LEX

Luxemburger Experimental-praktikum Journal

Journal of Psychological Research

Vol. 19, Jahrgang 2025

Université du Luxembourg
Bachelor in psychology (BAP)

Luxemburger Experimentalpraktikum Journal
Vol. 19, Jahrgang 2025

Table of Contents

Investigating smishing in Luxembourg: How people perceive SMS Phishing Attempts

Bruna Fernandes Alves, Jo Lorang, David Zhang

Supervisor: Dr. Anastasia Sergeeva

Luxembourg has experienced an increase in registered cases of phishing attacks. Our study aimed to investigate how people perceive SMS Phishing Attempts, which means analyzing the psychological responses of individuals. Interviewing people with Luxembourgish phone number in person, where data were collected from a sample of 10 participants aged between 21 and 69. The study employed quantitative and qualitative methods. The results are a beginning to this topic and will help for further research in the different aspects, such as possible prevention to combat smishing attacks. The study's limitations, including a small sample size, selection bias, and the fact that our participants are vigilant. Future research with larger and diverse samples is needed to validate these initial insights.

1. Introduction

1.1 What is Smishing?

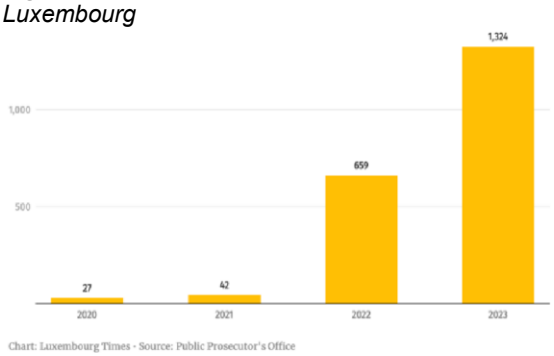
Phishing is a type of fraud where scammers try to trick people into sharing personal information like passwords or bank details. They pretend to be from trusted companies and send fake emails, websites or messages. This can lead to identity theft and financial loss.

Smishing is phishing through text messages (SMS). Scammers send fake texts that look like they are from real companies. Their goal is to get people to share personal details or click on harmful links.

1.2 Problematic of smishing in Luxembourg

Smishing has become an increasingly frequent and concerning topic, as evidenced by the graphic (see figure 1, (Luxembourg Times, 2024)) about the reported cases over recent years. In 2020, only 27 cases were officially registered, while in 2023 that number increased to 1,324. This increase highlights that smishing is not only a growing global issue but also a significant concern within Luxembourg. There is still a notable lack of data and studies on this topic in the Luxembourgish context, despite its relevance – a gap that motivated the development of our research.

Figure 1: *Number of new phishing attacks registered in Luxembourg*



1.3 Goals of our study

The limited existing research on smishing tends to focus primarily on the technical and cybersecurity facets, often overlooking the human and psychological aspects involved. However, understanding how individuals emotionally and cognitively respond to smishing attempts is essential for developing more effective prevention

strategies. Our study aims to explore how people interpret and react to smishing messages: how they experience such messages, how they evaluate them, and what factors influence their decisions to fall for those messages or ignore them.
By exploring the psychological reactions and behavioural patterns associated with smishing, particularly in a country where research on the topic remains limited. We hope to contribute meaningful insights that can support public awareness initiatives and guide the development of more effective protective strategies.

2. Research Questions

RQ1: How is the subjective experience with smishing in Luxembourg?
RQ2: What are the different tactics of smishing? *RQ3:* What are the possible solutions from people's perspective?

3. Methodology

3.1. Participants and recruiting

We conducted a lab-based study in the user-lab of the University of Luxembourg in Belval. We had the following eligibility criteria: (a) age 18 or over, (b) personal experience with use of mobile phone (c) proficiency in English. (d) Having Luxembourgish nationality or living /working in the country most of their life. We recruited 10 aged from 21-69 and with variety of professional experiences[e.g.] Most of them reported everyday use of the phones. 5 participants identified themselves as women and 5 as men. We used our professional and social networks to reach our participants. We also implemented snowballing method.

3.2. Procedure

Based on previous qualitative studies investigating user perceptions of phishing and smishing, we identified core themes such as trust, message plausibility, and emotional response. For instance, Ferreira et al. (2015) explored how end-users assess the credibility of suspicious messages and highlighted the role of contextual cues in decision-making. Similarly, Yasin et al. (2019) conducted a comprehensive review of social engineering strategies and user vulnerabilities, emphasizing the importance of personal and situational factors in threat perception. More recently, Schöps et al. (2024) examined user responses to simulated phishing scenarios in controlled environments, offering insights into how realistic threats are interpreted and processed. These studies informed the structure and content of our interview questions and helped ensure our approach was grounded in relevant empirical and theoretical work. The collected key points were reformulated into questions to enhance the understanding of the requirements and furthermore guided the development of our 10 research questions. Those 10 research questions were each used to create even more user-based questions to help us develop our interview questions. A document of the research questions can be found in the Appendix.

Before starting with the final questions for our semi-structured interview, an information sheet with a transparent description of what our study is about, our data collection and storage and the participants rights have been created. A detailed version of this document provided to participants can be found in the Appendix. A written consent form to receive consent for the participation on our study and a signature from the participant was also formulated. The participants received a 10 € gift voucher as a compensation of their time. This step is

important to respect and follow the ethical rules. This document can also be found in the Appendix.

In addition, we conducted an extended literature review and examined relevant theories in order to theoretically ground our interview questions and to ensure a certain degree of conceptual guidance and control. The theories we worked with were as follows:

1. The elaboration likelihood model focusing on peripheral and central processing (Petty & Cacioppo, 1986),
2. The persuasion principles focusing on urgency, scarcity, authority, social proof, etc. (Cialdini, 2001),
3. The protection motivation theory focusing on risk perception and adaptive behavior (Rogers, 1975),
4. The self-determination theory focusing on motivation, competence and autonomy
5. And the dual processing theory focuses on intuitive and analytical thinking (Kahneman, 2011).

After that, we designed and conducted a semi structured interview focusing on participants' encounters with smishing SMS. Several versions were created before finishing our final version. Our study protocol included questions about participants' experiences with smishing in Luxembourg, their strategies for distinguishing between the two types of SMS (legitimate or fraudulent), their perceptions of the tactics used by fraudsters and what they think which solutions could be implemented to effectively combat smishing and support users in recognizing and handling such threats. The full study protocol is provided in the Appendix.

A questionnaire with a small number of questions to collect personal background information for a better understanding of the participants responses and opinions was additionally created. We were asking about their gender, their age, their professional background, how often they use their mobile phone and how often they receive SMS messages, both personal or from organizations, in a week or in a month. It was sent by mail along with the 10€ gift voucher after the interview. All our mails got confirmed afterwards by the participants. The document with the questionnaire can be found in the Appendix.

Simultaneously we were also working on creating categories based on real smishing messages taken from Smishing Dataset I:

Phishing SMS Dataset from Smishtank.com (Timko and Rahman, 2024) and Grumbletext (n.d.) for our analytical research, although these were not from Luxembourg and we compiled them in a word document. In order to avoid a selection bias, we put the document containing the samples in ChatGPT and asked it to create a list containing 90 samples that are randomly selected from the main data set.

To make sure the categories were reliable; we used SPSS to check the agreement between different raters. This helped us define clear and consistent categories for smishing content. Those categories supported the identification of real smishing examples which we demonstrated to our participants to trigger the participants reflection on different tactics. In detail, we asked them what tactics they think the fraudster tried to use, what their feelings are about the strategies, if they saw something like that before and for which people they think this strategy could be the most effective.

The interview was conducted in the user-lab with at least two researchers and our empirical research supervisor present. All

answers were recorded by a microphone and afterwards analyzed inductively by the researchers.

3.3. Analysis

We used a combined inductive-deductive approach to analyse transcripts of the qualitative interviews. We used a combined deductive and inductive approach to analyse the interview transcripts. In the deductive part, we worked with predefined codes based on existing theories (like persuasion or protection

motivation). This means we looked in the data for ideas that we already knew from the literature. This helped us stay focused on our research questions and check how well the theories matched what participants said. In the inductive part, we allowed new ideas and patterns to emerge from the data itself. We used open coding, which means reading the transcripts closely and labelling meaningful segments of text without predefined categories. In many cases, we used participants' own words to label these ideas, helping us stay close to their personal experiences and perspectives. Combining both approaches gave us a strong structure while also letting us stay open to new and surprising findings.

Our main goal was to identify the factors that influence why people perceive SMS as
smishing and understand the measures people believe should be applied to tackle this fraud; as the study was performed in Luxembourgish context we try to discover which factors of
cultural context of the country (e.g. high level of digitalisation and
multilingualism) contributed to the susceptibility or protection from smishing
attacks. The process of analysis was as follow:

1. Each member of research group read all the transcripts and marked the findings according to the answers to RQs as well as making preliminary coding scheme
2. All four authors make a discussion session where they discuss the findings and finalise the scheme
3. When three authors separated the following analysis according to RQ's; the preliminary findings are presented in the results sections
4. The final coding and analysis were discussed between all four authors

4. Smishing Strategies

4.1. Goal of research

The goal for the qualitative research consists in revealing the different strategies used by fraudsters to trick people into interacting with the SMS messages. For our research we took inspiration from the article Principles of Persuasion in Social Engineering and Their Use in Phishing (Ferreira et al., 2015) in creating our smishing strategy categories.

