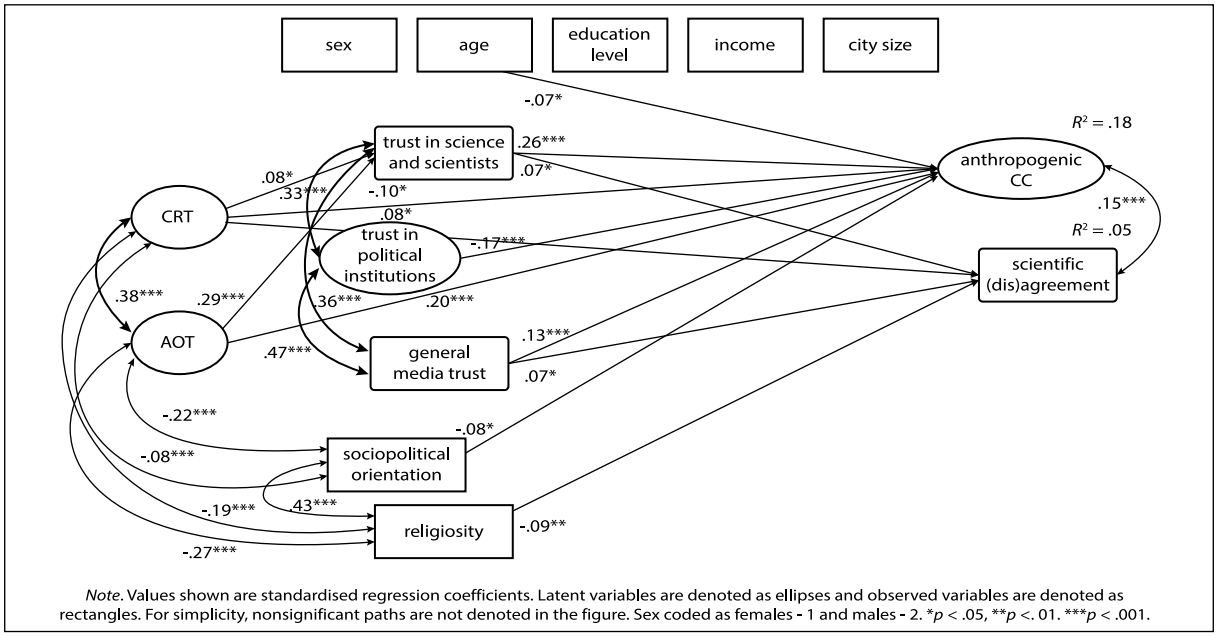
In the next model, the mediating role of different dimensions of institutional trust was tested (Figure 3). The model[2] achieved an adequate fit (scaled RMSEA = .037, 90% CI [.035, .040], scaled CFI = .917, SRMR = .041) and overall explained 18% of the variance in anthropogenic CC beliefs and only a negligible portion (5%) of the variance in perceptions of scientific (dis)agreement on the issue. The results demonstrate that the effect of the AOT on the acceptance of anthropogenic CC was partially mediated by higher science trust (25% of the total effect, with 71% attributed to the direct effect), while the effects of the CRT were generally relatively weak, and interestingly in the opposite direction of what was expected in the case

[1] The robust WLSMV (weighted least squares means and variance adjusted) method of diagonally weighted least squares was used to estimate the parameters, suitable for modelling categorical or ordinal data (Beauducel & Herzberg, 2006; Li, 2016).

[2] The robust WLSMV (weighted least squares means and variance adjusted) method of diagonally weighted least squares was used to estimate the parameters, suitable for modelling categorical or ordinal data (Beauducel & Herzberg, 2006; Li, 2016).

Figure 3

Structural equation modelling of political trust, trust in science and scientists, and general media trust as mediators of the relationship of the AOT and CRT with anthropogenic CC beliefs and the perception of scientific consensus on anthropogenic CC



Note. Values shown are standardised regression coefficients. Latent variables are denoted as ellipses and observed variables are denoted as rectangles. For simplicity, nonsignificant paths are not denoted in the figure. Sex coded as females - 1 and males - 2. *p < .05, **p <. 01. ***p < .001.

of the perception of the scientific consensus (Figure 3 and Table 2).

Beyond their mediating roles, science and media trust exhibited positive relationships, while political trust showed a negative relationship with the acceptance of anthropogenic CC (Figure 3). The relationships of science trust and media trust on the perception of scientific consensus on anthropogenic CC also were statistically significant and positive (though rather weak) (Figure 3).

Table 2

Estimation of indirect and total effects

Estimation of indirect and total effects	*b*	*SE*	z-value	*p*	β
AOT → science → anthropogenic CC	0.09	0.01	6.36	< .001	.07
AOT → political → anthropogenic CC	0.01	0.01	0.99	.324	.01
AOT → media → anthropogenic CC	0.001	0.01	0.27	.791	.001
Total AOT effect on acceptance of anthropogenic CC	0.34	0.05	7.58	< .001	.28
AOT → science → scientific disagreement	0.04	0.02	2.36	.018	.02
AOT → political → scientific disagreement	0.001	0.003	0.44	.659	.001
AOT → media → scientific disagreement on CC	0.001	0.005	0.27	.791	.001
Total AOT effect on scientific disagreement on CC	0.16	0.07	2.28	.023	.07
CRT → science → anthropogenic CC	0.02	0.01	2.42	.015	.02
CRT → political → anthropogenic CC	-0.01	0.01	-0.93	.355	-.01
CRT → media → anthropogenic CC	< 0.001	0.004	0.08	.938	< .001
Total CRT effect on acceptance of anthropogenic CC	-0.08	0.04	-2.05	.040	-.08
CRT → science → scientific disagreement	0.01	0.01	1.72	.085	.01
CRT → political → scientific disagreement	-0.001	0.002	-0.44	.663	-.001
CRT → media → scientific disagreement	< 0.001	0.004	0.08	.938	< .001
Total CRT effect on scientific disagreement on CC	0.15	0.06	2.37	.018	.09

Among the control variables, sociopolitical orientation exhibited as a statistically significant negative (though rather weak) relationship with the acceptance of anthropogenic CC, while religiosity exhibited a statistically significant (though weak) negative relationship with the perception of the scientific consensus on CC (Figure 3). Basic sociodemographic characteristics (except for age in the case of anthropogenic CC beliefs) did not exhibit statistically significant relationships with either of the outcome variables (Figure 3).

Discussion

The good thing is, as the results of this study suggest, that Croatian citizens do seem to acknowledge that human activity is the primary cause of recent CC. Cumulatively, just over 88% of them agree to some degree that human activity is the primary cause of global warming. Moreover, more than 90% agree that human-caused CC will increase the frequency of hurricanes, floods, and droughts and that the average global temperature has been continuously rising since the Industrial Revolution. Additionally, almost 73% of them disagree that all recent CC is primarily caused by the Sun (see Figure 1 and https://osf.io/2yndf/). These findings, combined with other data suggesting that Croatian citizens generally perceive CC as a serious problem (e.g., European Commission, 2019, 2021; Pavlović et al., in press), are encouraging.

On the other hand, we can observe the results from a different angle—around 10% of citizens are sceptical to some extent about human influence on CC. Also, other findings suggest that (Cik, 2021; Pavlović et al., in press) Croatian citizens (similar to the average European citizen) are uncertain about the anthropogenic causes of CC versus the natural processes, when the two are pitted against each other. Moreover, the results of this study show that 63% of them perceive that there is significant disagreement among scientists on the issue of anthropogenic CC, highlighting a potential gap in public understanding of scientific consensus. As noted in the introduction and evidenced by relatively weak (to medium, see Funder & Ozer, 2019; Gignac & Szodorai, 2016) associations with political orientation and religiosity, it is fortunate that the issue of CC is not greatly politicised in Croatian public discourse[3]. However, as already noted, Croatian citizens do not live in a vacuum; they are exposed to vast amounts of information and misinformation, both locally and globally. Thus, under certain circumstances, CC could become an ideologically divisive issue, for example, by framing it as an issue of economic growth versus care for the environment.

Furthermore, although Croatian citizens take issues of CC and environmental care seriously (e.g., Pavlović et al., in press), other findings show that other societal concerns often take precedence, such as economy, healthcare, poverty, crime, and alike (e.g., Cik, 2021). This suggests that they may fail to see the interconnectedness and complexity of these issues or the far-reaching consequences of CC. Additionally, research also indicates that Croatian citizens express relatively low personal responsibility for mitigating CC (Pavlović et al., in press). In interpreting these findings, institutional trust should also be considered. While Croatian citizens express relatively high trust in science and scientists (see also Tonković et al., 2023; Verboord et al., 2023), research consistently shows low levels of trust in political institutions (e.g., Franc et al., 2020; Maglić, 2023b). Given that CC mitigation requires coordinated structural changes, it is unreasonable to expect it to be tackled effectively solely at the national level, let alone at the individual level (e.g., reducing personal carbon footprints, recycling). In this context, political distrust could pose a significant obstacle—if political institutions lack credibility, citizens may be less willing to support or comply with climate policies and may even resist them, perceiving them as unfair, ineffective, or motivated by ulterior interests.

[3] Especially compared to the effect of political orientation on anthropogenic CC beliefs typically observed on participants from the USA (e.g., Drummond & Fischhoff, 2017; Pennycook et al., 2020, 2023).

Concerning the role of analytical thinking, the main findings of the present research highlight the importance of an AOT disposition for understanding beliefs about the causes of CC.

The results of the first tested model (nominally) confirmed three out of the four hypotheses. Specifically, the AOT was shown to be a positive predictor of acceptance of anthropogenic CC (H1a) and the perception of scientific consensus on the issue (H1b), although the effect on the latter was rather weak. The effect of the CRT was also weak and in line with the expectation in the case of the perception of scientific consensus on anthropogenic CC (H2b). However, in the case of acceptance of anthropogenic CC, the CRT exhibited a weak predictive effect but *in the opposite direction* than what was hypothesised (H2a).

Here, it is important to note several things. Firstly, the two outcome measures differ in the number of manifest variables used to measure them. Specifically, acceptance of anthropogenic CC is measured using a four-item scale, whereas perception of scientific consensus on anthropogenic CC is assessed using a single-item measure. Although it is not uncommon that single-item measures are used to tap CC beliefs (e.g., Drummond & Fischhoff, 2017; Pennycook et al., 2020, 2023), this may have contributed to the observed differences in the predictive effects of the AOT on the two outcome variables. Also, the relationship of the two outcome variables could be tested within different models, such as the Gateway Belief Model (see Van der Linden, 2021), which postulates perceived consensus on CC as an antecedent of CC beliefs (and support for public action).

Furthermore, as mentioned in the introduction, both the CRT and AOT, as indicators of analytic thinking, have been shown to predict lower susceptibility to a wide range of epistemologically unfounded beliefs. The two measures—a performance-based and a self-assessment measure—although moderately to strongly correlated (e.g., Bronstein et al., 2019; Haran et al., 2013; Pennycook et al., 2022)—could also be capturing different aspects of analytic reasoning. Namely, the established difference in the predictive effect of the AOT and CRT could be due to the AOT tapping not only the quantity, but also the quality of thinking, while the CRT capturing (in addition to cognitive capacity) a reflectiveness/impulsivity disposition (Baron, 2019; Baron et al., 2015, 2023; see also Erceg et al., 2020; and Bago & De Neys, 2020; Raoelison et al., 2020 for a discussion of logical intuitions).

Regarding the weak predictive effect of the CRT on acceptance of anthropogenic CC, which is in the opposite direction of what was hypothesised (H2a), this does not imply that cognitive reflection inherently reduces acceptance of anthropogenic CC. Instead, it might reflect its complex relationship with the AOT and their shared variance. Namely, the CRT shows negligible bivariate correlation with acceptance of anthropogenic CC (Table 1), but when the CRT and AOT are entered together in the model (Figure 2), the CRT becomes a negative (albeit weak) predictor of anthropogenic CC, and a positive (albeit weak) predictor of the perception of scientific consensus. This suggests a potential suppressor effect[4], which should be explored in future research more fully.

Furthermore, the AOT was expected to positively predict trust in science and scientists and, by extension, CC beliefs, which was tested in the mediation model. Given that the AOT reflects principles of sound reasoning and decision-making (Baron, 2019; Baron et al., 2023; Pennycook et al., 2020), individuals with higher AOT tendencies should be more adept at fairly evaluating evidence and assessing the trustworthiness of information sources. In line with H3, the results show that individuals higher in AOT place greater trust in science and scientists (possibly because they could more easily identify more credible sources regarding a scientific issue). At the same time, in line with H4a, science trust is confirmed as a positive predictor of the acceptance of anthropogenic CC but a rather weak positive predictor of the perception of scientific con-

[4] Additional analyses (https://osf.io/2yndf/) showed that when the CRT is entered in the model as the sole predictor, its predictive effects on the acceptance of anthropogenic CC are negligible, while its effect on the perception of scientific consensus is positive. When only the AOT is entered in the model as the sole predictor of the two outcomes, its predictive effects on the two outcomes are positive, albeit weaker in the case of the perception of scientific consensus.

sensus (H4b). Thus, higher trust in science and scientists, who have been in the vast majority consistently messaging overwhelming evidence of anthropogenic CC, is positively associated with the acceptance of anthropogenic CC (see for example Bogert et al., 2024), yet weakly with the perception of scientific consensus.

Finally, consistent with H5a, the results confirmed that science trust partially mediated the effect of the AOT on acceptance of anthropogenic CC (see also Cohen et al., 2022). However, the indirect effect of the AOT via science trust, although statistically significant, was negligible (H5b) and the total effect was rather weak. In addition to the already mentioned limitation of using a single-item measure as the outcome variable of the perception of scientific consensus, it is important to note further limitations. Specifically, the item on scientific consensus itself was not completely explicit about which scientists, i.e., the precise fields scientists who disagree belong to, which could have led the respondents to interpret the item differently. Also, science trust was measured with a single item that conflates trust in science with trust in scientists, which may have also introduced noise. Otherwise, the weak predictive effects of the AOT might be (partially) due to individuals high in AOT expect scientists, especially from different domains, to disagree as part of the usual scientific process, i.e., may value scientific disagreement (to an extent) as normative. This perspective could lead them to underestimate the degree of consensus on CC, even when it exists. For these individuals, the perceived consensus among scientists, possibly from different scientific domains, might signal dogmatism rather than robust agreement, attenuating the predictive power of the AOT.

Regarding the exploratory part of the study related to the role of two other dimensions of institutional trust, the results showed that media trust is a positive independent predictor of both acceptance of anthropogenic CC and the perception of scientific consensus. This effect was similar to that of science trust, although weaker in the case of acceptance of anthropogenic CC. Again, using a single-item measure in the case of media trust is far from optimal since different media sources vary in their credibility which should be addressed in future research.

Research generally suggests a positive relationship between political trust and CC concern, as well as support for various climate and environmental policies (e.g., Ejaz et al., 2024; Fairbrother et al., 2021; Kulin & Johansson Sevä, 2021; see also Fairbrother et al., 2019). In this study, however, political trust, though not related to anthropogenic CC belief on a bivariate level (Table 1), was predictive of lower acceptance of anthropogenic CC in the mediation model (Figure 3). This might suggest a potential suppressor effect[5], which should again be explored in future research more fully. But overall, the results show that the three dimensions are related to each other, sharing a common variance, but differentially predict CC beliefs. When controlling for this shared institutional variance, political trust reveals a negative relationship with the acceptance of anthropogenic CC. In this context, individuals who show higher levels of political trust are more likely to disagree that anthropogenic CC is happening. One can only speculate why this might be the case in the Croatian context. A right-wing coalition (in differing compositions) led by the HDZ has been governing Croatia since 2016. Although the government nominally acknowledges anthropogenic CC, it often promotes narratives tied to national, conservative, and nationalistic interests, thus implicitly downplaying the climate crisis as a global challenge, secondary to domestic issues (or even as an agenda imposed by external actors, e.g., the European Union). Consequently, for citizens who trust these institutions, this may contribute to CC scepticism. This is, clearly, just one tentative possibility, which could be further explored by testing whether ideological alignment moderates the relationship between political trust and CC scepticism (see also Fairbrother et al., 2019 for an interaction effect of political trust CC beliefs on a climate policy support).

Finally, additional limitations of this study should be explicitly mentioned and considered when

[5] Additional analyses (https://osf.io/2yndf/) showed that when political trust is entered in the model as the sole predictor, its predictive effects on the two outcome variables are negligible. However, when either of the two other dimensions of institutional trust were entered in the model alongside political trust as predictors of the two outcomes, the predictive effect of political trust on acceptance of anthropogenic CC became negative.

interpreting the results. Firstly, the study carries all the limitations of a correlational cross-sectional design, which does not allow for causal conclusions about the observed relationships. Secondly, within the dual-process framework, other measures of cognitive sophistication, such as cognitive abilities, different cognitive styles, scientific reasoning and scientific literacy, conspiracy beliefs, and other measures of contaminated mindware (see Rizeq et al., 2021) and their interplay could play a significant role when considering CC beliefs. Additionally, as mentioned, one of the outcome variables, i.e., perception of scientific consensus, and constructs of media and science trust, were measured using single-item measures, which may not fully capture their complexity and multidimensionality and might even introduce noise due to lack of specificity.