4.2. Evaluation

4.2.1. Findings of 1st trial

This list of 90 samples is individually evaluated by 3 people. For the evaluation we assigned 1 of the 5 categories established by Ferreira et al. (2015) to each sample based on how well the content of the message matched the description of the category and at times 2 categories are assigned to a sample when the meaning of the message was rather ambiguous. After our first trial we noticed that not every category was used as frequently as some others (i.e. the categories Liking, Similarity & Deception,

Distraction and Authority were more frequent, whereas Commitment, Reciprocation & Consistency (CRC) and Social Proof (SP) have barely seen any use). We also noticed that the existing categories were too broad and general, and they often didn't fully fit the message they were assigned to. We came up with 5 preliminary categories that were more specific, and they were chosen to cover messages that didn't fit the existing categories. The following categories were established: "Financial and Banking scams" (FB), "Fake competition and Prizes" (FP), "Subscription scams" (S), "Fake customer support & delivery scams" (FD) and "Job & identity theft scams" (JI). The 5 new categories were then associated with the categories by Ferreira et al. (2015) to maintain a certain level of validity:

- FB □ Authority & Social Proof
- FP □ Liking, Similarity & Deception
- S □ Commitment, Reciprocation & Consistency
- FD □ Distraction & Authority
- JI □ Liking, Similarity & Deception

4.2.2. Findings of 2nd trial

On our 2nd trial, we noticed that once again certain categories were overused (FB, S, FD) while some others didn't see any use at all (JI, FP). So, we redefined the underused ones whilst we specified the overused ones, since they were still too general. We then came up with 6 more refined categories: loss (urgent messages threatening the loss of money, a package or an account), opportunity (messages offering money, free vacations or a job opportunity), social (messages appealing to the need of romance, friendship and altruism), competition (messages offering the quizzes or tests to assess your knowledge about a certain topic), authority (messages sent by the government, the bank or a renowned company). We've once again associated the newly defined categories with the ones established by Ferreira et al. (2015):

- Loss □ Distraction
- Opportunity □ Distraction
- Social □ Liking, Similarity, Deception
- Competition □ Distraction
- Authority □ Authority and Commitment, Reciprocity & Consistency

4.2.3. Findings of 3rd trial

On our 3rd trial with the newly established categories on a new list of 90 samples. We compared our results and noticed that the categories "Authority", "Social", "Loss" and "Opportunity" had a decent level of agreement between the 3 raters, but we also noticed that while "Authority" and "Social" were well defined and fitting. Although the categories "Loss" and "Opportunity" saw frequent use, more than half of the messages were assigned to them which indicates that they are still too broad and required more refinement. So, we kept the categories "Authority" and "Social" (renamed to "Social bonding"), split up "Loss" into "Scarcity" and "Urgency" while we split up "Opportunity" into "Compensation" and "Lucky Winner". We redefined "Competition" into "Curiosity" as the former was too specific. We made sure to give the categories clearer definitions in order to discriminate the messages more efficiently. So, we came up with the following descriptions: Authority (messages incorporating trusted orangs to have people act without questioning the context), Social bonding (messages aiming to exploit a person's emotional needs, i.e. romance, friendship and altruism), Curiosity (messages aiming to trigger interaction by offering novelty or by presenting a seemingly trivial offer +

any gifts that don't involve money), Scarcity (messages presenting a limited stock of a desirable item, thus enticing user to act quickly),
Urgency (messages containing elements of time pressure or potential vital losses), Lucky winner (messages aiming to make the user feel special about winning a big price) and Compensation (messages offering a high amount of money that seemingly requires a small effort to claim,
i.e. answering a survey, a job that pays exceptionally well or a recompense for past efforts).

4.2.4. Findings of 4th trial

On our 4th trial we once again used a new list of samples, and we noticed a high agreement between the 3 raters for most of the categories. So, we decided to run the results through SPSS in order to check for reliability. First, we transformed the findings of our categorization into sheet format which was suitable for SPSS. Then in SPSS we checked the kappa value of the overall agreement, the Z-statistic and the confidence intervals of our findings.

Table 1

Inter-rater agreement of categories

Overall Agreement			Asymptotic 95% Confidence Interval	
	Kappa	z	Lower Bound	Upper Bound
Authority	,689	6,535	,482	,895
Curiosity	,241	2,284	,034	,447
Lucky Winner	,584	5,445	,374	,794
Social Bonding	,788	7,478	,582	,995
Urgency	,135	1,277	-,072	,341
Compensation	,259	2,460	,053	,466
Scarcity	-,023	-,216	-,229	,184

The categories Authority, Lucky winner and social bonding have performed fairly well. Reason for that is that the sample contained many messages with these elements which allows the odds to even out.

The problem with urgency performing so badly was because it was often used as secondary category or totally omitted from the get-go. In another testing with different samples, it saw more usage and the results were also showing a high % of agreement.
For Curiosity and Compensation there was consistently a confusion between these two and I'd argue that the categories per se are good, but they require stricter definitions.
Scarcity being the worst performance wise, one would think it should be discarded but in during the whole testing it was only used twice, hence why the given result of -0.23 might be a distorted due to a too small sample size. It could be argued that this category is redundant and that it should be either discarded or revised.
We decided to split up "Authority" into "Trivial issues" and "Urgent issues" as it was still too broad and general, since that category contained 30% of the messages. Then we discarded "Scarcity" as most messages contained an element of scarcity which renders this category redundant. "Social bonding" remained the same, since it was well defined and also showed a high inter-rater agreement. Category wise "Lucky winner" and "Curiosity" remained the same, but we gave them a clearer definition in order to separate them more clearly from the other categories. Lastly, we renamed "Compensation" to "Reciprocity" and gave it a more polished definition.
In the end, we received the following descriptions for the categories:

- Reciprocity: Big compensation for small efforts (i.e. survey's, big job offers, quizzes with big money offers), compensation for past efforts (tax refund, reparation for technical issues), surreal deals at an absurdly low cost (discounts)
- Curiosity: Presentation of unusual information (i.e. news headlines, quizzes

with no monetary gain, …), obscure links or gifts with vague descriptions, not involving an exact amount of money

- Social bonding: Messages appealing to social needs such as romance, friendship (invitations to groups, "friendly" messages), altruism (charity, donation)
- Lucky winner: Being chosen as the main winner for huge prizes (money, vacations, valuable items like iPhone, jewellery) or for huge upgrades (to a service) for free
- Verification (common): Notification about a parcel in transit, an account suspended (from Amazon, Netflix, …) or a message about a verification process
- Urgent issues (urgent): Messages about suspension of an account (bank accounts, credit card) or the expiry of governmental documents

For our 4th trial we applied the same method as before and after running the results through SPSS in order to check for inter-rater agreement.

Table 2

Inter-rater agreement of categories

Overall Agreement			Asymptotic 95% Confidence Interval	
	Kappa	z	Lower Bound	Upper Bound
Urgent Issues	.909	14.929	.789	1.028
Verification	.908	14.922	.789	1.027
Lucky Winner	.540	8.877	.421	.660
Social Bonding	.936	15.374	.816	1.055
Curiosity	.701	11.524	.582	.821
Recompense	.759	12.478	.640	.879

4.3. Conclusion

We deemed the results as satisfactory and subsequently chose some of the messages assigned to the categories as examples that are meant to be presented to our interview participants in order to assess the user's opinion and perception about the different smishing tactics.

5. Results

5.1. Findings RQ1:

In Luxembourg, smishing is very common and increasingly sophisticated, although participants still perceived many messages as flawed, noting issues such as bad grammar, vague or out-of-context wording, strange links, or foreign numbers. During the interview, people's reactions ranged from confusion and anger to growing indifference because they got used to it over time. Most people now delete or ignore messages they identify as smishing. Some people are blocking the messages, especially when there are multiple attempts using the same phone number. However, people are rarely reporting them due to unclear, timeconsuming processes and the belief that reporting is pointless.

5.2. Findings RQ2:

The majority of smishing messages were generic and mimicked Luxembourgish companies such as POST, LuxTrust and CNS, thus creating a trusted environment. The phrasing and presentation of the messages would fall under the category of "Trivial issues" as they entice the user to complete seemingly simple tasks by clicking on a link. In order to avoid suspicion the messages are sent in the local administrative languages such as French and Luxembourgish. The smishing is most dangerous when the given context matches the recipient's life situation which creates an additional layer of trust. Occasionally messages disguised as official Police or bank writings imposed a tight deadline in order to trigger emotional response (panic and fear) which in return increases the chance of interacting with the message. In rare cases there were messages that praised the recipient for winning a huge prize money or congratulated them for being accepted to a

job application. These messages were once again phrased in a very general manner and lacked personalization,
i.e. the person's name, detailed information
about the person's life situation.

5.3. Findings RQ3:

During the interviews, various groups were identified as particularly vulnerable, including older adults, young people/students, individuals facing language barriers, people with limited digital experience, and even those in financially precarious situations. Several potential solutions were suggested to address these vulnerabilities. These included awareness campaigns through different channels, such as advertisements on social media, public transportation (e.g., trams), radio, and television. Another important strategy involves education and training, both in schools and in professional environments. By integrating digital safety trainings programs in school and workplace, individuals of all ages can develop the necesssary skills to navigate the digital world more safely and confidently. In addition, technical measures were proposed, such as the implementation of systems that allow users to report suspicious activity, as well as filtering and blocking systems to prevent exposure to harmful content. Warning mechanism would also be a preventive measure to reduce smishing. Lastly, the importance of support within the family context was emphasized, particularly through the role of "tech guides"; family members who are more digitally experienced and can provide guidance and assistance in navigating technology safely to those who are less comfortable with it.

6. Highlights

People need straightforward (easy-to-use) mechanisms to report suspicious messages, such as clear feedback to know their report was received and taken seriously. This would help to build trust and encourages more people to report such cases.
In Luxembourg, the presence of multilingualism provides challenges and opportunities in the fight of smishing. It can be harder to share messages with a great and correct grammar, but using different languages also allows messages to reach more people in ways they understand best. Additionally, the increasing use of artificial intelligence (AI) in Luxembourg introduces also creates new risks. AI can be exploited by scammers to create more convincing and personalized messages, making it harder for individuals to distinguish between the legitimate and the fraudulent messages. Finally, it is important to create different solutions for different groups because people have different needs and levels of experience with technology. There is no one- size-fits-all approach addressing the smishing threat. Tailoring solutions helps protect more people in a more effective way.

7. Limitations

Firstly, the small sample of 10 interviewees may not sufficiently represent the variations of opinions which a larger group would certainly surface. Further, any outlier position would disproportionally weight in the results. The small sample may therefore not be statistically relevant on all dimensions of the study.

Secondly, there was a selection bias because participants were recruited through our personal social networks and snowball sampling which may have resulted in a less diverse sample and therefore less diverse responses and perspectives.
Lastly, we mostly had vigilant participants. This may have led to a limited understanding of the challenges and vulnerabilities associated with smishing. The populations that are potentially more susceptible for such attacks were underrepresented which means that again, important aspects of user experience may be missing from the data.

8. Future directions

To gain a better and more detailed insight on this topic in Luxembourg, longitudinal research should be done to explore further changes in awareness, perception and behavioral responses to smishing attacks.
In addition, investigating cultural and linguistic influences may represent a compelling subject for further study because Luxembourg's multilingualism may influence the perception of smishing and message interpretation.
Lastly, it is essential to develop more precise methods for identifying at-risk populations for a better understanding of their vulnerabilities and tailor preventions strategies to support them accordingly.

9. Acknowledgement

Lastly, we want to express our sincere gratitude for our supervisor Dr. Sergeeva Anastasia for the amazing support and feedback throughout the entire process of our project.
We hereby declare that we are the sole authors of the work entitled, and we confirm that this work was independently written in our own words. AI was employed solely to assist with grammar and stylistic improvements. We also confirm that this work has not been submitted elsewhere.

10. Literature

Abbasi, A., Dobolyi, D., Vance, A., & Zahedi, F. M. (2021). The phishing funnel model: A design artifact to predict user susceptibility to phishing websites. Information Systems Research, 32(2), 410–
436. https://doi.org/10.1287/isre.2020.0973

Butavicius, M., Parsons, K., Pattinson, M., & McCormac, A. (2016). Breaching the human firewall: Social engineering in phishing and spear-phishing emails. arXiv. https://arxiv.org/pdf/1606.00887

Cialdini, R. B. (n.d.). Dr. Robert Cialdini's seven principles of persuasion. Influence at Work. https://www.influenceatwork.com/7-principles-of-persuasion/

Das, S., Nippert-Eng, C. E., & Camp, L. J. (2022). Evaluating user susceptibility to phishing attacks.
Information & Computer Security, 30(1), 1–18. https://doi.org/10.1108/ICS-12-2020-0204

Ferreira, A., Coventry, L., & Lenzini, G. (2015). Principles of persuasion in social engineering and their use in phishing. In Proceedings of the 3rd International Conference on Human Aspects of

Information Security, Privacy, and Trust (pp. 36–47). Springer. https://orbilu.uni.lu/bitstream/10993/20301/1/FerreiraAna-CameraReady.pdf

Heindrichs, T. (2024, November 13). Explosion in phishing cases in Luxembourg. Luxembourg Times. https://www.luxtimes.lu/luxembourg/explosion-in-phishing-cases-inluxembourg/26604747.html

Norris, G., Brookes, A., & Dowell, D. (2019). The psychology of internet fraud victimisation: A systematic review. Journal of Police and Criminal Psychology, 34(3), 231–245. https://doi.org/10.1007/s11896-019-09334-5

Open Newcastle University. (n.d.). Protection motivation theory. https://open.ncl.ac.uk/theories/10/protection-motivation-theory/

Rahman, M. L., Timko, D., Wali, H., & Neupane, A. (2023). Users really do respond to smishing. In Proceedings of the 13th ACM Conference on Data and Application Security and Privacy (pp. 49–60). Association for Computing Machinery. https://doi.org/10.1145/3577923.3583640

Rajivan, P., & Gonzalez, C. (2018). Creative persuasion: A study on adversarial behaviors and strategies in phishing attacks. Frontiers in Psychology, 9, Article 135. https://doi.org/10.3389/fpsyg.2018.00135

ScienceDirect. (n.d.). Dual-process theory. https://www.sciencedirect.com/topics/psychology/dualprocess-theory

Schöps, S. (2024). Presented at the USENIX Security Symposium 2024. https://www.usenix.org/conference/usenixsecurity24/presentation/sch%C3%B6ps

SpringerLink. (n.d.). Chapter 4. https://link.springer.com/chapter/10.1007/978-3-319-20376-8_4

Tally, A. C., Abbott, J., Bochner, A., Das, S., & Nippert-Eng, C. (2023). What mid-career professionals think, know, and feel about phishing: Opportunities for university IT departments to better empower employees in their anti-phishing decisions. Proceedings of the ACM on Human-Computer Interaction, 7(CSCW1), Article 113. https://doi.org/10.1145/3579547

Yasin, A., & Abuhamad, M. (2019). Contemplating social engineering studies and attack scenarios: A review study. ResearchGate. https://www.researchgate.net/publication/333843845_Contemplating_social_engineering_studies_and_attack_scenarios_A_review_study

Appendix

10 Research Research Questions:

1) Does a stressful environment influence the susceptibility / success rate of phishing?
- Is a person under stress more likely to process information only on surface-level?
- Do people under stress only rely on certain keywords in order to judge the legitimacy of a message?
- Does workload induced stress make a person more likely to consider the phishing message as yet “another” task to be dealt with?
- Is a person under stress more likely to do a simple favor?

2) How do phishing scams impact a consumers trust in Luxembourg’s online services?
- What impact have phishing scams on consumer trust in Luxembourg’s online services?
- How does consumer trust in Luxembourg’s digital platforms change after exposure to phishing scams?
- What impact can phishing scams have on consumer trust in Luxembourg’s online services?

3) What psychological factors influence individuals in Luxembourg to fall victim to phishing scams?

4) Do different attributes of messages have an impact on the success rate of smishing?
- How do urgency cues (limited time offers, threats, rewards) influence user responses to smishing messages?
- Does the inclusion of personal information (name, location) increase the likelihood of a successful smishing attack?
- What role do grammar, spelling, and message length play in determining a smishing message's credibility?

5) Can training be relied on as a mechanism to improve user's security awareness/knowledge?

6) Does the socio-cultural context of Luxembourg play a role in which kinds of phishing strategies are applied?
- Is there a correlation between GDPs per capita and the amount of phishing attempts?
- Do most phishing messages pretend to be associated with bank insurance companies?
- Is the method of gaining confidential information adapted to the advanced digital processes?
- Is it more based on simple manual input rather than via calls or text messages?
- Do they resort to the Luxembourgish language to build trust on a more personal level?

7) How does a user’s impulsivity in making decisions affect their ability to judge the safety of a link in an SMS?
- How does a user’s tendency for impulsive decisions affect their ability to evaluate the security of links in SMS messages?
- How does user impulsivity influence the evaluation of SMS link security?
- To what extent does impulsivity affect a user’s capacity to judge the reliability of links in SMS messages?

8) Which factors make people detect smishing attacks? / How does trust influence user’s likelihood of clicking on phishing links, and what factors contribute to this trust?
- How does prior cybersecurity training impact an individual’s ability to recognize smishing attempts?
- To what extent do generational differences affect the likelihood of recognizing smishing messages?

- How do spam filters and AI-driven detection methods affect user perception and response rates to smishing?

9) To which extend do phishing attacks exploit the emotional vulnerability of users and how successful are the different strategies?
- To what extent do phishing attacks take advantage of users' emotional vulnerability, and how effective are different strategies?
- How do phishing attacks exploit emotional vulnerability, how effective are different approaches?
- In what ways do phishing attacks manipulate users' emotions, and which strategies prove most successful?

10) What were the factors that affected users' perceptions behavior and actions after receiving the smishing?
- What factors influence whether users report a smishing attempt or ignore it?
- Does being a victim of smishing make individuals more likely to adopt proactive cybersecurity measures?
- What role does cognitive dissonance play in how users justify their actions after interacting with a smishing message? / Do users experience guilt or embarrassment after falling for smishing, and how does this shape future behavior?

INFORMED CONSENT FORM in relation to "SMIR (Smishing: Methods, Impacts and Reactions)" project

I have read the information sheet and I have been informed by Researcher (dr. Anastasia Sergeeva) orally and in writing (see pages 2 and following) about the nature and the potential consequences and risks of the above-mentioned research project (the Research Project), and I have had sufficient opportunity to ask any questions.

I understand that my data will be collected and used in connection with this Research Project and to enable publication of the research results.
I have been informed that I am entitled to withdraw my consent to participate in the Research Project at any time without giving a reason and without negative consequences to myself. Furthermore, I may object to further processing of my personal data and/or samples or request that these be deleted. I may do so by contacting dr. Anastasia Sergeeva or prof. dr. Christine Schiltz (principle investigator)

Please tick the appropriate boxes in the table below:

I consent to the collection and use of my personal data in relation to the Research Project • Yes • No
I agree to the data I provide being archived on-line in the University of Luxembourg servers and being used in pseudonymised form for other research in the area of studies about privacy concerns in Extended Reality applications beyond the Research Project • Yes • No
I consent to my interviews being recorded in audio format for the purposes of the Research Project • Yes • Only if my identity is not disclosed • No
I consent to my personal data, as described in the information sheet, being processed for the purposes of understanding the smishing attacks presented in the pseudonymized form in scientific publications and conference presentations of findings. • Yes • Only if my identity is not disclosed • No

I am happy to be contacted after • Yes • No

this Research Project to ask whether I would be interested in taking part in a follow-up study

I voluntarily agree to take part in this Research Project. PARTICIPANT
Last name: First name:
Date of birth:

Place & date:

Signature of the participant:

Signature of the legal representative:

RESEARCHER
I have informed the above-mentioned participant orally and in writing (see pages 1 and following) about the nature and the potential consequences and risks of the Research Project, and I have given the participant the opportunity to ask any questions.