Conclusion

A substantial majority of Croatian citizens accept that human activities are the primary cause of recent CC. At the same time, 63% of individuals still believe, to some extent, there is significant disagreement among scientists on this issue.

In the tested SEM models, the AOT emerged as one of the strongest positive predictors of pro-scientific beliefs towards CC, highlighting its role in fostering evidence-based beliefs. The AOT also predicted science trust, which partially mediated its effect on CC beliefs. However, the effects of the AOT on perceiving the scientific consensus on anthropogenic CC were rather weak with little evidence of mediation via trust in science and scientists. In contrast, the CRT showed weak and less consistent effects on CC beliefs compared to the AOT.

Science trust exhibited a positive relationship with pro-scientific CC beliefs. Media trust showed a modest positive relationship, while political trust was negatively linked to the acceptance of anthropogenic CC. The relationships between each of the three dimensions of institutional trust and the perception of scientific consensus on anthropogenic CC were very small to negligible.

Acknowledgments

This research was carried out as part of the project “DISINFO KLIMA”, funded by European Union programme NextGeneration EU (01/108-73/23_2519-6) and based on data collected within a doctoral dissertation supported by the Excellence Scholarship from the Faculty of Humanities and Social Sciences, University of Zagreb.

References

Bago, B., & De Neys, W. (2020). Advancing the specification of dual process models of higher cognition: A critical test of the hybrid model view. *Thinking & Reasoning*, *26*(1), 1–30. https://doi.org/10.1080/13546783.2018.1552194

Baron, J. (2019). Actively open-minded thinking in politics. *Cognition*, *188*, 8–18. https://doi.org/10.1016/j.cognition.2018.10.004

Baron, J., Isler, O., & Yilmaz, O. (2023). Actively open-minded thinking and the political effects of its absence. U: V. C., Ottati I C., Stern (ur.), *Divided: Open-mindedness and dogmatism in a polarized world* (str. 162–182). Oxford University Press. https://doi.org/10.1093/oso/9780197655467.001.0001

Baron, J., Scott, S., Fincher, K., & Metz, S. E. (2015). Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition*, *4*(3), 265–284. https://doi.org/10.1016/j.jarmac.2014.09.003

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling, 13*(2), 186–203. https://doi.org/10.1207/s15328007sem1302_2

Binnendyk, J., & Pennycook, G. (2022). Intuition, reason, and conspiracy beliefs. *Current Opinion in Psychology*, *47*, 101387. https://doi.org/10.1016/j.copsyc.2022.101387

Bogert, J. M., Buczny, J., Harvey, J. A., & Ellers, J. (2024). The effect of trust in science and media use on public belief in anthropogenic climate change: A meta-analysis. *Environmental Communication*, *18*(4), 484–509. https://doi.org/10.1080/17524032.2023.2280749

Bowes, S. M., Costello, T. H., & Tasimi, A. (2023). The conspiratorial mind: A meta-analytic review of motivational and personological correlates. *Psychological Bulletin, 149*(5–6), 259–293. https://doi.org/10.1037/bul0000392

Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G., & Cannon, T. D. (2019). Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *Journal of Applied Research in Memory and Cognition*, *8*(1), 108–117. https://doi.org/10.1016/j.jarmac.2018.09.005

Broz, T. (2024). *Fabricating doubt and persecuting science: Analysis of misinformation about the climate crisis in Croatia. Faktograf*. Faktograf – Association for an Informed Public. https://faktograf.hr/wp-content/uploads/2024/04/faktograf-analysis-of-misinformation.pdf

Cik, T. (2021). *Izgubljeno desetljeće: Stavovi i mišljenja o zaštiti okoliša, klimatskim promjenama i energetskoj tranziciji u Republici Hrvatskoj: Izvještaj u sklopu projekta: METAR do bolje klime (Mreža za edukaciju, tranziciju, adaptaciju i razvoj)*. Projektni partneri projekta METAR do bolje klime http://idiprints.knjiznica.idi.hr/1043/1/Izgubljeno%20desetlje%C4%87e.pdf

Cohen, A. S., Lutzke, L., Otten, C. D., & Árvai, J. (2022). I think, therefore I act: The influence of critical reasoning ability on trust and behavior during the COVID-19 pandemic. *Risk Analysis, 42*, 1073–1085. https://doi.org/10.1111/risa.13833

Cologna, V., Kotcher, J., Mede, N. G., Besley, J., Maibach, E. W., & Oreskes, N. (2024). Trust in climate science and climate scientists: A narrative review. *PLOS Climate*, *3*(5), e0000400. https://doi.org/10.1371/journal.pclm.0000400

Cushman, F. (2020). Rationalization is rational. *Behavioral and Brain Sciences*, *43*, e28. https://doi.org/10.1017/S0140525X19001730

De Neys, W. (2014). Conflict detection, dual processes, and logical intuitions: Some clarifications. *Thinking & Reasoning*, *20*(2), 169–187. https://doi.org/10.1080/13546783.2013.854725

de Zúñiga, H. G., Ardèvol-Abreu, A., Diehl, T., Patiño, M. G., & Liu, J. H. (2019). Trust in institutional actors

across 22 countries. Examining political, science, and media trust around the world. *Revista Latina de Comunicación Social*, (74), 237–262. https://nuevaepoca.revistalatinacs.org/index.php/revista/article/view/279

Drummond, C. & Fischhoff, B. (2017). Individuals with greater science literacy and education have more polarized beliefs on controversial science topics. *Proceedings of the National Academy of Sciences*, *114*(36), 9587–9592. https://doi.org/10.1073/pnas.1704882114

Ejaz, W., Altay, S., Fletcher, R., & Nielsen, R. K. (2024). Trust is key: Determinants of false beliefs about climate change in eight countries. *New Media & Society*. https://doi.org/10.1177/14614448241250302

Epskamp, S. (2022). *semPlot: Path diagrams and visual analysis of various SEM packages' output*. [R package version 1.1.6]. https://CRAN.R-project.org/package=semPlot

Erceg, N., Galić, Z., & Ružojčić, M. (2020). A reflection on cognitive reflection–testing convergent/divergent validity of two measures of cognitive reflection. *Judgment and Decision Making*, *15*(5), 741–755. https://doi.org/10.1017/S1930297500007907

European Commission (2019). *Special Eurobarometer 490/Wave EB91.3: Climate change*. London, UK: Kantar. Prepared at the request of the European Commission. Retrieved December 29, 2022, from https://climate.ec.europa.eu/system/files/2019-09/report_summary_2019_en.pdf

European Commission (2021). *Special Eurobarometer 513/Wave EB95.1: Climate change*. London, UK: Kantar. Prepared at the request of the European Commission. Retrieved December 29, 2022, from https://webgate.ec.europa.eu/ebsm/api/public/deliverable/download?doc=true&deliverableId=75838

European Commission, Secretariat-General (2023). *Democracy: report*. Publications Office of the European Union. https://data.europa.eu/doi/10.2792/455273

European Environment Agency (2017). *Climate change, impacts and vulnerability in Europe 2016: An indicator-based report* (EEA Report No. 1/2017). European Environment Agency. https://www.eea.europa.eu/publications/climate-change-impacts-and-vulnerability-2016

European Social Survey (2020). *ESS10 Source Questionnaire HR.* London: ESS ERIC Headquarters c/o City, University of London. Retrieved December 29, 2022, from https://stessrelpubprodwe.blob.core.windows.net/data/round10/fieldwork/croatia/ESS10_questionnaires_HR.pdf

Fairbrother, M. (2017). Environmental attitudes and the politics of distrust. *Sociology Compass*, *11*(5), e12482. https://doi.org/10.1111/soc4.12482

Fairbrother, M., Arrhenius, G., Bykvist, K., & Campbell, T. (2021). Governing for future generations: How political trust shapes attitudes towards climate and debt policies. *Frontiers in Political Science*, *3*, 656053. https://doi.org/10.3389/fpos.2021.656053

Fairbrother, M., Sevä, I. J., & Kulin, J. (2019). Political trust and the relationship between climate change beliefs and support for fossil fuel taxes: Evidence from a survey of 23 European countries. *Global Environmental Change*, *59*, 102003. https://doi.org/10.1016/j.gloenvcha.2019.102003

Fawzi, N., Steindl, N., Obermaier, M., Prochazka, F., Arlt, D., Blöbaum, B., ... & Ziegele, M. (2021). Concepts, causes and consequences of trust in news media–a literature review and framework. *Annals of the International Communication Association*, *45*(2), 154–174. https://doi.org/10.1080/23808985.2021.1960181

Franc, R., Maglić, M., & Sučić, I. (2020). Političko (ne)povjerenje kao odrednica glasanja i sklonosti protestnim oblicima političkog sudjelovanja [Political (dis) trust as a determinant of voting and protest participation]. *Revija za sociologiju*, *50*(3), 381–406. https://doi.org/10.5613/rzs.50.3.3

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Funder, D. C., & Ozer, D. J. (2019). Evaluating Effect Size in Psychological Research: Sense and Nonsense. Advances in Methods and Practices in Psychological Science, *2*, 156–168. https://doi.

org/10.1177/2515245919847202

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Pantheon Books.

Haran, U., Ritov, I. i Mellers, B. A. (2013). The role of actively open-minded thinking in information acquisition, accuracy, and calibration. *Judgment and Decision Making*, *8*(3), 188–201. http://doi.org/10.1017/S1930297500005921

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). *semTools: Useful tools for structural equation modeling.* [R package version 0.5-6]. https://CRAN.R-project.org/package=semTools

Ladd, J. M. (2012). *Why Americans hate the media and how it matters.* Princeton University Press

Lewandowsky, S., Ecker, U. K., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of applied research in memory and cognition*, *6*(4), 353–369. https://doi.org/10.1016/j.jarmac.2017.07.008

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods, 48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Lynas, M., Houlton, B. Z., & Perry, S. (2021). Greater than 99% consensus on human caused climate change in the peer-reviewed scientific literature. *Environmental Research Letters*, *16*(11), 114005. https://doi.org/10.1088/1748-9326/ac2966

Maglić, M. (2023a). *Analitičko rasuđivanje i uvjerenja koja proizvode polarizaciju* [Analytical reasoning and socially polarizing beliefs] (Doctoral dissertation, University of Zagreb. Faculty of Humanities and Social Sciences. Department of Psychology). https://doi.org/10.17234/diss.2023.102859

Maglić, M. (2023b). Trust issues and Suspicious Minds? Political Distrust as a Determinant of Endorsement of Conspiracy Theories: Evidence from Multiple International Datasets. *Proceedings of the XXIX scientific conference Empirical studies in psychology*. Institute of Psychology, Laboratory for Experimental Psychology, Faculty of Philosophy, University of Belgrade. http://empirijskaistrazivanja.org/wp-content/uploads/2023/09/EIP23_proceedings.pdf

Mercier, H. i Sperber, D. (2017). *The enigma of reason*. Harvard University Press.

Mirhoseini, M., Early, S., El Shamy, N., & Hassanein, K. (2023). Actively open-minded thinking is key to combating fake news: A multimethod study. *Information & Management*, *60*(3), 103761. https://doi.org/10.1016/j.im.2023.103761

Myers, K. F., Doran, P. T., Cook, J., Kotcher, J. E., & Myers, T. A. (2021). Consensus revisited: Quantifying scientific agreement on climate change and climate expertise among Earth scientists 10 years later. *Environmental Research Letters*, *16*(10), https://doi.org/10.1088/1748-9326/ac2774

Kulin, J., & Johansson Sevä, I. (2021). Who do you trust? How trust in partial and impartial government institutions influences climate policy attitudes. *Climate Policy*, *21*(1), 33–46. https://doi.org/10.1080/14693062.2020.1792822

Nisbet, E. C., Cooper, K. E., & Garrett, R. K. (2015). The partisan brain: How dissonant science messages lead conservatives and liberals to (dis) trust science. *The ANNALS of the American Academy of Political and Social Science*, *658*(1), 36-66. https://doi.org/10.1177/0002716214555474

Oldrati, V., Patricelli, J., Colombo, B., & Antonietti, A. (2016). The role of dorsolateral prefrontal cortex in inhibition mechanism: A study on cognitive reflection test and similar tasks through neuromodulation. *Neuropsychologia*, *91*, 499–508. https://doi.org/10.1016/j.neuropsychologia.2016.09.010

Pavlović, T., Maglić, M., & Franc, R. (in press). Differences between major Croatian regions in climate change attitudes. In *Mediterranean Issues, Book 5.* Zagreb: Institute of Social Sciences Ivo Pilar, VERN University, the Scientific Council for Anthropogical Research, Croatian Academy of Sciences and Arts

(HAZU). https://www.mic-vis.eu/Conference/BookOfPapers

Pennycook, G. (2023). A framework for understanding reasoning errors: From fake news to climate change and beyond. *Advances in Experimental Social Psychology*, *67*, 131–208. https://doi.org/10.1016/bs.aesp.2022.11.003

Pennycook, G., Bago, B., & McPhetres, J. (2023). Science beliefs, political ideology, and cognitive sophistication. *Journal of Experimental Psychology: General, 152*(1), 80–97. https://doi.org/10.1037/xge0001267

Pennycook, G., Cheyne, J. A., Koehler, D. J., & Fugelsang, J. A. (2020). On the belief that beliefs should change according to evidence: Implications for conspiratorial, moral, paranormal, political, religious, and science beliefs. *Judgment and Decision Making*, *15*(4), 476–498. https://doi.org/10.1017/S1930297500007439

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015a). Everyday consequences of analytic thinking. *Current Directions in Psychological Science*, *24*(6), 425–432. https://doi.org/10.1177/0963721415604610

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015b). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, *80*, 34–72. http://dx.doi.org/10.1016/j.cogpsych.2015.05.001

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2015c). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision Making*, *10*(6), 549–563. https://doi.org/10.1017/S1930297500006999

Pechar, E., Bernauer, T., & Mayer, F. (2018). Beyond political ideology: The impact of attitudes towards government and corporations on trust in science. *Science Communication*, *40*(3), 291-313. https://doi.org/10.1177/1075547018763970

Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making. 29*(5), 453–469. https://doi.org/10.1002/bdm.1883

Raoelison, M., Thompson, V. A., & De Neys, W. (2020). The smart intuitor: Cognitive capacity predicts intuitive rather than deliberate thinking. *Cognition*, *204*, 104381. https://doi.org/10.1016/j.cognition.2020.104381

Revelle, W. (2023). *psych: Procedures for Personality and Psychological Research* [R package version 2.3.3]. Northwestern University, Evanston, Illinois, USA. https://CRAN.R-project.org/package=psych

Rizeq, J., Flora, D. B., & Toplak, M. E. (2021). An examination of the underlying dimensional structure of three domains of contaminated mindware: Paranormal beliefs, conspiracy beliefs, and anti-science attitudes. *Thinking & Reasoning*, *27*(2), 187–211. https://doi.org/10.1080/13546783.2020.1759688

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Shehata, A., & Strömbäck, J. (2022). Media use and societal perceptions: The dual role of media trust. *Media and Communication*, *10*(3), 146-157. https://doi.org/10.17645/mac.v10i3.5449

Stanovich, K. E. (2011). *Rationality and the* reflective *mind*. Oxford University Press

Stanovich, K. E. (2018). Miserliness in human cognition: The interaction of detection, override, and mindware. *Thinking & Reasoning*, *24*(4), 423–444. https://doi.org/10.1080/13546783.2018.1459314

Stanovich, K. E. & Toplak, M. E. (2023). Actively open-minded thinking and its measurement. *Journal of Intelligence*, *11*(2), 27. https://doi.org/10.3390/jintelligence11020027

Stanovich, K. E., West, R. F. i Toplak, M. E. (2016). *The rationality quotient: Toward a test of rational thinking*. MIT Press.

Strömbäck, J., Tsfati, Y., Boomgaarden, H., Damstra, A., Lindgren, E., Vliegenthart, R., & Lindholm, T. (2020). News media trust and its impact on media use: Toward a framework for future research. *Annals of*

the International Communication Association, *44*(2), 139–156. https://doi.org/10.1080/23808985.2020.1755338

Tonković, M., Vranić, A., & Blanuša, N. (2023). Bihevioralni korelati teorija zavjere o COVID-u 19. U: N. Blanuša i A. Brakus (Eds.), *COVID-19 dezinformacije i teorije zavjera u Hrvatskoj* (pp. 137–152). Gong.