In addition, the participant has received a copy of the information sheet(s) and of this consent form.
Name:

Place & date:

Signature of the researcher:

Information sheet
Title of research project: Investigating smishing: How people react to SMS Phishing Attempts
Acronym: SMIR (Smishing: Methods, Impacts, and reactions)

What is this study about?
This study is about phishing, a type of fraud where scammers try to trick people into sharing personal information like passwords or bank details. They pretend to be from trusted companies and send fake emails, websites or messages. This can lead to identity theft and financial loss.
We focus on smishing, which is phishing through text messages (SMS). Scammers send fake texts that look like they are from real companies. Their goal is to get people to share personal details or click on harmful links.
In this study, we want to understand how people react to smishing messages. We will discuss:
1. Your experience with smishing – what influenced your reactions and actions.
2. Different smishing tactics – how they work and your thoughts on them.
3. Possible solutions – what should be done to prevent smishing in the future
The interview will take about 45-60 minutes, and you will receive 10 € gift voucher as compensation for your time.

What data do we collect?
During the interview, we will record your voice as you answer our questions. We will also collect your opinions on examples of smishing messages we show you.

In addition to this, we will ask you for some background information to help us better understand your responses. This includes:
- What is your age?
- What is your gender?
- What is your professional background?
- How often do you use your mobile phone?

• How often do you receive SMS messages (both personal and from organizations)? (e.g., per week or per month)?

Your data is pseudonymized, as your name is replaced with a code to protect your identity. A secure key list that links your name to the code will be stored separately. Your response will only be used for research purposes to gain a better understanding of smishing.

How will we use your data?
The collected data will be used only for research. It may be included in scientific publications (journals or conferences) but always in an anonymized and pseudonymized format.

How do we store and protect your data?
Your name and other personal details will not be linked to your recordings.
The data will be stored securely and only accessible to researchers working on this study.
After 36 months, all personal data will be deleted automatically.
Pseudonymized data (without names) will be kept for up to 10 years for research purposes.

What are your rights? You have the right to:
Access, correct or delete your data at any time.
Withdraw from the study whenever you want without giving a reason. Receive a summary of the study's results if you are interested.
For more details on your rights, visit the University of Luxembourg's Data Protection page: https://wwwen.uni.lu/university/data_protection/your_rights
If you have an questions, feel free to ask before participation. Thank you for your time !

If you have a concern about any aspect of your participation, please raise this with the researcher, or with :
Dr. Anastasia Sergeeva Postdoctoral researcher (+352) 621 384 503
Anastasia.sergeeva@uni.lu

Questionnaire

• What is your age?
• What is your gender?
• What is your professional background?
• How often do you use your mobile phone?
• How often do you receive SMS messages (personal C organizations)? (in a week or in a month)

Interview protocol

A. Introduction and Consent
- Welcome participants and introduce the research objective.
- Explain the procedure and anonymized data storage.
- Obtain consent for participation and recording of the interview.

B. Introductory questions

1. Are you still using SMS or prefer using other messages services (like whatsapp)?
2. Have you ever heard about the term „smishing "? (if not, explain it to the person)

C. Experience with smishing

3. Have you ever encountered some SMS messages that seemed suspicious to you?

- Can you tell us about your experience with this sms ?
- What made the SMS seem suspicious to you?
- Can you describe the situation in which you received the SMS (home, workplace etc.…)
- How did you react after receiving a smishing SMS and realizing it was scam? What were your actions? Can you tell maybe step by step how you interacted with them?
- What did you feel after an encounter with such a message?
- Did you think about it later-on?
- Is your trust in Luxembourg's digital services less due to smishing? (Only if the message tried to impersonate some legal institution/banking institution)
- Do you believe that having more time would have changed your actions? (only if the description of the incident included being affected by the message)
- Have you/would you report the incident to any officials? Why/why not?

D. Strategy Appraisal

4. What typically affects your decision to open an SMS or not? * Open or klick on the link?
5. What influenced your judgment about whether the message was fraudulent or legitimate? **
6. What strategies do you think they use to influence people? Why do you think they do it and how?

E. Examples of Smishing:

1. What do you think the fraudster tried to use here?
2. What are your feelings about this strategy? Did you saw something like that before? (it can also be the loop back to the previous block of questions)
3. For which people this strategy can be the most effective?

F. Perceived self-efficacy in identifying smishing

7. Have you ever suspected a message to be fraudulent that in the end turned out to be legitimate?
8. Have you ever avoided or even ignored an important message from your bank/insurance company/the state, because you deemed it as fraud?

G. Impact and behavioral changes

9. Do you think reporting it makes a difference or is important to prevent further consequences?
10. What would we need to combat such messages? What do you think can be done against such SMS?

H. Conclusion

11. Is there anything else you would like to share about your experience with smishing?
12. Do you have any other remarks?

(name of the company, the specific deal you are awaiting, language or the way the proposal is presented) *
(e.g., sender information, URL, language used, urgency, etc.) **

Description of categories:
Reciprocity:

- Small effort/favors – big wins (i.e. survey's, big job offers, quizzes with big money offers)
- compensation for past efforts (tax refund, reparation for technical issues)
- huge offer at a low cost (discounts)

Curiosity:
- presentation of unusual information (i.e. news headlines, quizzes with no monetary gain, …)
- obscure links
- gifts with vague descriptions, not involving an exact amount of money

Social bonding:
- social needs
- romance
- friendship (invitations to groups, « friendly » messages)
- altruism (charity, donation)

Lucky winner:
- being chosen as the main winner
- huge prices (money, vacations, valuable items like iPhone, jewelry)
- huge upgrades (to a service) for free

Verification/confirmation (common)
- parcel in transit
- account suspended (from Amazon, Netflix, …)
- verification processes

Urgent issues (urgent)
- accounts suspended (bank accounts, credit card)
- expiry of governmental documents

Set of examples from categories

1. Amazing news!, your application has been accepted. Finish the process: [link to the external site]

2. Hi, are you who played football together last time? I'm Lisa

3. Hey [PERSON'S NAME], help us answer a fast survey. We will give you a free watch in return.[link to the external site] Stop to end

4. Pending Delivery: Hi, your package containing tracking code 71623881LS5 is waiting for you to confirm the shipment address: [link to external site]

5. [link to external site] [NAME OF THE COMPANY] Profile locked because of unusual activities, kindly restore. Reply STOP to unsubscribe

6. BANK [NAME OF THE BANK] ALERT SERVICES: Your EDD prepaid Visa debit card '4427' has been locked. Click to reopen: [link to external site]

7.[NAME OF THE COMPANY]Customer: Claim $100: [link to external site]

Does remembering nature decrease loneliness in older people?

Marianna Kubiak, Sarah Kupke, Paula Althaus

Supervisor: Prof. Dr. Adrian Stanciu

This quantitative study investigates whether recalling a positive nature experience through guided imagination can reduce feelings of loneliness and enhance nature connectedness. To assess potential changes, measures of loneliness and nature connectedness were collected before and after the intervention. Participants first completed a pre-questionnaire, followed by a four-minute audio guided imagination and then a post-questionnaire using the same items. While it was hypothesized that recalling a natural experience would lead to positive effects — particulary, decreased loneliness and increased nature connection – the results did not support this assumption. Although the changes were not statistically significant, an unexpected trend was observed: participants reported slightly increased feelings of loneliness and decreased nature connectedness after the intervention. These findings indicate that guided recall of nature-based experiences may influence emotional states, but not necessarily in the intended direction. This complexity might be explained by the way the nature memories were recalled —whether they were purely about being in nature or emotionally charged situations that happened to occur in natural settings. Future studys are needed to investigate how to decrease loneliness in older adults.

1. Introduction

Loneliness is a major issue among older adults, with one in three reporting feelings of loneliness (WHO-5, 2023). It is best described as a “distressing feeling that accompanies the perception that one’s social needs are not being met by the quantity or especially the quality of one’s social relationships” (Hawkley & Cacioppo, 2010, S. 218). This sense of disconnection can have serious consequences for physical and mental health. Loneliness has been linked to increased risk of depression, suicide, Alzheimer’s disease, and other adverse outcomes (Hawkley & Cacioppo, 2010). Among older adults, it is additionally associated with cardiovascular disease, stroke, diabetes, cognitive decline, anxiety, and reduced quality of life (Mikton et al., 2021). Given the rapidly growing proportion of older adults worldwide, addressing loneliness in this population is both urgent and essential for promoting healthy aging.

PROTECTIVE ROLE OF NATURE.

A growing body of research suggests that exposure to natural enviroments plays a protective role in mental-health. Nature can foster calm, reduce stress and increase feelings of connection and belonging, which are essential in combating loneliness (Lavelle Sachs et al., 2024). Moreover, regular contact with nature, has been linked to preventive effects against mental health disorders, as sleep improvement and stress relief were recorded, which are related to depression and anxieties. Nature connectedness—the subjective sense of feeling emotionally and cognitively linked to the natural world—has been positively correlated with the frequency and quality of nature-based experiences (Cleary et al., 2020). This is supported by the Biophilia hypothesis (Kellert & Wilson, 1995), which posits that humans have an innate evolutionary need to connect with nature. As such, fostering nature connectedness may be one meaningful pathway to support mental health and combat loneliness.

ACCESSIBLE INTERVENTIONS.

Despite the benefits, many older adults have limited access to natural environments due to mobility issues, environmental constraints, or institutional living settings. Moreover, societal trends such as urbanization and increased screen time have contributed to decreased time spent in nature across all age groups (Bratman et al., 2019). Given these challenges, there is a growing interest in developing nature-based interventions that do not rely on physical access to nature. Kalantari et al. (2022) evaluated a virtual reality (VR) nature intervention with older adults. They found improvements in mood and engagement through immersive nature experiences. However, such interventions often require expensive technology and may not be easy accessible for all populations. An alternative approach—guided nature imagery—may offer a low-cost, accessible solution. Drawing on memory and imagination, asking individuals to recall a positive experience in nature may activate similar emotional and physiological benefits to those gained from real nature exposure.

THE PRESENT RESEARCH.

The current study aims to investigate whether imagining a positive experience in nature can reduce loneliness and increase nature connectedness in older adults. This approach is grounded in prior findings suggesting that indirect (non-physical) contact with nature can evoke beneficial psychological states similar to those elicited by actual nature exposure (Kalantari et al., 2022). Additionally, reconnecting with meaningful nature memories may activate the innate human need to connect with nature (Kellert & Wilson, 1995), thereby fostering a greater sense of belonging and reducing feelings of isolation. Participants in this study completed an online pre-questionnaire assessing loneliness, nature connectedness, and well-being. They then listened to a brief, four-minute audio-guided imagination designed to help them recall a positive experience in nature. Following the intervention, the same questionnaire was completed again to assess any changes.

Two hypotheses were tested: (H1) Imagining a positive nature experience decreases loneliness in older adults; and (H2) Imagining a positive nature experience increases nature connectedness in older adults.

2. Methods

2.1 Participants

The present study used a quantitative method to explore the study objectives. N = 45 participants aged fifty years or older participated in our study by filling out the questionnaire. Participants were recruited via snowball sampling. We started with a small group of initial participants, which were primarily family members. These participants referred our study to others who fit our inclusion criteria. Inclusion criteria were proficiency in German while exclusion critera was the presence of severe mental illness. Upon completing the questionnaire participants could put down their email address for a chance of winning a voucher worth. Informed consent was obtained at the experiment's onset which emphasized voluntary participation. Ethical approval was given from the University of Luxembourg.

2.2 Study design

The study consisted of our intervention and two data collections via three questionnaires. The intervention consisted of a short guided meditation with the intending to allow the participants to re-experience a positive memory in nature (e.g. Appendix). Before the intervention participants filled out questionnaires and therefore provided the baseline data which was to be influenced by our intervention. After the intervention participants again filled out the same questionnaires to enable the measurement of changes at pre- and post-intervention. The intervention consisted of a guided imagination, helping the participants recall a positive nature experience. Both the intervention and questionnaire were piloted before the data collection.

2.3 Materials and Procedures

2.3.1 DEMOGRAPHICS

At the beginning of the study, participants indicated their age, gender and occupational status.

2.3.2 NATURE RELATEDNESS SCALE

To assess individual differences in participants' perception of their relationship with nature, the Nature Relatedness Scale (NRS) was included in the questionnaire (NRS – Nisbet & Zelenski, 2013). Participants were asked to rate each item on a 5-point Likert scale ranging from 1 ("I strongly disagree") to 5 ("I strongly agree"). To ensure the reliability of the measurements, the internal consistency of both scales was evaluated separately at baseline and after the intervention. All scales demonstrated strong internal consistency, with Cronbach's alpha exceeding the accepted threshold of .80, indicating that the items within the scale consistently measure nature relatedness.

2.3.3 THE WORLD HEALTH ORGANIZATION-FIVE WELL-BEING INDEX (WHO-5)

To ensure that our participants did not suffer from any severe mental illness and potentially could distort our results, the WHO-5 Well-Being Index was included to exclude any participants that showed abnormalities regarding their well-being. Furthermore, the index was included to observe potential changes in mental well-being before and after the intervention. The WHO-5 is a self-report measure assessing mental well-being relating to the previous two weeks and consisting of five items (WHO-5 - World Health Organization, 2024). Participants were asked to rate each statement on a 6-point scale ranging from 5 ("All of the time") to 0 ("At no time") and higher scores indicating a better mental well-being. Ensuring the reliability of the measurements, the internal consistency of both scales was evaluated separately at baseline and after the intervention. All scales demonstrated strong internal consistency, with Cronbach's alpha exceeding the widely accepted threshold of .80. This indicates that the items within each scale consistently measure mental well-being.

2.3.4 UCLA LONELINESS SCALE

Changes in loneliness after the intervention were of main interest, therefore, to assess participants loneliness at baseline and after intervention, the UCLA Loneliness Scale 6-item Short Form was used (RULS-6, Wongpakaran et al., 2020). The scale consists of six questions asking how often individuals feel certain ways related to loneliness and social connectedness. Participants were asked to rate each question on a 4-item scale, ranging from 1 ("never") to 4 ("always") with higher scores indicating higher levels of loneliness. To ensure the reliability of the measurements, the internal consistency of both scales was evaluated separately at baseline and after the intervention. The scales demonstrated a strong internal consistency, as Cronbach's alpha exceeds the accepted threshold of .80, allowing the assumption that the items within each scale consistently measure loneliness.

2.4.5. PROCEDURE AND OVERALL STUDY DESIGN

In the beginning of the questionnaire, participants were introduced to the study's procedures, were provided with a declaration of consent, and generated a unique, pseudo-anonymized code. Afterwards, participants answered demographic questions on their respective digital device. Baseline nature relatedness was evaluated using the NRS, loneliness was measured using the RULS-6 and mental well-being was assessed with the WHO-5 Well-Being Index. After answering the first round of questions, participants were asked to find a quiet place to listen to the guided imagination. Subsequently, participants were asked to fill out the second round of questionnaires, which were the same as in the first round, as well as provided in the same order (NRS, RULS-6 and WHO-5 Well-Being Index post). Finally, participants were quickly

debriefed about the use of their data and the purpose of the study.

2.4.6. ETHICAL CONSIDERATIONS

Participants were informed about the procedure of the study and their right to withdraw from the study at any time without any negative consequences as participation was voluntary. Participants suffering from any severe mental illness were exempt from participating in the study to prevent any potential harm being caused.

PRIVACY AND CONFIDENTIALITY. Participants were informed that the data would be collected anonymously which would under no circumstances allow the identification of the participant. The participants were also assured that the data would only be used for research purposes.

3. Results

3.1. Participant sample

The sample consisted of *N* = 45 participants out of which 5 were excluded for either dropping out of the study or missing data. The sample had *N* = 32 participants identifying as female (71.1% of the sample) and *N* = 7 self-identifying as male (15.6% of the sample), as well as one participant identifying as diverse (2.2% of the sample), who we included in the measure but excluded in the further exploration of the data. The mean age was 60.92 (*SD* = 6.749) years old, with the oldest participant being 76 years of age and the youngest 50 years of age. Most of our participants still worked full-time (*N* = 17) or were already retired (*N* = 14).

Table 4

Demographic Information

Variable		*N*	%
Age		60.92	
Sex	Female	32	17.5
	Male	7	71.1
	Diverse	1	2.2
Occupation	Full-time	17	43.6
	Part-time	8	20.5
	Retirement	14	35.9

3.2 Descriptive Statistics

Descriptive statistics for well-being, loneliness and nature-connectedness at pre- and post-intervention are presented in Table 5.

Table 5

Descriptive statistics of outcome variables before and after intervention

Measures	Pre M (SD)	Post M (SD)
Well-being	3.01 (1.03)	3.01 (1.10)
Loneliness	1.94 (0.64)	1.93 (0.57)
Nature-connectedness	3.97(0.7)	3.93 (0.71)

3.4. Inferential Statistics

A repeated measures ANOVA was conducted to evaluate within-subject differences in outcomes before and after the intervention, allowing for assessment of changes attributable to the intervention over time and was used to test the hypotheses. To test for this we ran the repeated measures ANOVA for pre-intervention and post-intervention and checked whether or not there was a significant change in loneliness over time. The repeated measures ANOVA revealed no significant changes in loneliness over time ($F(1) = 1.581$, $p = 0.218$). Furthermore, the repeated measures ANOVA revealed no significant interaction between nature-connectedness after the guided imagination and changes in loneliness over time ($F(1) = 1.215$, $p = 0.279$), indicating that nature-connectedness cannot explain or predict changes in loneliness over time. Bearing no significant revelations regarding the repeated measures ANOVA, our first hypothesis (“Imagining a positive nature experience decreases loneliness in older adults”) had no evidence in support of it, as there was no significant result

regarding the interaction of nature-connectedness and changes in loneliness over time, as well as no significant change in loneliness over time in general.

3.5 Further exploration of the data

To explore any potential abnormalities or deviants in our collected data, we conducted further measurements. To determine if there is a statistically significant difference between the means of our data pre- and post-intervention, we evaluated this for nature-connectedness, well-being and for loneliness.

NATURE CONNECTEDNESS.

Regarding the second hypothesis ("Imagining a positive nature experience increases nature-connectedness in older adults"), a paired-samples *t*-Test showed no significant differences between nature connectedness pre-intervention (M = 4.0364; SD = 0.6698) and post-intervention (M = 3.9333; SD = 0.7065), $t(32) = 1.254$, $p = 0.219$. determine if there is a statistically significant difference between the means of two related groups.Therefore not confirming our second hypothesis.

WELL-BEING.

A paired samples *t*-Test showed no significant difference in the mean for pre-intervention well-being (M = 3.0742; SD = 1.04745) and post-intervention well-being (M = 3.0468; SD = 1.09171), $t(30) = 0.507$, $p = 0.616$.

LONELINESS.

The paired samples *t*-Test showed no significant difference in the pre-intervention scores of loneliness (M = 1.895; SD = 0.5707) and post-intervention scores (M = 1.927; SD = 0.5725), $t(31) = 0.937$, $p = 0.338$.

GENDER DIFFERENCES.

As we yielded no evidence allowing for the confirmation of the hypotheses, we explored the data further to assess for potential explanations or abnormalities. Using a independent-samples *t*-Test en enabling us to verify if there is a significant difference between the mean of the two group conditions, we checked for potential gender differences. Although the result were not significant, meaning that the mean did not differ significantly between women (pre: M = 4.0258; SD = 0.74; post: M = 4.0615; SD = 0.65) and men (pre: M = 3.7143; SD = 0.54; post: M = 3.5333; SD = 0.78) (Pre-intervention $t(36) = -1.047$, $p = 0.302$; Post-intervention $t(30) = -1.725$, $p = 0.095$) there is, as illustrated in Figure 1, a possible trend visible that men's nature-connectedness decreased more post-imagination than women's, with men potentially being more sensitive to the intervention than women.