Thomson, K. S. & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, *11*(1), 99–113. https://doi.org/10.1017/S1930297500007622

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning, 20*(2), 147–168. http://dx.doi.org/10.1080/13546783.2013.844729

Tsfati, Y., & Cappella, J. N. (2003). Do people watch what they do not trust? Exploring the association between news media skepticism and exposure. *Communication research*, *30*(5), 504-529. https://doi.org/10.1177/0093650203253371

Tsfati, Y., Strömbäck, J., Lindgren, E., Damstra, A., Boomgaarden, H. G., & Vliegenthart, R. (2022). Going beyond general media trust: An analysis of topical media trust, its antecedents and effects on issue (mis) perceptions. *International Journal of Public Opinion Research*, *34*(2), edac010. https://doi.org/10.1093/ijpor/edac010

Van der Linden, S. (2021). The Gateway Belief Model (GBM): A review and research agenda for communicating the scientific consensus on climate change. *Current Opinion in Psychology*, *42*, 7-12. https://doi.org/10.1016/j.copsyc.2021.01.005

Verboord, M., Janssen, S., Kristensen, N. N., & Marquart, F. (2023). Institutional trust and media use in times of cultural backlash: A cross-national study in nine European countries. *The International Journal of Press/Politics.* https://doi.org/10.1177/19401612231187568

Vitali Čepo, D. (2021). *Klimatske promjene u Hrvatskoj.* CO2go-Priče za klimatsku akciju. https://www.znanost-klima.org/wp-content/uploads/2021/12/Klimatske-promjene-u-Hrvatskoj.pdf

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (2nd ed.). New York, NY: Springer-Verlag. https://ggplot2.tidyverse.org/

10

Metacognitive Monitoring in Wason Selection Task: The Influence of Content Abstractness, Conditional Type, and Social Content

Pavle Valerjev
Department of Psychology, University of Zadar
valerjev@unizd.hr

Abstract

In meta-reasoning studies, solvers are asked to make a metacognitive judgment about the correctness of their reasoning. It is often found that solvers have no insight into the actual accuracy of their responses and that they make confidence judgments based on cues such as the fluency of the response. According to the dual-process theory of reasoning, responses to reasoning tasks generated by heuristic processes (Type 1) are automatic, quick, and easy, while responses generated by analytic processes (Type 2) are often slower, require mental effort, and are deliberate. Participants tend to have more confidence in their Type 1 responses than in their type 2 responses, regardless of their accuracy. The Wason selection task involves a form of conditional reasoning. Typical results show that accuracy in the abstract version of the task is very low. This can be explained by cognitive biases such as confirmation and matching bias, which are based on heuristic processes. The aim of this study was to investigate how variations in content abstractness, conditional type, and social content affect judgments of confidence in Wason tasks. 128 participants took part in the 2x2x2 experiment. They solved variants of Wason tasks and rated their confidence after each response on a 0-100% scale. The results showed a statistically significant effect of social content (the participants were more confident in non-social content tasks than in social content tasks) and abstractness (more confidence in abstract content tasks than in concrete tasks). It appears that heuristically generated responses promote higher metacognitive scores. This result is consistent with the general views of the meta-reasoning framework and dual-process theories of reasoning.

Keywords: *Wason selection task, metacognition, meta-reasoning, dual-process theory*

Introduction

Conditional reasoning is one of the most important areas of research in the psychology of reasoning. Researchers are interested in it because it involves a number of interesting effects that can be studied in tasks as simple as conditionals. The classical conditional reasoning task consists of two premises and a conclusion. The first premise is a conditional sentence with the form If P, then Q.

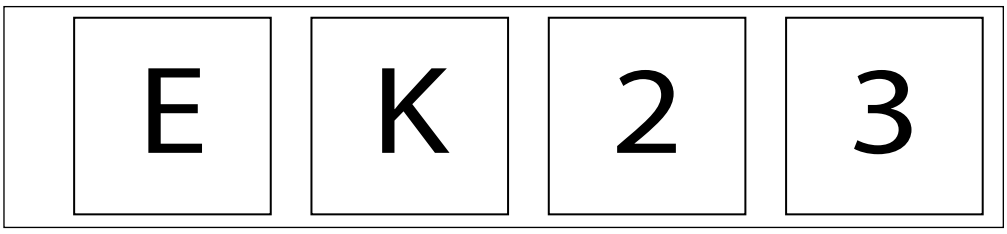
P is a simple proposition called the antecedent, and Q is a simple proposition called the consequent. The second premise and the conclusion are simple propositions that refer to the affirmation or negation of the antecedent and the consequent. This results in four possible forms of conditional tasks (see Table 1).

Table 1
Basic conditional task forms

Name	Conditional premise	Second premise	Conclusion	Validity
Affirmation of antecedent (Modus ponens)	If P, then Q	P	Q	Valid
Affirmation of consequence	If P, then Q	Q	-	No valid conclusion follows
Negation of antecedent	If P, then Q	Not P	-	No valid conclusion follows
Negation of consequence (Modus tollens)	If P, then Q	Not Q	Not P	Valid

In addition to the simple conditional reasoning tasks with two premises, one of the most common conditional reasoning tasks is the Wason selection task (Wason, 1966, 1968). Alternative names for this problem are the Card selection task and the Four-card problem. This task has been studied for a long time and is an important tool in research on deductive and inductive reasoning, in research on scientific reasoning, and in research on cognitive biases. The classic version of the task (Wason, 1966, 1968) is an abstract version with a rule about the occurrence of letters and numbers on the cards. It reads as follows:

TASK DESCRIPTION:
Each card in the set has a letter on one side and a number on the other. Four cards are chosen at random from the set and placed on the table. For example:



The experimenter then presents a hypothesis (conditional rule)

RULE:
If there is a vowel on one side of the card, then there is an even number on the other side of the card.

TASK FOR THE PARTICIPANT:
Select all the cards, and only those that need to be turned over, to find out if the hypothesis of the four cards on the table is valid or invalid.

Note that according to the conditional rule *If P, then Q*, the four cards in the example above represent simple propositions P, not-P, Q, and not-Q. In this version of the task, the correct response is to select two cards: P and not-Q. For the rule to be valid, there must be an even number on the back of card E and no vowel on the back of card 3. The other two cards (not-P and Q) can have any letter or number on the back, so that they are not relevant to the validity of the rule.

However, the majority of participants select the incorrect combination of cards in this version of the task. They choose cards P and Q, but also other incorrect combinations. In the work of Wason and Johnson-Laird (1972), only 4% of participants chose the logically correct combination of cards. Among the incorrect responses, P Q was the most common, given in 46% of trials. When participants were differentially instructed to choose exactly two cards (and not any number of them), the dispersion of their responses decreased, but there was still an even stronger tendency to choose the typical incorrect response P Q (74%) over the correct response P not-Q (14%) (Valerjev, 2000; Valerjev & Pedisić, 2001). Initially, Wason (1968) explained this universal error as confirmation bias. Confirmation bias is the tendency to consider only the evidence and information that confirms our knowledge and beliefs and to ignore the evidence and information that does not confirm them. In other words, people tend to test their hypothesis only positively (verify it by choosing the P-card) and not try to test it negatively (falsify it by choosing the not-Q-card). Later, Evans (1984, 1989) attempted to attribute this effect to matching bias—the tendency to choose the cards that match the information contained in the question. Both biases are a typical result of heuristic reasoning processes that simplify the task and tend to speed up its solution. For a more detailed insight into the study of confirmation bias using card selection tasks and other reasoning tasks, see Evans (1989) and Valerjev (2024).

There are several variants of the Wason selection task, and some of them have led to a dramatic change in responses. One of the best-known is the manipulation through the content of the task. Instead of using abstract content, Griggs and Cox conducted a series of experiments using concrete (and social) content (see Griggs & Cox, 1982; Cox and Griggs, 1982). They used the conditional rule, which is more concrete, familiar, and social in nature: *If a person drinks beer, then the person must be over 19 years old*. The four cards presented (for the P, not-P, Q, and not-Q cases) were: *Drinks beer*; *Drinks Coke*; *22 years old*; *16 years old*. Participants' task was to select the cards that had to be turned over to determine whether someone was breaking the rule or not. The correct response was the same as the abstract version of the task: P and not-Q card (*Drinks beer* and *16 years old*). The change in content led to a clear shift towards correct responses. 60% of participants gave the correct responses (Cox and Griggs, 1982), and in the other version of the concrete task, where exactly two cards had to be selected, 56% of participants were correct (Valerjev, 2000; Valerjev & Pedisić, 2001). It was also found that participants tended to represent the abstract and neutral (not-social) conditional as biconditional, and that in this type of conditional the case P Q would also be the correct one.

The increase in accuracy has been described as the effect of thematic content (Wason & Shapiro, 1971). Although it was not always possible to replicate the effect when specific rules were used (see Pollard, 1981), this effect sparked a massive debate about which theory explained it better. The theory of pragmatic reasoning schemas (Cheng & Holyoak, 1985) assumes that when reasoning with concrete rules, we draw on learned rules (pragmatic schemas) from real life, which facilitates reasoning. The innate modules theory of reasoning (Cosmides, 1989) holds that we are sensitive to the violation of the social contract implicit in the rule because this sensitivity has been a useful evolutionary adaptation of humans. Mental model theory (Johnson-Laird & Byrne, 1991; Johnson-Laird, 2001; Johnson-Laird et al., 2015) states that we can use adequate mental models to generate the counterexample, but in the abstract version, the models may be incomplete and/or represent biconditional, which leads to an incorrect response. Formal rule theories of reasoning (e.g., Braine, 1978; Rips, 1983, 1994) have a problem explaining this

effect since both abstract and concrete rules have the same logical form. According to relevance theory (Sperber et al., 1995), we use the interpretation that is pragmatically relevant in the given context, which affects differences in responses. The probabilistic approach (Oaksford & Chatter, 1994, 2007) assumes that people draw conclusions that are rational in a probabilistic sense, i.e., that they choose answers that lead to the reduction of uncertainty. According to Evans' (1984, 1989, 2006) heuristic-analytic theory, the content and type of the conditional can influence the greater or lesser activation of heuristic or analytic thinking. Heuristic thinking leads to a typical incorrect answer, and analytic thinking leads to a correct one. For a more detailed comparison of the implications of the 16 theories of reasoning in the Wason selection task, see the analysis by Ragni et al. (2018).

For this study, we have adopted the heuristic-analytic approach proposed by Evans (2006), as well as other theories that fit into the dual-process approach to reasoning (for a detailed overview of developments in this area, see De Neys, 2018). The main idea is that there are two types of processes: Type 1, which tends to be fast and intuitive, more automatic and based on heuristics, and Type 2, which is slower and more deliberate, based on reflective and analytical thinking. Type 2 processes relate to the conflict in the responses and the resolution of the conflict by re-evaluating the response. Specifically, in the Wason selection task, this would mean that rapid heuristic processes (causing confirmation and matching bias or simplification of the conditional representation to a biconditional representation) lead to the typical incorrect response (P and Q). Manipulating the task material could increase the likelihood that the analytic processes are activated, which would lead to a higher proportion of alternative responses and, among these, to correct responses (P and not-Q).

Another important theoretical background for this study is the meta-reasoning framework by Ackerman and Thompson (2015, 2017). It provides explanations for the metacognitive monitoring and control that accompany processes of reasoning and problem-solving. According to this framework, people typically use a cue utilisation approach during reasoning. This means that they usually do not have access to the cognitive process itself (at least not to the automatic parts of it), which makes it difficult to assess how good the reasoning process and its responses are. Therefore, in metacognitive monitoring, people have to rely on indirect cues that are necessary for performing metacognitive control (e.g., decisions about strategy choice, further engagement in solving, etc.). The main cues that people typically rely on are fluency, ease and speed of responses, feeling that something is wrong, recognizing a conflict, duration of the task, etc. The cue utilisation approach has been observed in various reasoning tasks. For example, in the base-rate neglect task, the conjunction fallacy task, and the covariation detection task (see Dujmović & Valerjev, 2018; Dujmović, et al., 2020; Valerjev & Dujmović, 2019), metacognitive judgments were generally found not to correlate with actual reasoning accuracy (people often tend to overestimate their efficiency), but to correlate positively with fluency and negatively with the feeling of conflict. Participants are generally more confident when responses are faster and easier to produce and less confident when they are slower and more conflictual. In addition, responses produced by the dominant heuristic alone are more fluent and faster and are therefore judged with greater confidence in accuracy. In contrast, when there is a conflict between two or more intuitions or two or more processes, the final response is slower, generated with more effort, and therefore judged with less confidence.

A similar mechanism is expected in the Wason selection task. A typical incorrect response (P Q cards) results from heuristic intuition, which is quick and easy, and should, therefore, be associated with a high degree of confidence. The correct response (P not-Q) requires analytical thinking that involves rejecting the quick intuitive response and a deeper understanding of the conditional rule. Therefore, we expect that task variants with a high proportion of P Q responses will generally have higher confidence levels than tasks that elicit more non-heuristic response alternatives (including correct P not-Q responses).

A few studies have measured metacognitive parameters in the Wason selection task. In the work of Thompson et al. (2013), participants completed a modified abstract version of the Wason selection task in a two-response paradigm. The modification required participants to decide for each card whether it was useful in deciding the validity of the conditional rule and to rate their felling of rightness (FOR) for their response. Each card was rated in two steps, once quickly and a second time giving participants enough time to re-evaluate their responses. The results showed that the responses consistent with the matching heuristic were faster, had higher FOR ratings, and were analysed less retrospectively. In another metacognitive study using a Wason selection task (Valerjev & Dujmović, 2017), participants were randomly presented with 30 Wason tasks (10 abstract, 10 concrete neutral, and 10 concrete with a social contract integrated into the conditional rule). Participants were instructed to respond as quickly as possible, which increased the strength of heuristic intuition, and resulted in participants relying almost exclusively on heuristic responding in all task variants, including the variant with a social contract. As a result, the proportion of correct responses was very low (less than 5% in all three task variants), and the proportion of incorrect heuristic P Q responses was high (55-58.67%). There was no difference between the proportions of the two response types in the different task variants. Confidence judgments were high (78.81-82.17%), and there was no difference in confidence judgments between task variants. The confidence ratings correlated with the response times so that the faster responses were generally rated with higher confidence.

Previous research has shown the expected effect of heuristic reasoning on metacognitive judgments in the Wason selection task. However, there is insufficient evidence for the change in metacognitive judgement levels associated with the shift in responses from heuristic to less heuristic reasoning in the Wason selection task. We expect that experimental manipulations in the Wason selection task that change responses from predominantly heuristic (P Q) to more analytic (P not-Q) and other non-heuristic responses would generally reduce confidence in such task variants.

Aim

The aim of the study was to investigate changes in metacognitive monitoring in the Wason selection task. To this end, confidence judgments of response accuracy were measured in selection tasks that differed in terms of abstractness, conditional type, and social content.

Method

Participants. The sample consisted of N = 128 (*Nmale* = 22) psychology students from Zadar, Rijeka, and Mostar universities. The mean age of the participants was *Mage* = 20.42 (*SDage* = 1.1; *minage* = 19; *maxage* = 24). The participants were rewarded with an appropriate research point.

Research design. The experiment was designed as a factorial 2x2x2 within-subjects experiment. Each of the three factors (abstractness, conditional type, social content) was manipulated by the two levels (abstract vs. concrete content; classical (non-deontic) conditional type vs. deontic conditional type; no social situation in content vs. social situation in content of the task). The experiment consisted of eight conditions. The recorded responses were selected cards and the confidence judgment for each response. Although the main interest was on the main factor effects, the balanced factorial design was chosen for better experimental control. By combining factor levels in the task variants, we wanted to capture broader task variants (e.g., the social effect was tested in both abstract and concrete tasks and in both deontic and non-deontic tasks).

Materials and procedure. Eight Wason selection tasks were constructed for this study. Each experimental condition was represented by one task. Table 2 lists the conditional rules used in the eight tasks. It was decided not to use more than one task for each condition to keep the procedure economical and simple. This was to avoid the undesirable uniformity of responses that occurs in experiments with too many tasks. Participants seem to give the same heuristic response for each task and ignore the experimental manipulation (see Valerjev & Dujmović, 2017) when there are many selection tasks and when there may be time pressure. The experiment was conducted in the classroom in several group sessions (before the lecture). The participants were presented with booklets with printed selection tasks. The order of the tasks in the booklets was rotated, so that the participants started with different tasks and had a different order. The participants were instructed to read the conditional rule carefully and then select the cards (out of four cards presented) to be flipped to determine whether the rule was valid or invalid. They then rated their confidence in the answer on a 0 to 100% scale. The zero value was anchored as *"no confidence; I am just guessing"*, while 100% was *"complete confidence in the correctness of the response"*. The participants then moved on to the next task. The procedure took about 10 minutes.

Table 2
Conditional rules that were used in eight variants of Wason selection task

Experimental condition	Conditional rule to be tested in the task
A-nD-nS	If a card has a vowel on one side, then it has an odd number on the other side.
A-D-nS	If there is an even number on one side, then the other side of the card must have a capital letter.
K-nD-nS	If the weather is sunny, then the bees fly.
K-D-nS	If it's rainy, then the frogs have to croak.
A-nD-S	If a person has occupation X, then they are in cahoots with people from city ZZ.
A-D-S	If a woman has an oc cupation DDD, then she must be in contact with people from block C.
K-nD-S	If the athlete uses the sports hall, then they are a member of the "Athletics" club.
K-D-S	If a person drinks beer, then they must be of legal age.

Note. A = abstract and K = concrete, nD = not deontic conditional (classical conditional) and D = deontic conditional, nS = not social content and S = social content.

Results

There are 15 ways to select four cards, resulting in 15 possible responses for each task. However, some of the combinations are pretty unlikely and uninteresting, given the typical responding patterns grounded in theories of reasoning. Two responses are essential for the aims of this study: P Q, which is very common and is an incorrect heuristic response resulting from heuristics that lead to confirmation and matching biases, and a correct P not-Q response that reflects understanding of the conditional rule and suppresses misleading heuristics and can therefore be considered an analytic response. All other responses are considered non-heuristic and were categorised as Other. Thus, there are three response categories: P Q, P not-Q, and Other.