An independent-samples *t*-Test was also conducted for possible gender differences regarding loneliness before and after the imagination. These results were also not significant ($t(35) = 0.049$, $p = 0.962$; $t(29) = -0.423$, $p = 0.676$) indicating no significant differences between the means for women (pre: M = 1.9389; SD = 0.68; post: M = 1.9467; SD = 0.63) and men (pre: M = 1.9524; SD = 0.56; post: M = 1.8333; SD = 0.32) but as Figure 2 illustrates there is a trend visible that men's loneliness seemingly increased more than women's after the intervention.

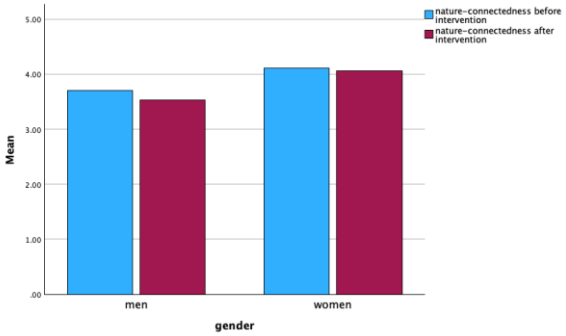


Figure 1: changes in nature-connectedness

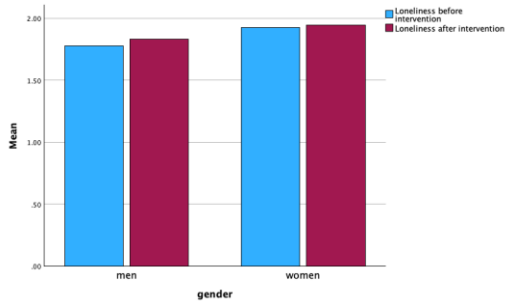


Figure 2: changes in loneliness

It should be noted, that, again, the independent-samples *t*-Tests yieled no significant results but we can observe a general trend in our data for possible changes.

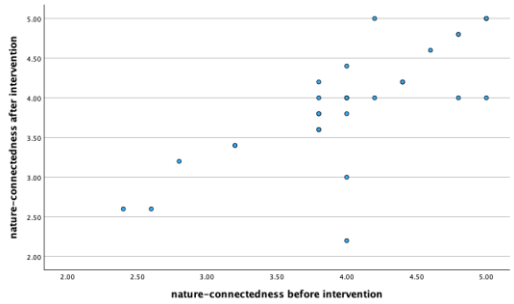


Figure 3: scatter-plot of the correlation of nature-connectedness before and after intervention

The correlations between the items before and after intervention were also analyze to understand the consistency of the individual participant's score before the intervention and after the intervention and while the correlations for the composite score for loneliness and well-being before and after were highly positively correlated with loneliness before and after the intervention being correlated with r = 0.949, N = 32 (p = <0.001) and well-being with r = 0.961, N = 31 (p = <0.001) the correlation of the composite score for nature-connectedness deviated. The correlation for nature-connectedness before and after intervention was still highly positive (r = 0.766, N = 33, p = <0.001) but did show a deviation, indicating that compared to well-being and loneliness, nature-connectedness decreased after the intervention. This is being presented in Figure 3.

4. Discussion

Although our hypotheses were not confirmed, our data revealed some interesting oberservations, which lead us to a discussion. Contrary to our expectations, loneliness increased and nature connection decreased after our intervention. It is important to note, however, that these changes were not statistically significant. Nevertheless, we aim to explore possible explanations for these trends.

SOCIAL AND EMOTIONAL IMPACT. Given that we do not know the experiences of the participants' imagination, we explored possible reasons for the observed increase in loneliness. One hypothesis is, that our intervention might have a social-emotional impact. Participants may have recalled a positive experience in nature associated with a social context (e.g., with children, friends or family). Research by Capaldi et al. (2015) has shown that nature connectedness is often intertwined with social connectedness. So remembering a nature-connected situation could evoke memories of a social connected situation. Those meomories could elicit nostalgic emotions such as longing or sadness. This emotional reaction aligns with the self-discrepancy theory (Higgins, 1987), which suggests that reflecting on emotionally significant experiances may highlight the gap between one's current and desired social state, leading to negative emotions, such as loneliness. Our intervention could remind the person of the current social isolation. This brings up the following question: Can remembering a nature experience trigger emotions more strongly tied to social disconnection than to environmental connectedness, thereby contributing to increased feelings of loneliness?

GENDER-RELATED.
An interesting trend emerged when examining at gender differneces: Men appeared to be more affected by the intervention, showing greather changes after the intervention in lonliness and nature connection compared to women. One possible explanation is that men may be less frequently exposed to meditative or emotionally reflective environments, and

therefore respond more strongly to such interventions. This observation is consistent with findings that men often report lower baseline levels of emotional expressiveness and environmental concern (Zelezny et al., 2000), which might mean that nature-based or emotion-focused interventions affect them more strongly—either positively or negatively—due to unfamiliarity.
Moreover, we observed that men were overall less connected to nature both before and after the intervention. Previous research suggests that women tend to report higher emotional affinity toward nature (Tam, 2013). Future studies should aim for a more balanced sample to better investigate gender-related responses to nature-based interventions.

LIMITATIONS.

We acknowledge the limitations that may have influenced our results of our study. Our small and imbalanacd sample size likely contributed the lack of statistically significant results. Additionally, the study was not standardized, as we had no control over the enviroments in which participants completed the intervention. Some may not have been alone, or they might have been in noisy sourroundings, which burdend them to fully engage in the guided imagination. It is also possible that some participants did not listen to the audio intervention, for example due to not working headphones or the absence of headphones. This lads to another limitation: We did not investigated subjective experiances participants had during the imagination. We do not know what they imagined. Did they recall an experience in nature or did they remember an experience outside within a social context?

FUTURE RESEARCH.

Based on our observations, serval improvments and extensions can be proposed for future research. It may be helpful to integrate more real sensory elements into the intervention for example - smell of the ocean or the rain, touching the texture of leaves, moss or sand, or listening to brids singing. Such sensory approaches might facilitate a deeper connection to nature and increase the effecttivness of the intervention.

Furthermore, it would be valuble to explore whether repeated exposure over a longer time frame (for example: weekly sessions for serval months) could lead to higher capacitys of imagination and therefore to lasting increases in nature connection and decreases in loneliness. A long term study could reveal whether imagination-based nature interventions can foster a greater sense of belonging. Finally, incorporating open-ended questions or qualitative interviews to explore participants's experiences of the imagination intervention could provide deeper inisghts into the mechanisms underlying the observed effects.

5. Conclusion

Lonliness is an increasing problem in modern society. As the population of older adults grow, finding accessible and effective interventions becomes increasingly urgent. Nature can play a role in decreasing loneliness. Our study investigated whether simply remembering a positive experience in nature could decrease loneliness and foster nature connection. We conducted an online-based questionnaire measuring loneliness, nature connection and well-being, before and after our guided imagination as intervention. Our results were not significant and our hypotheses were not confirmed. However, despite a small and imbalanced sample, non-standardized conditions and a lack of insight into participants' subjective experiences during the intervention, some interesting observations emerged. It is possible that imagination-based nature experiences evoke emotional responses tied to social memories, which may have produced effects contrary to our expectations—loneliness increased and nature connectedness decreased. Gender-related differences in response to the intervention were also observed, suggesting the need for further investigation. Future research is needed to investigate the interplay between remembering nature experiences and social connectedness and nature connection. Understanding these mechanisms may lead to low-cost, accessible interventions that contribute to the prevention of loneliness.

6. References

Bratman, G. N., Anderson, C. B., Berman, M. G., Cochran, B., De Vries, S., Flanders, J., ... & Daily, G. C. (2019). Nature and mental health: An ecosystem service perspective. *Science advances*, *5*(7), eaax0903.

Capaldi, C. A., Passmore, H. A., Nisbet, E. K., Zelenski, J. M., & Dopko, R. L. (2015). Flourishing in nature: A review of the benefits of connecting with nature and its application as a wellbeing intervention. *International Journal of Wellbeing*, *5*(4).

Cleary, A., Fielding, K. S., Murray, Z., & Roiko, A. (2020). Predictors of nature connection among urban residents: Assessing the role of childhood and adult nature experiences. *Environment and Behavior*, *52*(6), 579-610.

Hawkley, L. C., & Cacioppo, J. T. (2010). *Loneliness matters: a theoretical and empirical review of consequences and empirical review of consequences and mechanisms*. Annals of behavioral medicine: a publication of the Society of Behavioral Medicine, 40(2), 218–227.

Higgins, E. T. (1987). Self-discrepancy: a theory relating self and affect. *Psychological review*, *94*(3), 319.

Kalantari, S., Bill Xu, T., Mostafavi, A., Lee, A., Barankevich, R., Boot, W. R., & Czaja, S. J. (2022). Using a nature-based virtual reality environment for improving mood states and cognitive engagement in older adults: A mixed-method feasibility study. *Innovation in aging*, *6*(3), igac015.

Kellert, S. R., & Wilson, E. O. (1995). The biophilia hypothesis.

Lavelle Sachs, A., Kolster, A., Wrigley, J., Papon, V., Opacin, N., Hill, N., Howarth, M., Rochau, U., Hidalgo, L., Casajuana, C., Siebert, U., Gerhard, J., Daher, C., & Litt, J. (2024). *Connecting through nature: A systematic review of the effectiveness of nature-based social prescribing practices to combat loneliness*. Landscape and Urban Planning, 248, 105071. https://doi.org/10.1016/j.landurbplan.2024.105071

Mikton, C., Fried, L., Holt-Lunstad, J., Hussein, S., Lähesmaa, J., Ong, A., Prina, M., & Victor, C. (2021). *Social isolation and loneliness among older people: advocacy brief* (By A. Officer, E. Krug, World Health Organization, United Nations Department of Economic and Social Affairs, International Telecommunication Union, & UN Women) [Advocacy brief]. World Health Organization. https://iris.who.int/bitstream/handle/10665/343206/9789240030749-eng.pdf?sequence=1

World Health Organization. *Commission on Social Connection.* (2023, November 15). https://www.who.int/groups/commission-on-social-connection

Wongpakaran, N., Wongpakaran, T., Pinyopornpanish, M., Simcharoen, S., Suradom, C., Varnado, P., & Kuntawong, P. (2020). *Development and validation of a 6-item Revised UCLA Loneliness Scale (RULS-6) using Rasch analysis*. British journal of health psychology, *25*(2), 233–256. https://doi.org/10.1111/bjhp.12404

World Health Organization. (n.d.). *The World Health Organization-five well-being index (WHO-5)*. World Health Organization. https://www.who.int/publictions/m/item/WHO-UCN-MSD-MHE-2024.01.

Zelezny, L. C., Chua, P., & Aldrich, C. (2000). Elaborating on gender differences in environmentalism-statistical data included. *Journal of social issues*, *56*(3), 443-445.

7. Acknowledgements

We would like to express our sincere gratitude to the University of Luxembourg for funding and making it possible for us to conduct this study. Our deep appreciation also goes to our dedicated tutor, Adrian Stanciu, whose invaluable guidance and continuous support contributed greatly to the success of our research.

8. Appendix

INTERVENTION TEXT. Wir laden Sie im Rahmen dieser Übung nun ein, eine bequeme Sitzhaltung zu finden. Schließen Sie nun sanft die Augen und atmen Sie tief ein und aus. Spüren Sie, wie sich bei jedem Atemzug Ihr Bauch hebt und senkt – beim Einatmen füllt sich Ihr Bauch mit Luft, und beim Ausatmen wird er wieder leer. Atmen Sie noch zweimal tief und bewusst ein und aus, und lassen Sie dabei alle Gedanken los.

Erinnern Sie sich nun an einen besonderen Moment draußen in der Natur. Vielleicht waren Sie am Meer, in den Bergen oder im Wald? Versuchen Sie, diesen Moment wieder lebendig werden zu lassen.

Rufen Sie die Geräusche Ihrer Erfahrung in der Natur wieder auf. Hören Sie die Vögel zwitschern oder das beruhigende Rauschen des Wassers? Vielleicht hören Sie das Rascheln von Blättern. Spüren Sie das Wetter wieder. Scheint die Sonne auf Ihr Gesicht oder spüren Sie kühlenden Regen, der sanft Ihre Haut berührt? Ist es ein warmer Sommertag, an dem die Luft den Duft von Blumen trägt, oder ein klarer Herbstmorgen, an dem der Boden von den letzten Regentropfen glänzt? Haben Sie in Ihrer Erinnerung kalte oder warme Hände?

Vielleicht spüren Sie noch den Sand zwischen Ihren Zehen oder den Geruch von nassen Blättern? Sehen Sie in Ihrer Erinnerung noch die Sonne, die auf dem Wasser glitzert, oder das sanfte Schaukeln der Bäume im Wind? Können Sie in Ihrer Erinnerung vielleicht einen Schmetterling beobachten, der durch die Luft flattert, oder eine Biene, die auf einer Blume landet?

Versuchen Sie, sich all diese Details vorzustellen und zu spüren: Wie hat sich Ihr Körper in diesem Moment angefühlt? Können Sie in dieses Körpergefühl wieder eintauchen? Haben Sie das Gefühl, mit der Natur verbunden und geerdet zu sein? Fühlen Sie sich ruhig und entspannt oder vielleicht glücklich und erfüllt von der Schönheit der Umgebung in Ihrer Erinnerung?

Nehmen Sie sich einen Moment, um diese Erinnerung in all ihren Details zu spüren und zu genießen. Wenn Sie bereit sind, atmen Sie noch einmal tief ein und aus und kehren langsam ins Hier und Jetzt zurück. Fangen Sie langsam an, Ihre Finger zu bewegen und Ihre Füße auf dem Boden wahrzunehmen. Wenn Sie so weit sind, öffnen Sie sanft Ihre Augen.

Effects of a One-Week Instagram Detox and Mindfulness on Body Image

Otilia Dragan, Elena Demuth, Kimberly Ingrid Gerneck, Alena Kožar, Anna Rita Sepe

Supervisor: Kriti Kelkar, Doctoral researcher

High use of visual platforms like Instagram has various problematic effects on health, well-being, and body image. Research suggests that interventions such as digital detox and mindfulness practice can mitigate these harmful effects. This online study experimentally tested the short-term effects of a one-week Instagram detox combined with a 10-minute daily mindfulness intervention on body image among young adults (18–34). Participants in the experimental group (N = 35) refrained from using Instagram for one week and completed brief daily mindfulness exercises, while participants in the control group (N = 35) continued their usual Instagram use. Body shape concerns, drive for muscularity, and appearance evaluation were measured at baseline (Time 1) and one week later (Time 2) for both groups using the platform SoSci survey. Data was analyzed using 2 (Group) x 2 (Time) mixed repeated measures ANOVAs. The daily mindfulness intervention was conducted on the “7Mind” app, and brief daily reflections on the process were recorded using the “m-Path” app. As predicted, the intervention had a significant effect, with a significant decrease in body shape concerns and drive for muscularity and a significant improvement in appearance evaluation. These findings suggest that a short “Instagram detox” combined with mindfulness may have promising benefits for young adults' body image..

Introduction

As of 2023, social media platforms reached around 60% of the global population, which led to a wide dissemination of highly idealised appearance standards and contributed to increased pressure to look as close as possible to these ideals (Markey & Daniels, 2022; Saiphoo & Vahedi, 2019; Statista, 2023; Thompson & Harriger, 2023).

Previous research showed that high social media use can lead to problematic outcomes such as neglecting important tasks, impairments in health and wellbeing, and difficulties in maintaining relationships, social activities, studies, or work (Kuss & Griffiths, 2017; Sun & Zhang, 2021). Frequent use was also associated with negative body image and lower self-esteem, notably among young women (Holland & Tiggemann, 2016; Saiphoo & Vahedi, 2019; Sherlock & Wagstaff, 2019).

Researchers have found that continuous social media engagement can also create “technostress” and fear of missing out (FOMO), leading to emotional exhaustion and compulsive checking behaviours (Neelakandan, 2024). Visual platforms like Instagram appeared to have a more harmful impact on body image than more text-based platforms such as Facebook (Vandenbosch et al., 2022).

High social media use strongly correlated with numerous mental health issues and cognitive difficulties including stress, generalised anxiety, depression and procrastination, especially in young people and adolescents (Keles et al., 2020). Conversely,

reducing the time spent on social media was shown to support emotional regulation and overall life satisfaction (Schmuck, 2022). Reduced exposure to idealised Instagram content (especially “thinspiration” and “fitspiration”) was shown to decrease body dissatisfaction, body-focused anxiety, and drive for thinness, whereas viewing such content was shown to lower self-esteem, positive mood, and body appreciation in young women (Hooper et al., 2024). Similarly, taking a one-week break from social media increased self-esteem and body image among young women (Smith et al., 2024).

1.1. Possible interventions against harmful social media effects

Digital detox is defined as the „voluntary reduction or temporary cessation of social media use“ (Neelakandan, 2024). Reducing screen time has been shown to reduce stress, anxiety, depression, and procrastination, while improving focus, sleep quality, and overall life satisfaction (Thomas et al., 2022; Schmuck, 2022). Particularly for young adults, who represent one of the most active and vulnerable user groups, digital detox interventions can protect mental health and body image by reducing exposure to unrealistic beauty standards (Smith et al., 2024).

Besides, mindfulness has been shown to moderate the effects of social media use on wellbeing. High trait mindfulness protects individuals from harmful outcomes by improving attentional control, present-moment focus, stress reactivity, non-judgmental bodily awareness, and emotional regulation (Kircaburun et al., 2019; Throuvala et al., 2020; Weaver & Swank, 2021). Individuals with high trait mindfulness tend to report less social media use and, when engaging with these platforms, they are more likely to share authentic content with less emphasis on idealised body representations (Baker, Krieger & Le Roy, 2016; Charoensukmongkol, 2016).

Conversely, low trait mindfulness is moderately correlated with higher risk for problematic social media use ($r = –.37$), poorer emotional regulation, difficulty sustaining focus on tasks, greater distractibility from notifications, and more habitual checking behaviors (Du et al., 2021).

Practicing daily mindfulness can help individuals observe their bodies without judgement or emotional reactivity, as studies have shown that it can reduce the tendency to engage in maladaptive behaviours and enhance understanding of one’s body image experiences (Stewart, 2004). Hooper et al. (2024) showed that integrating a 10-minute mindfulness exercise in combination with limited exposure to appearance-focused social media content helped buffer harmful social media effects. This was also highlighted by Miles et al. (2023), who found that mindfulness can address body image issues. According to Barrington et al. (2019), mindfulness can act as a protective factor that can be integrated into prevention and intervention for body dissatisfaction. Higher mindfulness can mediate between lower body comparison and higher body satisfaction.

The role of social media in body image can also be explained with three classical theories of social influence and self-perception: Thompson et al.'s Tripartite Influence Model (1999), Festinger's Social Comparison Theory (1954), and Fredrickson and

Roberts' Objectification Theory (1997).