Table 3
Responses and judgments of confidence for eight variants of Wason selection task (N=128)

Task	P Q(%)	Other (%)	P not-Q (%)	Conf. mean	Conf. SD	Conf. median	Skew.	Kurt.
AnDnS	93 (72.66)	27 (21.09)	8 (6.25)	88.68	20.25	100	-2.33	6.13
ADnS	89 (69.53)	32 (25.00)	7 (5.45)	86.27	23.45	100	-2.08	4.22
KnDnS	79 (61.72)	47 (36.72)	2 (1.56)	83.57	24.08	100	-1.78	3.12
KDnS	81 (63.28)	43 (33.59)	4 (3.13)	82.50	24.55	98.50	-1.61	2.35
AnDS	74 (57.81)	48 (37.50)	6 (4.69)	79.12	26.10	90	-1.26	0.97
ADS	74 (57.81)	47 (36.72)	7 (5.47)	77.90	28.43	90	-1.33	1.02
KnDS	64 (50)	56 (43.75)	8 (6.25)	75.56	28.43	80	-1.22	0.85
KDS	45 (35.16)	57 (44.53)	26 (20.31)	74.88	27.26	80	-1.20	0.98

Note. A = abstract and K = concrete, nD = not deontic conditional (classical conditional) and D = deontic conditional, nS = not social content and S = social content, Conf. = confidence judgment, Skew. = skewness, Kurt. = kurtosis.

The descriptive statistics of the results are shown in Table 3. It can be seen that the frequency of P Q responses, Other responses, and P not-Q responses changes in the different task variants. The highest frequency of P Q responses and the lowest frequency of Other and P not-Q responses were found for abstract/non-deontic/non-social task types. In contrast, the lowest frequency of P Q responses and the highest frequency of Others and P not-Q responses were found for the concrete/deontic/social task type. For the other task types, the frequencies for all three response categories were between the two stated frequencies.

The Cochran Q test showed that there was a statistically significant difference in the distribution of P Q responses across the eight variants of the task (Q = 89.22, df = 7, p < .001), with the highest frequencies observed for the abstract/not deontic/not social variant. The same test showed a statistically significant difference in the distribution of valid P not-Q responses across the eight task variants, with the highest frequencies observed in the concrete/deontic/social task variant (Q = 62.22, df = 7, p < .001) and in the distribution of all "Other" responses, which were more common in social content tasks than in not-social content tasks (Q = 41.13, df = 7, p < .001). The frequencies of the different response types are shown in Table 3.

The mean values and medians of the confidence judgments for eight task variants are also shown in Table 3. It can be seen that the participants were generally very confident about the accuracy of their responses, while at the same time, their accuracy was quite low. This is a clear example of overconfidence. The highest confidence judgments are paired with responses of the abstract/non-deontic/non-social task type, and the lowest confidence judgments are paired with responses of the concrete/deontic/social task type. Many participants tended to rate their confidence very highly on most tasks, including those who gave the highest rating. For the first three task types, more than half of the participants rated their responses with 100% confidence in their accuracy (see the "Confidence median" column in Table 3). For this reason, the distributions of confidence values are negatively asymmetric, which can be seen in the "Skewness" column. Kolmogorov-Smirnov normality tests showed that all distributions deviated statistically significantly from the normal distribution (all p < .01). Therefore, the non-parametric Wilcoxon matched-pairs test was chosen for the following analysis.

Three main factors with potential effects on confidence judgments were tested in the Wason rea-

soning task: abstractness, conditional type, and social content. Before the difference test, the confidence values of each task category had to be averaged for each participant (e.g., average of all non-social tasks vs. average of all social tasks). The average confidence values are shown in Table 4.

Table 4
Average confidence judgments in six task categories

Task category	Mean	SD	Median
Abstract content	82.99	20.51	90
Concrete content	79.13	22.07	86.88
Non-deontic conditional	81.73	20.35	87.5
Deontic conditional	80.39	20.16	87.5
Non-social content	85.26	19.23	95
Social content	76.86	23.30	82.5

The Wilcoxon matched-pairs test showed a statistically significant difference in average confidence ratings for the responses in the Wason selection tasks about the social content factor. Note that the effect size *r* was calculated using Rosenthal's (1994) formula. Confidence ratings obtained in tasks with non-social and social content differed ($Z = 5.63$; $N = 128$; $p < .001$; $r = 0.5$). The participants were generally more confident in their responses to non-social content tasks than social content tasks. When comparing the abstractness of task content, the participants were generally more confident in their responses for tasks with abstract content than those with concrete content ($Z = 2.86$; $N = 128$; $p = .004$; $r = 0.25$). For the type of conditional rule in the Wason selection task (non-deontic vs. deontic), the difference in overall confidence in responses was not statistically significant ($Z = 1.89$; $N = 128$; $p = .06$; $r = 0.17$).

In addition, the confidence values of the heuristic P Q responses were compared with the confidence values of all other (non-heuristic) responses for all eight tasks. This grouping of responses was the only way to compare, as the frequency of specific responses (including the correct P not-Q response) was too low for most tasks. The non-parametric Mann-Whitney U test was chosen for the tests. There was a general tendency towards higher confidence ratings for P Q responses than for all other responses in all task variants. However, the difference was statistically significant in three task variants. For the concrete/not-deontic/not-social task ($U = 1391.5$; $Z = 2.66$; $p = .008$; $r = 0.24$), concrete/not-deontic/social task ($U = 1468.5$; $Z = 2.76$; $p = .005$; $r = 0.24$), and concrete/deontic/social task ($U = 1386$; $Z = 2.40$; $p = .014$; $r = 0.21$), the confidence ratings for P Q responses were higher than the confidence ratings for all other responses. The differences were not statistically significant for the other five task variants.

Discussion

This study investigated whether the different variants of Wason selection tasks influence confidence judgments. The conditional rule manipulation included the social content factor, the abstractness factor, and the conditional type factor. The results showed that the participants were generally less confident in their responses in social content and concrete content tasks than in non-social content and abstract content tasks, while the conditional rule type (non-deontic vs. deontic) did not affect confidence. The participants tend to give fewer P Q responses (which are incorrect) in social content tasks and in concrete content tasks. In another words, the participants are more confident when inaccurate and less confident when they respond with non-heuristic responses (including the correct P not-Q response). At the specific task level, there is a tendency for the participants to make higher confidence judgments for P Q responses than for all other responses combined, at least for three task variants.

This finding is consistent with the general views of dual-process theories of reasoning (see De Neys, 2018; Evans, 2006), as well as with the explanations of the theoretical framework of meta-reasoning proposed by Ackerman and Thompson (2015, 2017). The heuristics activated in the Wason selection task led the reasoner to incorrect P Q responses that have been recognised as expressions of confirmation bias (Wason, 1968) and/or matching bias (Evans, 1984). There is overwhelming evidence that heuristic P Q responses are more frequent for abstract and non-social variants of the task than for concrete and social variants. This means that heuristic processes are stronger and more difficult to overcome in abstract and non-social variants of the task. On the other hand, it seems that concrete and social content in the task (for many different reasons—see the list of competing theories in Ragni et. al., 2018) promote overcoming the heuristic response and increase the likelihood of alternative, non-heuristic responses (including the correct P not-Q response). According to the meta-reasoning framework, metacognitive monitoring in reasoning is often based on cues rather than the actual accuracy of responses, and key cues include the fluency, ease, and speed of responses. Reasoners tend to assign higher confidence scores to heuristically generated responses because heuristically generated responses tend to be more fluent, easier, and faster than non-heuristically generated ones. Therefore, participants are generally more confident about incorrect responses when they are quick and easy, which is usually the case with heuristic responses. This finding is not uncommon and has been demonstrated in several other reasoning tasks, such as the base rate task (Dujmović & Valerjev, 2018) and the conjunction fallacy problem (Dujmović et. al, 2021), to name a few. It seems that there is a similar case with the Wason selection task.

This study has some limitations that need to be mentioned here. First, the study was not conducted in a laboratory but in classrooms using printed tasks, so it lacks rigour and control. The study also does not include measures of response times. Future studies should consist of response time measurement as it would be useful for comparing response times of different task variations. Finally, only one trial per experimental condition was used in this study (eight tasks for eight conditions), which opens up the possibility that some uncontrolled features of the tasks also influence the results. The reason for this decision was the fact that participants who have to solve too many Wason selection tasks would otherwise very soon start to give the same heuristic response (P Q) for all tasks (as shown in Valerjev & Dujmović, 2017). Even in the current study with a limited number of tasks, some participants gave the same P Q response and the same maximum confidence for all tasks. Therefore, the lesser of two evils was chosen and the participants were presented with limited tasks. To limit the number of tasks per participant, future studies could use a smaller number of main factors (perhaps only two) and/or independent samples of participants for different experimental conditions.

Conclusion

Participants generally tend to give a high proportion of incorrect P Q responses in Wason selection tasks, which is probably a combined consequence of matching bias, confirmation bias, and the tendency to simplify conditionals and represent them as biconditionals. All of these reasoning styles belong to the variants of heuristic thinking. The participants were very confident in their responses in the tasks where they gave the highest proportion of incorrect (and heuristic thinking-generated) P Q responses. The manipulation of task content reduced both confidence scores and the frequency of P Q responses. Confidence decreased statistically significantly in tasks with social and concrete content compared to tasks with non-social and abstract content. Apparently, participants are most confident in the tasks where they are most likely to give quick, heuristic (and incorrect) P Q responses without considering the alternatives. Confidence in P Q responses is higher than that in all other responses in the three variants of the task. When the content of the tasks encourages other responses (including the correct P not-Q response), participants feel less confident in the response they choose. These findings are consistent with dual-process theories of reasoning and the meta-reasoning framework.

References

Ackerman, R., & Thompson, V. A. (2015). Meta-reasoning: What can we learn from meta-memory? In A. Feeney & V. A. Thompson (Eds.), *Reasoning as Memory* (pp. 164-182). Psychology Press.

Ackerman, R., & Thompson, V. A. (2017). Meta-reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences, 21*(8), 607–617. https://doi.org/10.1016/j.tics.2017.05.004

Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review, 85*(1), 1–21. https://doi.org/10.1037/0033-295X.85.1.1

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17*, 391-416. http://dx.doi.org/10.1016/0010-0285(85)90014-3

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31*, 187–276. http://dx.doi.org/10.1016/0010-0277(89)90023-1

Cox, J. R., & Griggs, R. A. (1982). The effects of experience on performance in Wason's selection task. *Memory & Cognition, 10*(5), 496-502. https://doi.org/10.3758/BF03197653

De Neys, W. (Ed.). (2018). *Dual Process Theory 2.0*. Routledge. https://doi.org/10.4324/9781315204550

Dujmović, M., & Valerjev, P. (2018). The influence of conflict monitoring on meta-reasoning and response times in a base rate task. *Quarterly Journal of Experimental Psychology, 71*(12), 2548-2561. https://doi.org/10.1177/1747021817746924

Dujmović, M., Valerjev, P., & Bajšanski, I. (2020). The role of representativeness in reasoning and metacognitive processes: an in-depth analysis of the Linda problem. *Thinking & Reasoning, 27*(2), 161–186. https://doi.org/10.1080/13546783.2020.1746692

Evans, J. St. B. T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology, 75*, 451-468. http://dx.doi.org/10.1111/j.2044-8295.1984.tb01915.x

Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences*. Hillsdale, NJ: Lawrence Erlbaum.

Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review, 13*, 378–395. http://dx.doi.org/10.3758/BF03193858

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology, 73*, 407-420.

http://dx.doi.org/10.1111/j.2044-8295.1982.tb01823.x

Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences, 5*(10), 434-442. https://doi.org/10.1016/S1364-6613(00)01751-4

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Lawrence Erlbaum.

Johnson-Laird, P. N., Khemlani, S. S., & Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences, 19*, 201–214. http://dx.doi.org/10.1016/j.tics.2015.02.006

Klauer, K. C., Stahl, C., & Erdfelder, E. (2007). The abstract selection task: New data and an almost comprehensive model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 680-703. http://dx.doi.org/10.1037/0278-7393.33.4.680

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*, 608-631. http://dx.doi.org/10.1037/0033-295X.101.4.608

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198524496.001.0001

Pollard, P. (1981). The effect of thematic content on the 'Wason selection task'. *Current Psychological Research 1*, 21-29 (1981). https://doi.org/10.1007/BF02684422

Ragni, M., Kola, I., & Johnson-Laird, P. N. (2018). On selecting evidence to test hypotheses: A theory of selection tasks. *Psychological Bulletin, 144*(8), 779-796. https://doi.org/10.1037/bul0000146

Rips, L. J. (1983) Cognitive processes in propositional reasoning. *Psychological Review, 90*, 38-71. https://doi.org/10.1037/0033-295X.90.1.38

Rips, L. J. (1994). The psychology of proof. Cambridge, MA: MIT Press.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), The handbook of research synthesis. (pp. 231-244). New York: Russell Sage Foundation.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition, 57*, 31-95. http://dx.doi.org/10.1016/0010-0277(95)00666-M

Thompson, V. A., Evans, J. St. B. T., & Campbell, J. I. D. (2013). Matching bias on the selection task: It's fast and feels good. *Thinking & Reasoning, 19*(3–4), 431-452. https://doi.org/10.1080/13546783.2013.820220

Valerjev, P. (2000). Wasonov izborni zadatak: Stabilnost efekata pristranosti i tematskog materijala. [Wason selection task: Stability of bias and thematic material effects]. *Radovi: Razdio filozofije, psihologije, sociologije i pedagogije, 39*, 16, 97-111.

Valerjev, P. (2024). Pristranost potvrđivanja. [Confirmation bias]. In K. Damnjanović (Ed.), *Zbirka kognitivnih pristrasnosti i pogrešaka [Collection of cognitive biases and fallacies]*. (pp. 55-72) Heliks. https://doi.org/10.31212/kogn.prist.2024.vale.03

Valerjev, P., & Pedisić, A. (2001). Wasonov izborni zadatak – utjecaj upute, tipova kondicionala i tematskog materijala. [Wason Selection Task – Influence of Instruction, Conditional Types and Thematic Material]. *Radovi: Razdio filozofije, psihologije, sociologije i pedagogije, 40*, 17, 45-64.

Valerjev, P., & Dujmović, M. (2017). Metacognitive judgments during solving of Wason selection task. In N. Đapo, S. Zvizdić, A. Dautbegović, & M. Marković (Eds.), *Zbornik radova: Četvrti sarajevski dani psihologije*. (pp. 83-93). Filozofski fakultet Univerziteta u Sarajevu.

Valerjev, P., & Dujmović, M. (2019). Performance and metacognition in scientific reasoning: The covariation detection task. *Psychological Topics, 28*(1), 93-113. https://doi.org/10.31820/pt.28.1.5

Wason, P. C. (1966). Reasoning. In B. M. Foss, (Ed.), *New Horizons in psychology 1*. Penguin.

Wason, P. C. (1968). Reasoning about the rule. *Quarterly Journal of Experimental Psychology, 20*, 3, 273–281. https://doi.org/10.1080/14640746808400161

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content.* Harvard U. Press.

Wason, P. C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *The Quarterly Journal of Experimental Psychology, 23*(1), 63–71. https://doi.org/10.1080/00335557143000068

11

The Relationship Between Depressive Symptoms and Music Preferences

Luka Marija Delimar
Department of Psychology, University of Zadar
lukadelimar20@gmail.com

Pavle Valerjev
Department of Psychology, University of Zadar

Abstract

Music is important because it can influence emotions and mood and promote social relationships. However, the question of what determines a person-s music preferences, particularly in relation to mental health problems such as anxiety and depression, remains under-researched, despite the prevalence of these problems and their impact on daily life. The aim of this study was to investigate the relationship between depressive symptoms and preferences for five music preference dimensions: Intense, Classical, Contemporary, Jazzy, and Unpretentious. These five music preference dimensions comprised 17 music genres, with each music genre being categorised into the corresponding music preference dimension. The study comprised two phases: an online (n=101) and a laboratory measurement of music preferences (n=43). The participants first completed a Depression subscale (DASS-O) and rated their preferences for 17 music genres in an online survey. In the laboratory phase, a subgroup of participants from the first phase rated audio recordings of 34 music clips representing 17 music genres. In both phases of the study, depressive symptoms were positively associated with Intense music, suggesting that individuals with higher depression scores have a greater preference for Intense music (e.g., heavy metal, rock). No significant correlations were found between depressive symptoms and the other four dimensions of music preference. In addition, the participants' ratings of the music genres which they rated online differed significantly from their ratings after listening to music clips, with the participants expressing a greater preference for Classical, Intense, Jazzy, and Unpretentious music preference dimensions after listening. Gender differences were found, with men preferring Intense music and women preferring Unpretentious music. The correlation between depressive symptoms and a preference for Intense music support the assumption that music preferences, particularly Intense music genres, may reflect emotional states such as depression. The study suggests that music may play a role in understanding and coping with depressive symptoms, which in turn has implications for future research and therapeutic interventions.

Keywords: *music, music preferences, depressive symptoms, intense music*

Introduction

Music plays an important role in people's lives as it reflects their emotions, values, and conflicts while fulfilling their emotional, social, and cultural needs (Ekinci et al., 2013; Miranda & Claes, 2007). The accessibility of online music has further increased its impact on adolescents and adults (Tarrant et al., 2000). Music is also an essential part of social experiences, such as concerts, where people come together to listen to music and express their musical preferences, which can vary greatly from person to person (Vella & Mills, 2017). Furthermore, the importance of music is linked to the messages it conveys or the identification it enables, with many people viewing musical preferences as an expression of personal identity (Rentfrow & Gosling, 2003). Consequently, music has received considerable attention in cognitive psychology (Chee et al., 2024), clinical psychology (van Sprang & Haeyen, 2024), and neuroscience (Greenberg et al., 2021), but there is still a limited understanding of why people prefer certain types of music. For example, personal music preferences vary widely by genre, which may be related to personality traits (Žauhar & Levak, 2020), social identity and the sense of self (Loureiro, et al., 2024), physiological arousal (Chee et al., 2024), and other individual factors such as self-concept and cognitive ability (Rentfrow & Gosling, 2003).

The fundamental question that interested researchers was the psychological organisation of music preferences. Originally, researchers relied on music genres as a unit of measurement due to technological limitations. However, genres are poorly defined and elusive constructs that often misrepresent listeners-musical tastes (Fricke et al., 2021). To overcome these limitations, research has moved away from looking solely at music genres and focused on assessing music preferences through a careful selection of audio clips presented to participants (Rentfrow et al., 2012). Thus, Rentfrow et al. (2012) set a framework for assessing music preferences by introducing a comprehensive five-factor model called MUSIC, which includes both auditory (e.g., loud, fast, etc.) and psychological (e.g., romantic, aggressive, sophisticated, etc.) aspects of music. Furthermore, the work of Bonneville-Roussy et al. (2017) refines previous categorisations and comes closest to a compromise between a focus on musical styles and a focus on musical characteristics by providing the clearest way to develop a causal model from within a psychometrics paradigm (Hallam et al., 2009, p. 277). Their line of research based on large-scale data analysis found that preferences for musical styles can be categorised into five factors: preferences for Unpretentious, Contemporary, Classical, Jazzy, and Intense music (Bonneville-Roussy et al., 2017). Each music preference dimension includes specific music genres represented by individual music clips, and their model has similarities with the aforementioned MUSIC model by Rentfow et al. (2012). For example, Unpretentious music is described as relaxing, slow, and romantic (e.g., soul, R&B); Classical music is artistic, complex, and intelligent (e.g., opera); and Intense music is loud, energetic, and thrilling (e.g., rock, punk, heavy metal) (Žauhar & Levak, 2020). Research by Rentfrow et al. (2012) also found that music preferences are driven by specific musical attributes and characteristics in the music (e.g., sad, complex, exciting, etc.) and that measuring participants' level of preference for individual music clips allows for a more nuanced assessment beyond genre labels as it does not require familiarity with specific music styles.

In addition to personality traits (Vella & Mills, 2017), music preferences have also been linked to internalising problem behaviours such as depression (Ekinci et al., 2013; Miranda & Claes, 2007; Ter Bogt et al., 2021b). Depression is a common and debilitating condition characterized by mood disturbances and a reduced capacity to enjoy positive experiences (anhedonia). Depressive symptoms are also associated with a pattern of stronger emotional and cognitive responses to negative stimuli and a reduction in positive emotions. Furthermore, early models of emotional dysfunction in depression emphasised the importance of inadequate physiological arousal; for example, depressed individuals show reduced and selective reactivity to positive stimuli, while negative stimuli elicit an enhanced response (Benning & Oumeziane, 2017). Therefore, when examining music and mental health problems, studies show that

music genres such as heavy metal, rap, and gothic rock can be associated with depression, self-harm, and suicidal tendencies (Ter Bogt et al., 2021b; Took & Weiss, 1994). For example, female adolescents who listen to heavy metal have shown higher levels of depressive symptoms, while pop and soul music can alleviate these symptoms. This could be due to the fact that heavy metal often contains pessimistic lyrics, while pop and soul music tend to offer more positive emotional content (Miranda & Claes, 2007). In addition, a study by Ter Bogt et al. (2021a) found that fans of all-out rock music had the highest peak levels of depressive symptoms and the lowest levels of well-being, while fans of metal music reported the most aggression. In general, music and its lyrics contribute positively to well-being by elevating mood, supporting identity formation, and providing comfort through specific and adaptive affect regulation strategies (Ter Bogt et al., 2021b; Yoon & Rottenberg, 2021). However, depressed individuals often show impaired emotion regulation and tend to rely on maladaptive strategies such as suppression, rumination, and avoidant coping when engaging with music (Kanagala et al., 2021). Furthermore, certain music preferences of individuals with depressive symptoms could reinforce these maladaptive mechanisms and therefore harm their psychological well-being (Miranda & Claes, 2008). Research by Kanagala et al. (2021) suggests that depressive symptoms are positively associated with unhealthy music use but not with healthy music use. This may be due to the fact that individuals with depressive symptoms tend to select mood-congruent music which perpetuates negative emotional states and reinforce distress rather than reduce it (Skånland, 2013). Moreover, mood congruency theory postulates that individuals are more likely to recall and engage with information that is congruent with their current affective state. In this context, negative moods might promote negative thinking and self-destructive behaviours that increase distress (Kanagala et al., 2021). Therefore, individuals with depression may listen to music that mirrors their negative emotional state due to difficulties in regulating their response to negative stimuli (Wilhelm et al., 2013). However, new evidence suggests that individuals with depressive symptoms do not select sad or low-energy music to reinforce negative mood states, but rather as an adaptive strategy aimed at relaxation, emotional calm, and emotional regulation. This suggests that music preferences in depression may serve multiple functions, including emotional self-regulation and mood modulation, rather than just the maintenance of sadness (Yoon & Rottenberg, 2021; Yoon et al., 2020).

While numerous studies have investigated the relationship between music preferences and personality traits, research establishing a link between music and mental health, including anxiety and depression, is much rarer, despite the significant burden these conditions pose (Miranda & Claes, 2007). Furthermore, the relationship between music preferences and psychiatric profiles is under-researched, and evidence is currently insufficient to establish a causal relationship between specific music genres and mental health outcomes (Ekinci et al., 2013; Kangala et al., 2021). However, the potential results of this correlational study could provide new insights into the relationship between music and mental health. Even if a causal relationship cannot be established, the results could help refine existing theories, generate hypotheses for future research, and highlight potential trends that can be investigated in experimental or longitudinal studies. Given the importance of the above findings and constructs, this study aims to determine whether there is a relationship between music preferences and depressive symptoms, more specifically, whether there is a relationship between the preference dimensions of Intense, Classical, Contemporary, Jazzy, and Unpretentious and depressive symptoms in two measurement settings (online and laboratory measurements). Based on previous research (Miranda & Claes, 2007; Ter Bogt et al. 2021a), participants with higher levels of depression should show a greater preference for Intense music. Furthermore, this study aims to examine gender differences in music preference**s**, with men expected to show a greater preference for Intense music compared to women (Bonneville-Roussy & Rust, 2018). In addition, this study explores potential differences between the two methods of measuring music preferences (online and laboratory settings) to determine whether the way of measurement influences participants' reported preferences.

Method

Participants

This study was conducted in two phases. The first phase included 101 participants from Croatia (24.8% male) aged between 15 and 69 years (M = 27.71; SD = 11.56). The participants were recruited as volunteers via an online invitation that encouraged them to pass the questionnaire on to others (snowball sampling system). At the end of the survey, the participants had the opportunity to take part in the second phase of the study, which included a laboratory-based measurement of musical preferences. The second phase included a subsample of 43 participants (41.9% male) from the original group who were between 15 and 69 years old (M = 32.26; SD =15.08).

Materials

The participants completed demographic questions about their age and gender and completed the Depression Anxiety Stress Scale—DASS-O (Lovibond & Lovibond, 1995). The scale consists of 42 items and has a three-factor structure—it consists of the depression, anxiety, and stress subscales, each represented by 14 items (Reić Ercegovac & Penezić, 2012). Only the Depression subscale adapted by Reić Ercegovac and Penezić (2012) was used for this study. The reliability coefficient (Cronbach's alpha) for the depression subscale is 0.95. An example of an item is *"I felt that I had nothing to look forward to"*. The participants were instructed to respond according to how they felt in general, and responses were recorded on a 4-point Likert scale (0 = *not at all true for me* to 3 = *completely true for me*). The final score is formed as a linear combination of the ratings for each subscale, with the total score ranging from 0 to 42.

A list of 17 music genres, adapted from Bonneville-Roussy et al. (2017), was used to assess the participants' declarative, explicit preferences for specific music genres based on their labels. These genres could be categorised into five major music preference dimensions: Unpretentious, Intense, Contemporary, Classical, and Jazzy. The participants rated their preference for each genre on a 5-point scale, with 1 indicating strong dislike and 5 indicating strong preference.

In addition to the genre ratings, the participants' actual preferences were measured by rating 34 music clips, with two clips representing each of the 17 music genres. These genres were categorised into five major music preference dimensions. The clips used in this study were taken from the list of clips provided by Bonneville-Roussy et al. (2017) and Rentfrow et al. (2012) depending on availability (as their models share similarities—for example, both categorise the rock music genre under the Intense dimension). The music examples selected by the authors were instrumental and were of unknown music pieces. Both groups of authors selected music examples by consulting experts: Rentfrow et al. (2012) consulted 10 musicologists to identify representative pieces for each music genre while Bonneville-Roussy et al. (2017) asked four genre experts to curate and refine music clips to reflect the full range of each genre. The clips were 15 seconds long and normalised in volume using Audacity software. The Participants rated each clip on a scale from 1 (*"I do not like the clip at all"*) to 5 (*"I like the clip very much"*) using response sheets with an ordinal numbering system and corresponding scales for the featured clips. The reliabilities of the music preference dimensions were generally higher in the laboratory setting compared to the online setting (Table 1). The Intense dimension demonstrated the highest reliability in the laboratory (α = .92) and remained acceptable online (α = .71). In contrast, the Unpretentious dimension showed the greatest deviation, going from questionable reliability (α = .67) in the laboratory to unacceptable reliability (α = .39) online. The Contemporary dimension also showed lower reliability online (α = .58), suggesting possible measurement inconsistencies. Additionally, the reliabilities of the music preference dimensions in this study (both online and in the labo-

ratory) follow similar trends to the study by Bonneville-Roussy and Rust (2018), with the Intense, Classical, and Jazzy dimensions having the highest reliability, and Contemporary and Unpretentious music preference dimensions having the lowest reliability.

Procedure

The study was conducted in two phases, combining online and laboratory settings to measure music preferences. In the first phase, the participants completed several online questionnaires, starting with a demographic survey, followed by a Depression subscale and finally ratings for 17 music genres. The aim of this phase was to determine the participants' declarative preferences for music genres based solely on their names. In the second phase, the participants who chose to continue were invited to the Laboratory for Experimental Psychology within a week to participate under controlled conditions. There they listened to 34 randomised audio clips, with two clips representing a specific music genre. Each clip was played once at 80% of the device's maximum volume, with enough time between clips for the participants to record their responses. The participants were not given any information about the clips other than the ordinal number, ensuring an unbiased assessment of their actual preferences for typical examples of each genre. To match the data in both phases, the participants used a unique code given to them before the study began using a template. This design allowed for a comprehensive comparison between declarative (online) and actual (laboratory based) music preferences.

Results

Preliminary Analysis

To determine whether there was a difference between the level of depressive symptoms of the participants that participated in the online measurement of music preferences (N=58) and those that participated in both online and the laboratory measurement (N=43), Mann-Whitney U test was calculated. The results of the test indicated that there was no statistically significant difference found in depressive symptoms between those that participated in online and those that participated in both online and laboratory measurement of music preferences (z= 0.84, p= 0.40). The effect size was small (r= 0.08).

Main results

The skewness and kurtosis values are relatively low in both online and laboratory settings; however, the Kolmogorov–Smirnov (*K-S*) test indicated that the depressive symptoms, as well as the scores for the online measured Classical and Intense music preference dimensions, deviated significantly from a normal distribution (Table 1). For this reason, non-parametric statistical methods were used for further data analysis (in addition to the smaller sample size for the laboratory measurement of music preferences).

Table 1

Descriptive statistics and reliability of variables used in online (N=101) and laboratory (N=43) measurement of music preferences

		M (*SD*)	*C* (*IQR*)	*Skewness*	*Kurtosis*	*K-S d*	*p*	*α*
Online setting	Depressive symptoms level	8.71 (9.60)	5.00 (10.00)	1.31	0.72	0.20	<.001	0.95
	Classical	2.53 (1.09)	2.50 (1.50)	0.35	-0.71	0.13	.047	0.70
	Intense	2.50 (1.04)	2.33 (1.33)	0.59	-0.39	0.14	.039	0.71
	Contemporary	2.96 (0.96)	3.00 (1.33)	0.10	-0.54	0.10	.209	0.58
	Jazzy	2.96 (0.91)	3.00 (1.20)	0.09	-0.32	0.06	.786	0.77
	Unpretentious	3.05 (0.74)	3.00 (1.00)	0.07	0.23	0.09	.425	0.39
Laboratory setting	Depressive symptoms level	9.26 (9.60)	6.00 (12.00)	1.19	0.41	0.20	.048	-
	Classical	2.74 (0.92)	2.75 (1.25)	0.25	-0.58	0.11	.660	0.77
	Intense	3.00 (1.09)	3.00 (1.66)	-0.01	-0.74	0.10	.692	0.92
	Contemporary	3.19 (0.85)	3.33 (0.83)	-0.76	0.60	0.12	.467	0.66
	Jazzy	3.36 (0.61)	3.50 (0.90)	-0.96	0.28	0.15	.258	0.71
	Unpretentious	3.32 (0.63)	3.25 (1.00)	-0.37	-0.08	0.08	.897	0.67

Note. *C* = median, *IQR* = interquartile range, *K-S* = value of Kolmogorov-Smirnov test, *α* = Cronbach's alpha.

To investigate the relationship between the degree of depressive symptoms and the five musical preference dimensions (Classical, Intense, Contemporary, Jazzy, and Unpretentious), the Goodman-Kruskal gamma correlation coefficient was calculated (Table 2). This coefficient is particularly recommended when the data contains numerous tied ranks, as was the case in this study, where the participants rated on a 5-point scale, which increases the likelihood of multiple participants giving the same rating for certain genres (e.g., 3—*neither like nor dislike*).

Table 2

Gamma correlation coefficients between the participants' general level of depressive symptoms and five musical preference dimensions measured online (N=101) and in laboratory (N=43)

		CLA	INT	CON	JAZ	UNP			CLA	INT	CON	JAZ	UNP
Online setting	DSL	.03	.23**	.08	.12	-.13	Laboratory setting	DSL	.09	.24*	.19	.09	-.13
	CLA		.16*	.03	.28**	.33*		CLA		-.08	-.05	.28*	.38**
	INT			.22**	.45**	.04		INT			.26*	.10	-.22*
	CON				.19**	.02		CON				.14	-.23*
	JAZ					.44**		JAZ					.41**

Note. *p<.05; **p<.01; *DSL* = depressive symptoms level, *CLA* = Classical, *INT* = Intense, *CON* = Contemporary, *JAZ* = Jazzy, *UNP* = Unpretentious.

The results show a significant, albeit small, positive correlation between the level of depressive symptoms and Intense musical preference dimension in both the online and laboratory measures of music preference (Table 2). In particular, the participants who scored higher on the depression scale also showed a greater preference for Intense music. However, no significant correlation was found between depressive symptoms and preferences for the other four music preference dimensions—Classical, Contemporary, Unpretentious, and Jazzy (Table 2).

The correlations between the music preference dimensions in the online and laboratory measures of music preferences are reasonably similar, although notable differences were found. Significant positive

correlations were found between the Classical, Jazzy, and Unpretentious music preference dimensions, as well as between the Intense-Contemporary and between Jazzy-Unpretentious music preference dimensions in both the online and laboratory measures of music preference (Table 2). These correlations were low to medium.

In the online measurement, however, additional small positive correlations were observed between the Classical and Intense preference dimensions, as well as between Contemporary and Jazzy and a medium correlation between Jazzy and Intense music (Table 2). These correlations were not present in the laboratory measurement of music preferences. Interestingly, in the laboratory measurement, the Intense and Contemporary music preference dimensions were negatively correlated with the Unpretentious music preference dimension (Table 2), suggesting that the participants who preferred Unpretentious music showed a lower preference for Intense and Contemporary music.

Additional analyses were also conducted to explore potential gender differences and discrepancies between the two methods of measuring music preferences (online and in the laboratory).

Additional Analyses

Moreover, Wilcoxon matched pairs test was conducted to compare the participants' declarative ratings of the music preference dimensions (online) with their ratings after listening to music clips (laboratory phase) (*N*=43) (Table 3). Additionally, gamma correlation coefficients have been calculated between declarative and listened ratings of musical preference dimensions for the subsample (*N*=43) (Table 4).

Table 3
Results of Wilcoxon matched pairs test between declarative and listened ratings of musical preference dimensions (N=43)

	Declarative preferences		Listening preferences				
	M (*SD*)	*C* (*IQR*)	*M* (*SD*)	*C* (*IQR*)	*z*	*p*	*r*
CLA	2.26 (0.86)	2.50 (1.50)	2.74 (0.92)	2.74 (1.25)	3.87	<.001	0.66
INT	2.55 (1.14)	2.33 (1.67)	3.00 (1.09)	3.00 (1.67)	3.35	<.001	0.54
CON	3.14 (0.99)	3.00 (1.67)	3.19 (0.85)	3.33 (0.83)	0.63	.528	0.10
JAZ	3.01 (0.89)	3.20 (1.40)	3.36 (0.61)	3.50 (0.90)	3.31	<.001	0.52
UNP	2.91 (0.72)	3.00 (1.00)	3.32 (0.63)	3.25 (1.00)	4.38	<.001	0.69

Note. *C* = median, *IQR* = interquartile range, *r* = Rosenthal's effect size, *CLA* = Classical, *INT* = Intense, *CON* = Contemporary, *JAZ* = Jazzy, *UNP* = Unpretentious.

A statistically significant difference was found between the declarative ratings and the ratings after listening to the clips for the Classical, Intense, Jazzy, and Unpretentious preference dimensions. Classical and Unpretentious music preference dimension showed the most pronounced difference with a large effect size, while the Intense and Jazzy music preference dimensions showed a significant difference with a moderate effect size (Table 4) according to Bartz (1999, p.184). In these cases, the participants expressed a higher level of liking for the music preference dimensions after listening to the clips compared to their initial declarative ratings (Table 3).

Table 4
Gamma correlation coefficients between declarative and listened ratings of musical preference dimensions (N=43)

Music dimension	*r* (listened vs. declarative rating)
Classical	.62**
Intense	.65**
Contemporary	.57**
Jazzy	.65**
Unpretentious	.59**

Note. **p<.01

The results show strong correlations between all five dimensions of music preference based on listening behaviour and the participants' declarative ratings, with the Intense and Jazzy dimensions being the most strongly correlated. This suggests that, despite significant differences between the declarative and listened ratings for four of the music preference dimensions (Table 4), the participants are consistent in their preferences as their preferences follow a similar trend regardless of the method of measurement (Table 4).

Mann-Whitney U tests were conducted to investigate gender differences in music preference ratings in the five dimensions of music preference, both online and in the laboratory (Table 5).

Table 5
Mann-Whitney U test results of gender differences in the participants' ratings for each of the five musical preference dimensions in online (N=101) and laboratory (N=43) setting

Online setting							
	Men (*N*=25)		Women (*N*=76)				
	M (*SD*)	*C* (*IQR*)	*M* (*SD*)	*C* (*IQR*)	*z*	*p*	*r*
CLA	2.18 (0.96)	2.00 (1.00)	2.64 (1.11)	2.50 (2.00)	1.84	.063	0.18
INT	2.91 (1.15)	3.00 (1.67)	2.36 (0.97)	2.33 (1.33)	-2.02	.042	-0.20
CON	3.24 (0.88)	3.00 (1.67)	2.86 (0.98)	2.67 (1.17)	-1.48	.136	-0.15
JAZ	2.87 (0.98)	3.00 (1.80)	2.98 (0.89)	2.90 (1.20)	0.26	.798	0.03
UNP	2.64 (0.70)	2.50 (1.00)	3.19 (0.71)	3.25 (0.75)	2.93	.003	0.29
Laboratory setting							
	Men (*N*=18)		Women (*N*=25)				
	M (*SD*)	*C* (*IQR*)	*M* (*SD*)	*C* (*IQR*)	*z*	*p*	*r*
CLA	2.40 (0.69)	2.50 (0.75)	2.98 (1.00)	3.00 (1.50)	2.08	.036	0.32
INT	3.43 (1.01)	3.33 (2.00)	2.69 (1.05)	2.83 (1.67)	-2.08	.037	-0.32
CON	3.27 (0.68)	3.33 (0.67)	3.13 (0.97)	3.33 (1.00)	-0.02	.980	-0.00
JAZ	3.28 (0.71)	3.40 (1.00)	3.41 (0.55)	3.60 (0.80)	0.49	.622	0.08
UNP	2.96 (0.60)	3.00 (0.50)	3.58 (0.52)	3.75 (0.88)	3.05	.002	0.47

Note. C = median, *IQR* = interquartile range, *r* = Rosenthal's effect size, *CLA* = Classical, *INT* = Intense, *CON* = Contemporary, *JAZ* = Jazzy, *UNP* = Unpretentious.

The online measurement revealed significant differences in the Intense and Unpretentious dimensions. Men reported significantly higher preferences for Intense music, whereas women rated Unpretentious music significantly higher. The effect sizes were small (r = -0.20 and 0.29) (Table 5).

In the laboratory measurement, significant gender differences were observed for Classical, Intense, and Unpretentious dimensions. Women indicated significantly higher preferences for Classical and Unpre-

tentious music, while men had a higher rating for Intense music. The effect sizes were small (r = ±0.32) to moderate (r = 0.47) (Table 5).

The results indicate gender differences in music preferences, with stronger effects observed in the laboratory setting compared to the online setting, particularly for Unpretentious music (Table 5).

Discussion

One of the aims of this study was to determine whether there is a relationship between depressive symptoms and preferences for the five dimensions of music preference: Intense, Classical, Contemporary, Jazzy, and Unpretentious. The same participants rated their music preferences in two phases of the study, and the results showed that the Intense music was positively associated with depressive symptoms in both the online and laboratory measurement of music preferences. However, no significant correlations were found between depressive symptoms and the other four music preference dimensions (Classical, Contemporary, Jazzy, and Unpretentious). As no differences in depressive symptoms were found between the participants in the online and laboratory measurements, it is unlikely that the differences in depression scores influenced the observed differences in music preferences in the two settings.

The relationship between depressive symptoms and a preference for Intense music obtained in this study supports the hypothesis that individuals with higher levels of depression are more likely to prefer Intense music, a finding that is consistent with the study by Ter Bogt et al. (2021a) and Miranda and Claes (2007). In their study, adolescent girls' preference for heavy metal music was found to be associated with increased depressive symptoms, while preference for soul and pop music was associated with lower levels of depression. According to their interactive influence perspective, individuals may initially gravitate towards music styles that resonate with their emotional state, with genres such as heavy metal, punk, and rock—characteristic of the Intense preference dimension—conveying negative, despairing messages that may increase depressive feelings. Conversely, genres such as soul and pop, with their uplifting melodies and comforting lyrics, can help to alleviate depressive symptoms. An alternative explanation for these findings could be found in early models of physiological underarousal in depression (Grossberg, 1972). According to the low positive emotion model, depressed individuals respond less strongly to pleasant stimuli, while the elevated negative emotion model states that they respond more intensely to negative stimuli (Benning & Oumeziane, 2017). It is possible that people with depression respond more strongly to the negative emotions expressed in Intense music genres such as heavy metal, rock, and punk, which emphasise aggression, despair, and loss, while responding less to the positive emotional cues offered by genres such as soul and gospel in the Unpretentious music preference dimension, which tend to be more soothing and romantic (Rentfrow & Gosling, 2003). In addition, individuals with depressive symptoms often have problems with emotion regulation and rely on maladaptive strategies such as rumination and avoidance when engaging with music (Kanagala et al., 2021). According to mood congruency theory, they are more likely to select music that reflects their current emotional state, which may exacerbate rather than alleviate stress (Skånland, 2013), contributing to the findings that depressive symptoms are associated with unhealthy music consumption, as individuals with depression listen to music that sustains negative emotions rather that adaptively regulating them (Kanagala et al., 2021). This may explain why the participants with higher depressive symptoms showed a stronger preference for Intense music genres such as rock, heavy metal, and punk, as these genres often convey themes of anger, sadness, and emotional intensity that are consistent with their affective state (Miranda & Claes, 2008).

Regarding the associations between the music preference dimensions, the results showed that Classical music was positively correlated with Intense, Unpretentious, and Jazzy music and that Intense music

was positively associated with Contemporary music. In addition, a positive correlation was found between Jazzy and Unpretentious music. Greasley and Lamont (2006) found through interviews that social interactions and environmental changes significantly influence music preferences, suggesting that preferences are not static but evolve over time. Music is made up of different attributes such as tempo, rhythm, and emotional tone, and individuals may prefer certain musical characteristics regardless of genre. For example, some people are attracted to sad or instrumental music regardless of whether it belongs to the rock, classical, or jazz genre (Rentfrow et al., 2012). Consequently, music preferences may not be limited to specific genres, but reflect a broader attraction to certain musical qualities, such as energy or emotional depth, which may be present in multiple genres. Interestingly, results from laboratory measures of music preferences have shown that the Intense and Contemporary music preference dimensions were inversely correlated with the Unpretentious music preference dimension. This finding is consistent with previous research by Bonneville-Roussy and Eerola (2018), who found that genres within the Unpretentious preference dimension (such as country, pop, and gospel) contrast strongly with genres within the Intense (e.g., punk, heavy metal) and Contemporary (e.g., electronic, hip-hop) preference dimensions. The contrasting characteristics of these genres—ranging from the melodic, simple patterns of Unpretentious music to the loud, distorted qualities of Intense music—suggest that these preference dimensions appeal to different emotional experiences (Bonneville-Roussy et al., 2017). The lack of similar results in the online phase of the study may be due to the use of genre-based measures, which may not have fully captured the participants' nuanced preferences for certain musical attributes.

Furthermore, this study aimed to investigate potential differences between the two methods of measuring music preferences (online and laboratory settings) to determine whether the type of measurement influences the preferences stated by the participants. Significant differences were found between the participants' declarative preferences (online measurements) and their ratings after listening to music clips (laboratory measurements) for the Classical, Intense, Jazzy, and Unpretentious preference dimensions, suggesting that listeners may be influenced by the specific characteristics of the music itself rather than by preconceived assumptions about music genres. This observation is consistent with research findings by Bonneville-Roussy and Rust (2018), which suggest that social influences are more likely to shape genre preferences, while preferences for actual music clips may be less influenced by external factors such as stereotypes or social desirability. This suggests that music clips, as opposed to declarative broad music genres, provide a more direct measure of individual preferences without the influence of social or cultural associations.

In addition, this study also aimed to examine gender differences in music preferences. Gender differences were found, with men preferring Intense music more than women, and women showing a stronger preference for Unpretentious music. This is consistent with previous research, suggesting that women tend to prefer softer, more vocal and melodic music (e.g., soul, R&B, pop), while men show a greater preference for harder, more aggressive music genres such as heavy metal and rock. These types of gender differences in music preferences have been documented since adolescence and persist into adulthood (Bonneville-Roussy & Rust, 2018).

Limitations and Suggestions

A potential limitation of the current study is the use of a music preference model that may be more appropriate for English-speaking populations (Bonneville-Roussy et al., 2017; Rentfrow et al., 2012) and may not fully capture music genres that are more prevalent in the Croatian context, such as turbofolk (Žauhar & Levak, 2020). In addition, the relatively small sample size limits the significance of the results. Another limitation is that the participants were not explicitly asked about their familiarity with the music

clips, meaning that familiarity was assumed rather than measured, which may have influenced their stated preferences.

Future studies could address these limitations by including a broader range of music genres relevant to the Croatian context. In addition, increasing the sample size would improve the reliability and generalisability of the results. To better control for the influence of familiarity, participants in future studies should also indicate whether they recognise the music clips presented in the study.

However, the two-phase design used in this study may have helped to alleviate some of these limitations by providing a more accurate representation of the relationship between depressive symptoms and music preferences. The combination of online surveys with controlled listening tasks in the laboratory may have improved the reliability of the data and reduced potential biases. These strengths allowed the study to explore the relationship between depressive symptoms and music preferences, particularly in relation to the Intense music preference dimension, which includes genres such as heavy metal, rock, and punk. Although a low positive correlation was found, the results are preliminary and should be interpreted with caution due to the limitations of the study. However, the results suggest that music may play a role in reflecting or enhancing emotional states in individuals with depressive symptoms. Future research is needed to further investigate these relationships and explore the potential of music-based interventions for individuals with depressive symptoms.

Conclusion

The study observed a small positive correlation between the level of depressive symptoms and preference for Intense music in both online and laboratory setting—the participants who scored higher on the depression scale also showed a greater preference for Intense music. The participants' music preferences were generally consistent across declarative and listening-based measures, though they showed a stronger preference for Classical, Intense, Jazzy, and Unpretentious music after listening (laboratory setting). Gender differences were also observed in both phases of the study—in the online setting, men preferred Intense music more than women, who favoured Unpretentious music, whereas in the laboratory setting, men also favoured Intense music more than women, who showed a greater preference for the Unpretentious, as well as for the Classical music.

References

Bartz, A. E. (1999). *Basic statistical concepts* (4th ed). Upper Saddle River, NJ: Merrill.

Benning, S. D., & Ait Oumeziane, B. (2017). Reduced positive emotion and underarousal are uniquely associated with subclinical depression symptoms: Evidence from psychophysiology, self-report, and symptom clusters. *Psychophysiology, 54*(7), 1010-1030. https://doi.org/10.1111/psyp.12853

Bonneville-Roussy, A., & Eerola, T. (2018). Age trends in musical preferences in adulthood: 3. Perceived musical attributes as intrinsic determinants of preferences. *Musicae Scientiae, 22*(3), 394–414. https://doi.org/10.1177/1029864917718606

Bonneville-Roussy, A., & Rust, J. (2018). Age trends in musical preferences in adulthood: 2. Sources of social influences as determinants of preferences. *Musicae Scientiae, 22*(2), 175-195. https://doi.org/10.1177/1029864917704016

Bonneville-Roussy, A., Stillwell, D., Kosinski, M., & Rust, J. (2017). Age trends in musical preferences in adulthood: 1. Conceptualization and empirical investigation. *Musicae Scientiae, 21*(4), 369-389. https://doi.org/10.1177/1029864917691571

Chee, Z. J., Chang, C. Y. M., Cheong, J. Y., Malek, F. H. B. A., Hussain, S., de Vries, M., & Bellato, A. (2024). The effects of music and auditory stimulation on autonomic arousal, cognition and attention: a systematic review. *International Journal of Psychophysiology*, 112328. https://doi.org/10.1016/j.ijpsycho.2024.112328

Ekinci, O., Bez, Y., Sabuncuoglu, O., Berkem, M., Akin, E., & Imren, S. G. (2013). The association of music preferences and depressive symptoms in high school students: A community-based study from Istanbul. *Psychology of Music*, *41*(5), 565-578. https://doi.org/10.1177/0305735612440614

Fricke, K. R., Greenberg, D. M., Rentfrow, P. J., & Herzberg, P. Y. (2021). Measuring musical preferences from listening behavior: Data from one million people and 200,000 songs. *Psychology of Music, 49*(3), 371–381. https://doi.org/10.1177/0305735619868280

Greasley, A. E., & Lamont, A. M. (2006). Music preference in adulthood: Why do we like the music we do? In *Proceedings of the 9th International Conference on Music Perception and Cognition*. Bologna, Italy.

Greenberg, D. M., Decety, J., & Gordon, I. (2021). The social neuroscience of music: Understanding the social brain through human song. *American Psychologist*, *76*(7), 1172. https://doi.org/10.1037/amp0000819

Grossberg, S. (1972). A neural theory of punishment and avoidance, II: Quantitative theory. *Mathematical Biosciences, 15(*3), 253–285. https://doi.org/10.1016/0025-5564(72)90038-7

Hallam, S., Cross, I., & Thaut, M. (Eds.). (2009). *Oxford handbook of music psychology*. Oxford University Press.

Kanagala, S. C., Schäfer, T., Greenberg, D. M., & Gabińska, A. (2021). Depression Symptoms are Linked to Music Use. *Music & Science*, *4*. https://doi.org/10.1177/20592043211057217

Loureiro, C., van der Meulen, K., & del Barrio, C. (2024). Why I listen to music: Emotion regulation and identity construction through music in mid-adolescence. *Empiria: Revista de metodología de ciencias sociales*, *60*, 145-168. https://doi.org/10.5944/empiria.60.2024.39285

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales.* Psychology Foundation. https://doi.org/10.1037/t01004-000

Miranda, D., & Claes, M. (2007). Musical preferences and depression in adolescence. *International Journal of Adolescence and Youth*, *13*(4), 285-309. https://doi.org/10.1080/02673843.2007.9747981

Miranda, D., & Claes, M. (2008). Personality traits, music preferences and depression in adolescence. *International Journal of Adolescence and Youth, 14*(3), 277-298. https://doi.org/10.1080/02673843.2008.9748008

Reić Ercegovac, I., & Penezić, Z. (2012). Skala depresivnosti, anksioznosti i stresa [The Depression Anxiety Stress Scale]. In: Proroković, A., Ćubela Adorić, V., Penezić, Z., & Tucak Junaković, I. (Eds.), *Zbirka psihologijskih skala i upitnika, Svezak 6*. University of Zadar. https://morepress.unizd.hr/books/index.php/press/catalog/view/32/28/374

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, *84*(6), 1236-1256. https://doi.org/10.1037/0022-3514.84.6.1236

Rentfrow, P. J., Goldberg, L. R., Stillwell, D. J., Kosinski, M., Gosling, S. D., & Levitin, D. J. (2012). The song remains the same: A replication and extension of the MUSIC model. *Music Perception*, *30*(2), 161–185. https://doi.org/10.1525/mp.2012.30.2.161

Skånland, M. S. (2013). Everyday music listening and affect regulation: The role of MP3 players. *International Journal of Qualitative Studies on Health and Well-Being, 8*(1), 20595. https://doi.org/10.3402/qhw.v8i0.20595

Tarrant, M., North, A. C., & Hargreaves, D. J. (2000). English and American adolescents' reasons for listening to music*. Psychology of Music, 28*(2), 166–173. https://doi.org/10.1177/0305735600282005

Ter Bogt, T., Hale III, W. W., & Becht, A. (2021a). "Wild years": rock music, problem behaviors and mental well-being in adolescence and young adulthood. *Journal of Youth and Adolescence, 50*(12), 2487-2500. https://doi.org/10.1007/s10964-021-01505-0

Ter Bogt, T., Hale, W. W., Canale, N., Pastore, M., & Vieno, A. (2021b). Goth music and depressive symptoms among adolescents: a longitudinal study. *Journal of Youth and Adolescence, 50*, 1925-1936. https://doi.org/10.1007/s10964-020-01294-y

Took, K. J., & Weiss, D. S. (1994). The relationship between heavy metal and rap music and adolescent turmoil: Real or artifact. *Adolescence, 29*(115), 613–621. https://www.proquest.com/scholarly-journals/relationship-between-heavy-metal-rap-music/docview/195926392/se-2

van Sprang, I., & Haeyen, S. (2024). Music therapy for therapeutic development in personality disorders: A qualitative case study. *Journal of Clinical Psychology*, *80*(7), 1596-1606. https://doi.org/10.1002/jclp.23682

Vella, E. J., & Mills, G. (2017). Personality, uses of music, and music preference: The influence of openness to experience and extraversion. *Psychology of Music*, *45*(3), 338-354. https://doi.org/10.1177/0305735616658957

Wilhelm, K., Gillis, I., Schubert, E., & Whittle, E. L. (2013). On a blue note: Depressed Peoples' reasons for listening to music. *Music and Medicine, 5*(2), 76–83. https://doi.org/10.1177/1943862113482143

Yoon, S., & Rottenberg, J. (2021). Listening to the blues: An ecological momentary assessment of music choice in depression. *Emotion*, *21*(6), 1177-1187. https://doi.org/10.1037/emo0000911

Yoon, S., Verona, E., Schlauch, R., Schneider, S., & Rottenberg, J. (2020). Why do depressed people prefer sad music? *Emotion, 20*(4), 613–624. https://doi.org/10.1037/emo0000573

Žauhar, V., & Levak, N. (2020). Glazbene preferencije prema modelu MUSIC: Povezanost s motivima slušanja glazbe i crtama ličnosti [Music preferences according to the MUSIC model: Relationships with uses of music and personality traits]. *Psihologijske teme, 29*(2), 311-337. https://doi.org/10.31820/pt.29.2.6

12

Conditions in Maternity Wards and Childbirth Experiences in Serbia, Bosnia and Herzegovina, Montenegro, and Croatia

Jasmina Šunjka Ćorković
Faculty of Media and Communications, Department of Psychology, Singidunum University, Belgrade, Serbia
j.n.sunjka@gmail.com

Marko Đurica
Faculty of Media and Communications, Department of Psychology, Singidunum University, Belgrade, Serbia

Tiana Alexandra Marić
Faculty of Media and Communications, Department of Psychology, Singidunum University, Belgrade, Serbia

Hristina Kamenović
Faculty of Media and Communications, Department of Psychology, Singidunum University, Belgrade, Serbia

Abstract

Various factors, including individual, interpersonal, and unexpected medical issues, play significant roles in shaping women's childbirth experiences. In response to the scarcity of research in this domain across the Balkan countries, this study offers foundational insights into the conditions prevalent in maternity wards and the intricacies of childbirth experiences. To achieve this, we constructed an online questionnaire which was completed by a total of 10,657 respondents from Serbia, Montenegro, Bosnia and Herzegovina, and Croatia, who gave birth between 2000 and 2021. In order to minimise the influence of potential confounding variables resulting from changes in obstetric care practices over the past two decades, data analysis was limited to the participants who gave birth between 2016 and 2021 (N= 7,779). Data were analysed using descriptive statistics, as well as content analysis for an open-ended question regarding adverse experiences during childbirth and hospitalisation. The most significant findings of this study highlight widespread issues of bribery and corruption within maternity wards, with nearly 50% of the respondents across all four countries reporting the use of personal connections with hospital staff to secure adequate care. Psychological support post-childbirth was alarmingly insufficient, with fewer than

5% of women receiving any form of assistance, despite 28.5% to 35.0% of the participants reporting symptoms of postpartum depression. Particularly noteworthy outcomes of the content analysis were instances within the staff category, where the respondents reported various forms of mistreatment and substandard care. These results underscore the necessity for further dedicated research on this pivotal topic.

Keywords: *pregnancy, maternity wards, childbirth, corruption in healthcare, postpartum depression*

Introduction

Childbirth and the perinatal period encompass significant physical, psychological, and social transformations in a woman's life, accompanied by a marked increase in the utilisation of healthcare services (Rodríguez & Rivières-Pigeon, 2007). Given these profound changes, a certain standard of care is essential to ensure that both the newborn and the mother can thrive and adapt during this critical transitional phase. The concept of "quality of care" has been defined in various ways throughout the literature, with fewer definitions specifically addressing maternity care. However, some studies have clarified this concept by identifying key elements: ensuring a minimum standard of healthcare for all pregnant women and their newborns, providing enhanced care for those who require it, and striving for optimal outcomes (Pittrof et al., 2002). Historically, maternal health was largely overlooked in policy development, with primary attention directed towards infant health. Maternal well-being was often reduced to technical and medical considerations until the feminist movement brought sexual and reproductive rights to the forefront, advocating for their recognition as critical global issues (Sen et al., 2018).

Since that time, numerous studies worldwide have explored various dimensions of childbirth and the quality of maternity care. Research has consistently demonstrated that both the quality of maternity care and interpersonal support play a pivotal role in shaping women's childbirth experiences. Inadequate care has been linked to adverse outcomes, such as postpartum depression and anxiety, underscoring the critical importance of compassionate and supportive care practices during this period (Rodríguez-Almagro et al., 2019; Leeds & Hargreaves, 2008). Furthermore, obstetric violence has been shown to have lasting effects on patients' willingness to seek medical care, extending beyond gynaecological services (Perrotte et al., 2020).

Key factors influencing women's satisfaction with childbirth include effective communication, physical comfort, and emotional support provided by healthcare professionals (Goodman et al., 2004; Gibbins & Thomson, 2001). For example, a woman's perceived sense of control during labour is a crucial determinant of a positive birth experience (Bryanton et al., 2008). Similarly, the provision of breastfeeding support in maternity wards is instrumental in enhancing maternal satisfaction, which, in turn, can facilitate early bonding and improve infant health outcomes (Pérez-Escamilla et al., 1994). Recent research on this topic recommends implementing a Baby-Friendly Hospital Initiative (BFHI) in maternity wards, as it promotes rooming-in, allowing mothers and infants to remain together to facilitate breastfeeding. This approach helps mothers recognise and respond to their infants' needs. Additionally, the implementation of a mandatory 20-hour World Health Organization breastfeeding counselling training for maternity ward staff has been proposed to enhance support (Živković Šulović, 2024). Despite these positive developments, research continues to highlight persistent issues of disrespect and abuse in maternity care, particularly in low-resource settings, which can significantly undermine the quality of the childbirth experience (Freedman et al., 2014; Sen et al., 2018).

These findings underscore a global concern; yet, in the Balkan countries—namely Serbia, Bosnia and Herzegovina, Montenegro, and Croatia—research and policy efforts addressing maternity ward conditions and the quality of childbirth experiences remain limited. Despite maternal health being a priority according to World Health Organization (WHO) standards (Chalmers et al., 2001), existing studies suggest that maternity wards in the Balkans frequently fail to meet essential care standards, with cultural factors potentially exacerbating these deficiencies (Chalmers, 1997). In Serbia specifically, Stanković (2017) identifies a range of challenges faced by women in public healthcare institutions, highlighting common themes of isolation, abandonment, poor communication, and a lack of agency. For instance, healthcare practitioners often controlled the birthing process with minimal communication or consideration for the woman's subjective experiences or preferences, fostering a sense of helplessness (Stanković, 2017). Similarly, a study conducted on a sample of over 1,000 women in Croatia revealed that nearly one-third of participants felt excluded from the decision-making process before, during, or after labour. Additionally, more than one-third of the respondents

reported experiencing various forms of both verbal and physical abuse. The research further underscores concerns regarding maternity ward conditions, particularly the failure to meet women's basic needs, such as access to food and water, as well as inadequate hygiene standards in maternity rooms (Cvitkušić, 2021). In like manner, research conducted in collaboration with the Ministry of Health and the UNICEF Country Office in Montenegro, aimed at evaluating the quality of maternal and neonatal services in the country, identified key areas for improvement, including the provision of private delivery rooms and companionship, the reduction of caesarean delivery rates, enhanced preparedness for obstetric and neonatal emergencies, improved counselling services, and increased maternal involvement in the decision-making process (Hodorogea et al., 2023).

The studies mentioned above, while significant, represent only a small subset of research addressing this issue within the Balkan countries. Therefore, the primary objective of this research was to broaden the scope by examining conditions in maternity wards and childbirth experiences across Serbia, Montenegro, Bosnia and Herzegovina, and Croatia. This study aims to provide a more comprehensive picture of the maternity care environment in these countries, focusing on factors such as healthcare staff interactions, facility conditions, emotional and psychological support, and potential issues like bribery and preferential treatment. By developing a comprehensive questionnaire that includes both quantitative measures and open-ended responses, we seek to capture the nuanced experiences of women in this region and explore how systemic issues—such as inadequate communication, emotional support, and a lack of patient agency—impact overall satisfaction with childbirth experiences. Ultimately, this study seeks to expand the discourse on maternity care in the Balkans by identifying areas in need of policy reform. By contributing empirical data, the findings can serve as a resource for policymakers and healthcare providers, helping to implement changes that ensure more respectful, supportive, and empowering childbirth experiences across the region. The main aim of the study was to present and make sense of the data collected in order to create a starting point and set a wider framework for future research which will dive deeper into more specific topics and analysis.

Method

Data collection

This study uses a descriptive research design to examine the situation in relation to maternity ward conditions and the delivery experience in the period between 2016 and 2021. This approach facilitated the collection of a diverse array of data, which was particularly significant given the absence of prior large-scale studies in Serbia, Croatia, Montenegro, and Bosnia and Herzegovina. The main aim of the study was to present and make sense of the data collected in order to create a starting point for future research.

The data used in this study were collected using an online questionnaire created by the authors of the study. The data collection lasted from August 27 to September 14, 2022. Within a two-week period, a sample of over 10,000 participants was obtained. The feasibility of this recruitment approach can be attributed to the increased accessibility of online survey distribution through social media platforms. Notably, influencers with substantial followings, some exceeding 100,000 and active in advocacy, significantly contributed to participant recruitment by sharing a link to the Google Form. This strategy generated over 2,000 responses within a few hours. Furthermore, influencers identified as "mother influencers" effectively targeted the desired demographic, amplifying both the reach and relevance of the survey. The sample was non-random; however, its substantial size and the inclusion of participants from multiple countries enabled a high degree of heterogeneity across socio-demographic variables. When participants clicked on the link shared on social media platforms, they were redirected to the Google Forms survey. First, they received a brief description of the study and its purpose, the voluntary nature of participation and the anonymity of their responses. If the participants

proceeded to answer the questions after reading the introduction, it was assumed that they had given their informed consent. Firstly, the participants completed a series of 8 questions on socio-demographic variables such as country of origin, their age at childbirth, year of birth, and the type of labour they had. They then answered a total of 41 questions about hospital conditions, treatment by hospital staff and their personal perceptions. The questionnaire contained both Likert and binary (yes/no) items as well as questions with open answers. In the first question, the participants were asked about their most positive memories of the whole childbirth experience and the hospital stay, and in the second question about their most negative memories.

Sample

Initially, the data gathered included all the participants who gave birth between 2000 and 2021. However, to minimise the influence of potential confounding variables resulting from changes in obstetric care practices over the past two decades, data analysis was limited to the participants who gave birth between 2016 and 2021. This approach ensured that only the most relevant and contemporary data were considered. Notably, this subsample represents 75.1% of the original sample.

The sample consisted of 7779 women from Serbia, Croatia, Montenegro and Bosnia and Herzegovina. Most of the participants were from Serbia (N = 5,302), followed by Bosnia and Herzegovina (N = 1,388). The majority were between 26 and 30 years old (42.5%) and between 20 and 25 (29.5%). For 81.8% of the participants, it was their first childbirth. Additionally, 57.3% gave birth vaginally, 21.3% had a cesarean section, and 21.3% received an epidural during labour. A detailed sample description is provided in Table 1.

The inclusion criteria required the participants to have given birth on the territory of Serbia, Montenegro, Bosnia and Herzegovina, and Croatia between 2016 and 2021. The individuals who did not meet either of these criteria were excluded to ensure that the data accurately reflected the experiences of the target population. The sample size was not determined on a power analysis; instead, the goal was to recruit as many participants as possible to maximise the representativeness of the findings.

Table 1
Sociodemographic profile of the sample by countries

		Serbia	Montenegro	Bosnia and Herzegovina	Croatia
Participants		68.2%	9.4%	17.8%	4.6%
Age	<19	1.3%	1.2%	1.5%	0%
	20-25	27%	28.5%	37.3%	38.9%
	26-30	42.5%	46.6%	40.3%	41.1%
	31-35	22.7%	18.8%	17.4%	16.3%
	36-40	6%	4.8%	3.2%	3.3%
	41+	0.5%	0.1%	0.3%	0.3%
First childbirth	Yes	81.8%	79.7%	82.9%	82%
	No	18.2%	20.3%	17.1%	18%
Due date	Expected	63.2%	64.2%	55.1%	48.5%
	Early	11.8%	11.6%	16.6%	14%
	Late	25%	24.2%	28.3%	37.5%
Type of childbirth	Natural	54.5%	64.2%	63.6%	61.1%
	Epidural	24.6%	14%	11.2%	27%
	Cesarean delivery	20.9%	21.8%	25.2%	11.9%

Data analysis

IBM SPSS 24.0.0.1 was used to analyse the data collected from the online survey. Since the main aim of this research was to present findings from the questionnaire, descriptive statistics were employed.

Results

Descriptive statistics

The descriptive analysis provides a detailed overview of the systemic and psychological challenges faced by women during and after childbirth across Serbia, Montenegro, Bosnia and Herzegovina, and Croatia. The findings highlight prevalent issues such as bribery, corruption, a lack of psychological support, and postpartum depression. The summarised data are presented in Table 2.

Table 2
Summary of key findings on maternity ward conditions and postpartum experiences by country

Category	Serbia N (%)	Croatia N (%)	Bosnia and Herzegovina N (%)	Montenegro N (%)
Used personal inside connections to ensure proper care	3044 (57.4)	145 (40.8)	628 (45.2)	307 (41.8)
Provided informal financial compensation to medical staff	2625 (49.5)	44 (12.4)	947 (68.2)	528 (71.9)
Received/were offered psychological support after childbirth	172 (3.2)	14 (3.9)	20 (1.4)	32 (4.4)
Experienced symptoms of postpartum depression	1527 (28.8)	101 (28.5)	457 (32.9)	257 (35.0)
Received support from family and/or partner for postpartum depression	999 (18.8)	62 (17.5)	300 (21.6)	172 (23.4)
Sought professional help for postpartum depression	256 (4.8)	13 (3.7)	58 (4.2)	35 (4.8)

Note. Data represent affirmative responses only. Values are shown as frequency and percentage N (%).

The findings presented in Table 2 reveal notable disparities in maternity care experiences across the analysed countries. In response to item 47, *"Did you use personal or professional connections to ensure you received appropriate care during childbirth?"*, the highest proportion of affirmative responses was recorded in Serbia (57.4%), while lower frequencies were reported in Bosnia and Herzegovina (45.2%), Montenegro (41.8%), and Croatia (40.8%). Similarly, item 42, *"Did you provide financial compensation to a doctor, anaesthesiologist, or midwife for their services?"*, showed substantial variation, with the highest prevalence in Montenegro (71.9%) and Bosnia and Herzegovina (68.2%), compared to significantly lower rates in Serbia (49.5%) and especially Croatia (12.4%). Access to psychological support was reported as limited across all countries. In response to item 43, *"Were you offered a conversation with a psychologist in the hospital to address any postpar-*

tum issues, such as bonding with your baby, body image concerns, the childbirth experience, breastfeeding challenges, or family relationships?", only 3.2% of the respondents in Serbia, 3.9% in Croatia, 1.4% in Bosnia and Herzegovina, and 4.4% in Montenegro indicated they were offered such support. Despite the low availability of psychological support, a considerable number of women reported experiencing postpartum mental health difficulties. In response to item 44, *"Did you experience symptoms of postpartum depression?"*, between 28.5% (Croatia) and 35.0% (Montenegro) of the respondents indicated they had experienced such symptoms. Among those who reported postpartum depression, item 45, *"If you experienced postpartum depression, did you feel supported by your partner or close family?"*, revealed that the proportion of women receiving family or partner support ranged from 17.5% (Croatia) to 23.4% (Montenegro). Finally, item 46, *"If you experienced postpartum depression, did you seek help from a psychologist or psychotherapist?"*, showed that seeking professional help was generally low across all countries, ranging from 3.7% in Croatia to 4.8% in Serbia and Montenegro.

Content Analysis

In addition to the descriptive statistics, a content analysis was conducted on one of the open-ended questions in the questionnaire (item 49: *"What would you highlight as the most negative memory of childbirth and your hospital stay? "*). This analysis employed a conceptual approach, aiming to quantify and examine the presence of a specific concept within the dataset. The primary objective was to explore the frequency of 15 predefined themes (as outlined in Table 3), which were identified through an initial review of the first 500 responses. It is important to note that the content analysis presented in this study covers only 10% of the total sample of the respondents who gave birth between 2016 and 2021, specifically 778 randomly selected responses to the open-ended question regarding the participants' most negative experiences during childbirth and hospital stay (item 49). The Inter-Rater Reliability (IRR) was also calculated, yielding a high reliability coefficient of 88%, with three independent raters coding the responses.

Table 3
Results of the content analysis

Description	N (%)
Staff-related complaints: *all complaints related to staff - negligence/behaviour/professionalism/ shortcomings (midwives, doctors, any other staff)*	351 (45.4)
Hygiene-related complaints*: all complaints related to hygiene*	25 (3.2)
Facility conditions complaints*: complaints about the conditions of stay of all kinds; institution conditions as a whole/room/bathrooms (overcrowding, temperature, appearance...)*	45 (5.6)
Food and meal complaints*: complaints about food and meals*	21 (2.5)
Multi-complaints: *multiple complaints about hygiene, conditions, food, staff, fellow patients (statements with at least 2 factors combined)*	97 (12.5)
Overall negative feedback: *complaints explicitly stating that everything is bad*	33 (4.3)
Complaints about fellow patients	11 (1.4)
No complaints*: everything is/was okay*	78 (10.1)
Emotion-related complaints*: complaints in the form of a spectrum of various negative emotions (fear, loneliness, uncertainty, etc.), whereas a lack of information and similar do not fall into this category as it is related to staff*	5 (0.7)
Visitor restrictions complaints: *complaints about visitor restrictions*	5 (0.7)

Pain and childbirth experience complaints: *complaints about the childbirth experience/pain (where specific complications and staff are not mentioned)*	23 (3.0)
Non-staff-related complications: *complaints about various complications before/during/after childbirth that are not related to staff (otherwise categorised as "S")*	35 (4.3)
Corruption-related complaints: *complaints explicitly mentioning corruption, bribery, "connections"*	17 (2.2)
Private hospital*: childbirth in a private hospital or suite*	1 (0.1)
Breastfeeding complaints*: complaints related to breastfeeding (only if not caused by staff incompetence, otherwise categorised as "S")*	2 (0.3)
(Other): *other complaints that cannot be categorised into any of the listed categories; empty fields; punctuation marks and other unintelligible symbols.*	25 (3.2)
Disagreement*: cases where all three raters assigned different categories to an item*	4 (0.5)

Note. The table presents the frequency and percentage of coded categories assigned by three independent raters. The category DIS (*Disagreement*) represents cases where all three raters assigned different categories to an item, indicating a complete lack of agreement. In cases where two raters assigned the same category while the third assigned a different one, the majority decision was adopted as the final categorisation.

As presented in Table 3, nearly half of the randomly selected respondents identified interactions with hospital staff as the most negative aspect of their hospital stay (45.4%). Representative responses include: *"I endured doctors and assistants pressing on my stomach, but the most terrifying experience during childbirth was the nervous and rude midwife."* (respondent from Serbia), *"I was accused of being spoiled, my stomach was pressed on, and there was a lack of interest in assisting with breastfeeding or helping me when my baby cried through the night."* (respondent from Bosnia and Herzegovina), and *"One of the nurses lost her temper when I asked for pain medication while she was watching a TV series. I didn't even dare to use a call button, I had to reach out to her myself, even though I was cut, torn, and stitched."* (respondent from Croatia). Following this, the next most frequent category of complaints (12.5%) involved multiple issues, such as hygiene, hospital conditions, food, staff behaviour, and the conduct of fellow patients. Examples from this category include: *"Well there are more things—no showerhead in the shower cabin, and the curtain is so blood-stained that it cannot be cleaned. I did not receive any medication for the headache following a challenging birth, despite waiting for three hours."* (respondent from Serbia) and *"Bad bathroom hygiene and the staff did not focus a lot on us women who had a C-section. There was no hot water, soap, or toilet paper; you had to bring those things from home. Horrible and ripped hospital gowns."* (respondent from Montenegro). The least frequent complaints were related to visitor restrictions, breastfeeding issues not involving staff treatment, and experiences related to childbirth in private hospitals or suites.

Discussion

The presented study on maternity ward conditions and childbirth experiences, which includes data from a substantial sample across Serbia, Croatia, Montenegro, and Bosnia and Herzegovina, makes an important contribution to addressing a significant gap in the existing literature. The findings underscore critical deficiencies, highlighting staff negligence, unprofessional behaviour, and inadequacies as the most significant issues. Facility conditions and hygiene also emerged as notable concerns, alongside pain and childbirth experiences. These insights provide a valuable foundation for future research and offer a framework to guide efforts aimed at improving maternity care in the region.

The findings of our study illustrate considerable gaps in maternity care practices when evaluated

against the World Health Organization (WHO) standards. WHO emphasises the necessity of high-quality maternity care that respects women's rights and ensures the health of both mothers and newborns. While our results indicate some positive aspects, such as general satisfaction among many participants, significant shortcomings were also noted in critical areas, including pain management, procedures that have been used, and postnatal care. These discrepancies are concerning, particularly in light of WHO recommendations that prioritise compassionate care and patient dignity during childbirth (Chalmers et al., 2001; Freedman et al., 2014).

Laws vs. Reality

National legislation, including the Law on Health Protection ("Official Gazette of RS", Nos. 25/2019 and 92/2023) and the Law on Patient Rights ("Official Gazette of RS", Nos. 45/2013 and 25/2019), establishes robust protections for the rights of pregnant women, aiming to ensure dignity and a minimum standard of care during childbirth. However, our findings reveal a concerning disparity between these legal provisions and the lived experiences reported by the participants, highlighting maternity care deficiency more as a systemic problem and less an individual one. This misalignment between legislative intent and practice highlights the urgent need for policy reforms, particularly in addressing bribery and neglect. Key rights enshrined in Serbian law include the right to privacy, ensuring the confidentiality of personal and medical information, and informed consent, which empowers women to make educated decisions regarding their care. These provisions are integral to fostering autonomy and respect throughout the childbirth process. The law also upholds the right to choose a birthing environment, the presence of a support person during labour, and access to adequate pain relief and anaesthesia, ensuring physical comfort and emotional support. Furthermore, postpartum care, including mental health services and breastfeeding guidance, is mandated to promote comprehensive well-being for mothers and newborns. Crucially, women have the legal right to file complaints regarding violations of these rights, promoting accountability within the healthcare system. Despite these protections, our results indicate that many women experience a significant gap between these standards and the realities of care, emphasising the need for stronger enforcement mechanisms and systemic improvements to align practice with legislative and ethical mandates.

Our findings reveal that while significant progress has been made in establishing legal protections for pregnant women, systemic issues persist that undermine the childbirth experience. Historically, problems such as neglect, a lack of privacy, and insufficient pain relief were pervasive concerns. Today, despite robust national legislation aimed at safeguarding dignity and rights during childbirth, these issues remain prevalent, along with newer challenges like bribery and inconsistent enforcement of care standards. Cultural factors and systemic deficiencies continue to shape negative experiences for a significant number of women, indicating a persistent gap between legal provisions and practice. This highlights the need for stronger implementation of existing laws, increased accountability, and culturally sensitive interventions to bridge the divide and ensure dignity, respect, and quality care for all.

From Past to Present

Building upon the identified discrepancies between legislative frameworks and practical challenges, it is crucial to examine the historical trends that may have contributed to these systemic issues. Our findings are consistent with prior research that highlights systemic neglect, as well as instances of disrespect and abuse in maternity care (Sekulic, 2016; Sen et al., 2018). Although the majority of women reported pos-

itive experiences, a significant proportion encountered negative circumstances, underscoring the need for targeted interventions aimed at fostering cultural sensitivity among healthcare providers. This observation is essential in addressing the broader issues of disrespect and abuse documented in the existing literature (Rodríguez-Almagro et al., 2019).

As noted in the introduction, previous studies have identified persistent and significant issues within maternity care, such as poor communication, exclusion from decision-making processes, various forms of verbal and physical abuse, and substandard hygienic conditions in maternity wards (Stanković, 2017; Cvitkušić, 2021). Our findings corroborate these concerns, with content analysis revealing that nearly half of the respondents cited interactions with hospital staff—particularly communication with doctors and nurses, as well as concerns regarding professionalism or negligence—as their most unpleasant memory during hospital care. In addition to staff interactions, respondents also highlighted poor hospital conditions and hygiene as some of their most negative experiences.

However, our results also illuminate a critical and under-explored issue within maternity care: the influence of personal connections and informal financial compensation. In all four countries under study, nearly half of the respondents reported utilising personal connections to secure proper care. Furthermore, in Serbia, Bosnia and Herzegovina, and Montenegro approximately half of the respondents acknowledged offering informal financial compensation to hospital staff, a practice far less prevalent in Croatia. These findings underscore the need for more in-depth, focused research on bribery and preferential treatment, given the potentially severe implications they may have on the manner in which healthcare providers (mis) treat women.

An additional, significant area that has been insufficiently studied in these countries is postpartum depression. Our results indicate that approximately one-third of women experienced symptoms of postpartum depression, yet only a small fraction sought professional help or received support from their social networks. Prior studies have indicated a correlation between postpartum depression and the use of forceps or vacuum extractions during delivery, as well as with undergoing a cesarean section, particularly when women had planned and desired a natural birth (Juric, 2020). As Juric (2020) noted, postpartum depression is a critical issue that impacts various aspects of a mother's life and the well-being of those in her surrounding environment, with the potential to evolve into chronic depressive episodes. Given the considerable challenges faced by women experiencing postpartum depression, the prevalence of this condition and a lack of help-seeking behaviour among our respondents warrant further investigation and a deeper understanding, particularly regarding the potential link to women's interactions with medical staff and the procedures performed during childbirth.

Study Limitations and Directions for Future Research

This study serves as a foundational step for future research on factors influencing the childbirth experience. Conducted entirely online using convenience sampling, it likely reflects a sample biased towards women of certain ages, socioeconomic statuses, cultural backgrounds, and digital literacy levels. As a result, the findings may not fully represent the population of women who gave birth between 2006 and 2021 in Serbia, Montenegro, Croatia, and Bosnia and Herzegovina. Additionally, the reliance on descriptive analysis limits the study's ability to explore relationships between variables that could influence the childbirth experience in greater depth.

An additional limitation of this study pertains to the sociodemographic variables, as certain key aspects were not included in the questionnaire. Notably, the type of living situation (urban or rural) and the specific age of the respondents were omitted, with the participants only able to select from predefined

age ranges. This omission restricts the potential for comparisons based on sociodemographic factors. We acknowledge these shortcomings in the design of the questionnaire and recommend that future research incorporate a comprehensive range of sociodemographic variables to enable more thorough and accurate analyses.

Despite these limitations, the study offers valuable insights and opportunities for further exploration. Future research could apply factor analysis to uncover underlying dimensions affecting childbirth experiences and use inferential statistics to examine correlations between key variables. Such analyses might identify specific areas of hospital care requiring improvement to enhance antenatal and postnatal care

To address current sample biases, systematic sampling should be employed to ensure more inclusive and representative findings. Furthermore, qualitative analysis of open-ended responses—only conducted on 10% of the current dataset—should be expanded to the entire dataset. This would provide deeper insights into mothers' lived experiences and highlight targeted areas for intervention. Moreover, the comparison between the countries may be influenced by the sample imbalance, as the majority of respondents are from Serbia.

By addressing these gaps, future research can build on the present findings to develop a more nuanced understanding of childbirth experiences and propose strategies to improve maternity care across the region.

Acknowledgement

The authors would like to acknowledge Assistant Professor Ana Milojković, PhD for her invaluable guidance and support throughout the research process.

References

Bryanton, J., Gagnon, A. J., Johnston, C., & Hatem, M. (2008). Predictors of women's perceptions of the childbirth experience. *JOGN Nursing, 37*(1), 24–34. https://doi.org/10.1111/j.1552-6909.2007.00203.x

Chalmers B. (1997). Childbirth in eastern Europe. *Midwifery, 13*(1), 2–8. https://doi.org/10.1016/s0266-6138(97)90026-7

Chalmers, B., Mangiaterra, V., & Porter, R. (2001). WHO principles of perinatal care: The essential antenatal, perinatal, and postpartum care course. *Birth (Berkeley, Calif.), 28*(3), 202–207. https://doi.org/10.1046/j.1523-536x.2001.00202.x

Cvitkušić, R. (2021). *Iskustva opstetričkog i ginekološkog nasilja nad trudnicama i rodiljama u hrvatskim bolnicama* (Doctoral dissertation, University of Zagreb, Faculty of Humanities and Social Sciences, Department of Sociology).

Freedman, L. P., Ramsey, K., Abuya, T., Bellows, B., Ndwiga, C., Warren, C. E., Kujawski, S., Moyo, W., Kruk, M. E., & Mbaruku, G. (2014). Defining disrespect and abuse of women in childbirth: a research, policy and rights agenda. *Bulletin of the World Health Organization, 92*(12), 915–917. https://doi.org/10.2471/BLT.14.137869

Gibbins, J., & Thomson, A. M. (2001). Women's expectations and experiences of childbirth. *Midwifery, 17*(4), 302–313. https://doi.org/10.1054/midw.2001.0263

Goodman, P., Mackey, M. C., & Tavakoli, A. S. (2004). Factors related to childbirth satisfaction. *Journal of Advanced Nursing, 46*(2), 212–219. https://doi.org/10.1111/j.1365-2648.2003.02981.x

Hodorogea, S., Kazakbaeva, C., & Maciulevicius, A. (2023). Assessment of the quality of maternal and neonatal services in Montenegro. *Curatio International Foundation*. Tbilisi: UNICEF.

Leeds, L., & Hargreaves, I. (2008). The psychological consequences of childbirth. *Journal of Reproductive and Infant Psychology, 26*(2), 108–122. https://doi.org/10.1080/02646830701688299

Pérez-Escamilla, R., Pollitt, E., Lönnerdal, B., & Dewey, K. G. (1994). Infant feeding policies in maternity wards and their effect on breast-feeding success: an analytical overview. *American Journal of Public Health, 84*(1), 89–97. https://doi.org/10.2105/ajph.84.1.89

Perrotte, V., Chaudhary, A., & Goodman, A. (2020). "At least your baby is healthy": Obstetric violence or disrespect and abuse in childbirth occurrence worldwide: A literature review. *Open Journal of Obstetrics and Gynecology, 10*(11), 1544. https://doi.org/10.4236/ojog.2020.10110139

Pittrof, R., Campbell, O. M., & Filippi, V. G. (2002). What is quality in maternity care? An international perspective. *Acta Obstetricia et Gynecologica Scandinavica, 81*(4), 277–283. https://doi.org/10.1034/j.1600-0412.2002.810401.x

Rodríguez-Almagro, J., Hernández-Martínez, A., Rodríguez-Almagro, D., Quirós-García, J. M., Martínez-Galiano, J. M., & Gómez-Salgado, J. (2019). Women's Perceptions of Living a Traumatic Childbirth Experience and Factors Related to a Birth Experience. *International journal of Environmental Research and Public Health, 16*(9), 1654. https://doi.org/10.3390/ijerph16091654

Rodríguez, C., & Rivières-Pigeon, C. D. (2007). A literature review on integrated perinatal care. *International Journal of Integrated Care, 7*(3). https://doi.org/10.5334/ijic.202

Sekulic, N. (2016). Culture of giving birth: Research on the issue of violence against women during delivery. *Sociologija, 58*(suppl. 1), 259–286. https://doi.org/10.2298/soc16s1259s

Sen, G., Reddy, B., Iyer, A., & Heidari, S. (2018). Addressing disrespect and abuse during childbirth in facilities. *Reproductive Health Matters, 26*(53), 1–5. https://doi.org/10.1080/09688080.2018.1509970

Stankovic, B. (2017). Women's Experiences of Childbirth in Serbian Public Healthcare Institutions: a Qualitative Study. *International Journal of Behavioral Medicine, 24*(6), 803–814. https://doi.org/10.1007/s12529-017-9672-1

Tabaghdehi, M. H., Kolahdozan, S., Keramat, A., Shahhossein, Z., Moosazadeh, M., & Motaghi, Z. (2019). Prevalence and factors affecting the negative childbirth experiences: a systematic review. *The Journal of Maternal-Fetal & Neonatal Medicine, 33*(22), 3849–3856. https://doi.org/10.1080/14767058.2019.1583740

Tavakoli, M., Emami, A., & Mirsaeedie, L. (2020). Environmental Factors Affecting Mother in the Maternity Ward (Case Study: Four Hospitals in Golestan Province). *Iran University of Science & Technology, 30*(1), 54–65. http://ijaup.iust.ac.ir/article-1-490-en.html

Živković Šulović, M. (2024). Korišćenje bejbi frendli usluga u porodilištima u Srbiji. *Srpski medicinski časopis Lekarske komore, 5*(1), 27–43. https://doi.org/10.5937/smclk5-49641