We found Thompson et al.'s Tripartite Influence Model (1999) particularly relevant in this context, as it described the emergence of sociocultural appearance pressure originating from three main sources (peers, parents or family and the media). This pressure may lead to internalisation of a “thin” and/or a “muscular” ideal (dependent on gender and other socio-cultural factors) and to further appearance comparison with others. The increased internalisation of beauty ideals also amplifies body dissatisfaction. Social media platforms intensify processes proposed by the Tripartite Influence Model by facilitating continuous exposure to media and peer influences within highly interactive digital environments.

The Social Comparison Theory (Festinger, 1954) holds that individuals evaluate themselves by comparing themselves to others, especially when they perceive notable similarities. On social media, this happens frequently: users scroll through images of peers and influencers (though these may often be edited and retouched images), and they compare their own appearance. This often results in upward comparisons with images of idealised bodies, which may increase body dissatisfaction and drive for thinness or muscularity (Thompson & Stice, 2001; Dane & Bhatia, 2023). The internalisation of “thin” and “fit” ideals further reinforces negative self-evaluations and heightens the risk for low self-esteem and disordered eating behaviors (Dittmar & Howard, 2004; Hogue & Mills, 2019).
More recent meta-analysis evidence confirmed that online social comparison correlates significantly with body image concerns and eating disorder symptoms (Bonfanti et al., 2025).

The Objectification Theory (Fredrickson & Roberts, 1997) explains how in societies that sexualise and objectify bodies (especially women), people internalise an observer’s view of their own body (self-objectification). On image-focused social media platforms, body parts are often presented, edited, likes/comments reinforce appearance value, and users may start seeing their bodies as objects for others’ gaze. This process promotes body shame, habitual monitoring of appearance, and impaired connection to internal cues. Appearance evaluation and external approval thus become an important part of a person's self-concept. Recent studies of social media exposure such as Castellanos Silva and Steins’ (2023) show that this objectifying environment correlates with poorer body image outcomes.

According to these theories, we conclude that social media fuels social comparison in a potent way: the constant access to modified visuals, the likes and comments acting as social feedback, and a wide spectrum of peers and influencers to compare with keep the comparison always active.

As social media use was proven to have a strong influence on body image (American Psychological Association, 2023), it is crucial to investigate strategies to curb adverse effects of social media use.

To identify a feasible short-term intervention that can buffer against these negative effects, we implemented a combined approach, consisting of a brief Instagram digital detox

and daily mindfulness exercises. We believe that this combined intervention would produce stronger effects than a digital detox alone.

This study aims to investigate the effects of a one-week digital detox from Instagram use in combination with a mindfulness intervention on the body image outcomes among young people.

1.2. Hypotheses in current study

Based on the studies discussed in the introduction, we initially formulated seven hypotheses exploring different possible effects of the Instagram digital detox and the mindfulness intervention on body image, which we later consolidated into three concise hypotheses covering all key aspects of the original set.

H1: A significant Time × Group interaction is expected, with the experimental group showing improvements in body image (BI) outcomes over time.

H2: Significant between-group differences are expected at post-test, with the experimental group showing more positive BI outcomes than the control group.

H3: No change in BI outcomes is expected for the control group over time.

Methods

2.1 Participants

An a priori power analysis was conducted to determine the required sample size using G*Power software (version 3.1.9.6) (Faul et al., 2007). For an ANOVA (repeated measures, within-between interaction, 2 groups and 2 measurement points) with a medium effect size of .25, a significance level of $\alpha = .05$, and a statistical power of .80, the a priori analysis resulted in the required sample size of 34 participants. To recruit participants, a poster was created and shared on WhatsApp and Instagram. Participants ($N = 70$) aged between 18 and 34 years ($M = 24.95$; $SD = 3.41$), took part in the study. 65.7 % identified as female and 34.3 % as male. SoSci Survey randomly assigned these seventy participants to two conditions. The experimental group and the control group included both 35 participants (50%) each. Inclusion criteria were gender (male or female), fluency in English, age between 18 and 34, as well as active and regular Instagram use (at least 30 minutes per day).

2.2 Study design and Procedure

This quantitative study investigated digital detox from Instagram in combination with mindfulness as potential lifestyle interventions to improve well-being, especially the body image of young adults. The research adopted an online-experimental approach with a short-term longitudinal design (Pre-Post), to evaluate changes in body image concerns.

Data collection was carried out using two secure online platforms: SoSci Survey for pre- and post-intervention questionnaires, and the m-Path application for daily reflections on mindfulness activity. Ethics approval was received from the Ethics commission at the University of Luxembourg. All participants received written information about the study procedures and provided informed consent electronically, while filling out the SoSci survey, before beginning the baseline survey.

Participants were required to be active Instagram users (at least 30 minutes per day), fluent in English, and willing to participate in a one-week study. Participants assigned to the experimental group were instructed to refrain from using Instagram and to complete daily online tasks consisting of mindfulness audio exercises lasting 10 to 15 minutes. Participants were randomly assigned by SoSci survey to either the experimental group (Instagram detox plus daily mindfulness activity; N = 35) or the control group (usual Instagram use; N = 35).

To ensure privacy and allow accurate matching of pre- and post-intervention data, participants created a unique personal code when completing the initial SoSci Survey. This code served as the only identifier within the dataset, ensuring full pseudonymisation. It was temporarily linked to participants' email addresses for organisational purposes, including matching questionnaires across time points and enabling the distribution of compensation afterwards.

Although email addresses were stored during data collection, they were irreversibly deleted immediately after the completion of data collection. As a result, the final dataset cannot be linked to individual participants. Non-identifiable research data will be stored securely for ten years in line with institutional guidelines.

All questionnaire data collected through SoSci Survey and daily reflections submitted via m-Path were stored on password-protected servers located in Europe and processed in accordance with GDPR regulations. Participants retained full control over any data generated within external applications used during the study (m-Path and 7Mind). These app-based data were not accessible to the researchers and could be deleted by the participants at any time.

Participants were informed that they were free to withdraw from the study at any time without penalty by contacting the researchers via e-mail. They were also provided with contact information should they experience discomfort during participation and wish to seek psychological support.

The participants were told that the study is conducted to understand how social media use and a break from it affects well-being over time (1 week).

They had to fill out demographic questions, some questions related to their current body satisfaction, mindfulness and social media use as well as questionnaires assessing body image (Body Shape Questionnaire-8C ; Cooper et al., 1987; Evans & Dolan, 1993), Drive for muscularity (McCreary & Sasse, 2000) and Appear-

ance evaluation (subscale of the MBSRQ ; Cash, 1990) at baseline (Time 1).

On day 1, the experimental group received instructions to completely refrain from using Instagram for one week. Participants downloaded two apps, 7Mind (for the mindfulness task) and m-Path (for daily reflections on the mindfulness task). They started the “Foundations” Course in 7Mind.

From day 2 to day 7, participants were instructed to remain logged out of Instagram and to complete daily mindfulness meditation sessions on the 7Mind app, by following the “Foundations” course. Following each meditation session, they answered brief reflection questions on the m-Path app regarding their mindfulness experience. Daily notifications were sent to participants reminding them to complete their meditation and reflection activities.

On day 8, participants took a screenshot of their meditation activity page (showing total minutes, number of days and number of sessions completed), and they uploaded evidence of at least 70 minutes of meditation across all 7 sessions. They also received a final questionnaire link via email to complete the post-test measurements (identical to baseline measures).

The control group continued using Instagram and the internet as usual without any mindfulness intervention, and they also completed the same questionnaires as the experimental group at baseline (Time 1) and after one week (Time 2).

After completing the study, the participants received a 30 Euro GoGift Voucher.

2.3 Scales

2.3.1 Body Shape Questionnaire

The Body Shape Questionnaire-8C (BSQ-8C) was used to measure body shape concerns (Cooper et al., 1987; Evans & Dolan, 1993). The questionnaire is a common self-report instrument, which is intended to measure body shape issues. The measure examines more specifically symptoms of body dissatisfaction such as feeling overweight, self-consciousness about appearance, and the fear of gaining weight. It is especially suitable to measure short-term changes and has demonstrated good psychometric properties like reliability and validity across various samples and languages (Gao et al., 2025).

Participants responded to each item using a 6- point Likert-type scale ranging from 1-Never to 6-Always, indicating how often they experienced specific thoughts/feelings about their body shape over the past four weeks. Scores are calculated by summing all items. Higher scores reflect higher body shape concerns.

BSQ-8C showed good internal consistency, with $\alpha = .866$ at Time 1 and $\alpha = .860$ at Time 2.

2.3.2 Appearance evaluation

The Appearance Evaluation (MBSRQ-AE; Cash, 1990) subscale of the Multidimensional Body-Self Relations Questionnaire was used to assess participants' overall satisfaction with their physical appearance. This 6-item subscale measures general feelings of attractiveness, comfort with one's looks, and satisfaction with body features. The MBSRQ-AE (AE) is widely used in body image research and demonstrates strong psychometric properties, including high internal consistency and good construct validity across different populations.

Participants completed the AE subscale at baseline (Time 1) and post-intervention (Time 2) using the online SoSci Survey platform. Each item is rated on a 5-point Likert scale ranging from 1 = Definitely disagree to 5 = Definitely agree, with higher scores indicating greater appearance satisfaction. Scale scores were calculated by computing the mean of all items.

AE also showed good internal consistency, with α = .885 at Time 1 and α = .868 at Time 2.

2.3.3 Drive for muscularity

The Drive for Muscularity Scale (DMS; McCreary & Sasse, 2000) was employed to assess muscularity-oriented body image concerns. This 15-item scale captures both attitudinal and behavioral components, including dissatisfaction with current muscularity, preoccupation with increasing muscularity, and engagement in muscle-enhancing behaviors (e.g., exercise and diet). The DMS has demonstrated good psychometric properties in previous research, with strong reliability and validity across genders and age groups. Participants completed the DMS (DM) at baseline (Time 1) and post-intervention (Time 2) via SoSci Survey. Each item is rated on a 6-point Likert scale ranging from 1 = *Never* to 6 = *Always*. Items were reverse coded where necessary, so that higher total scores reflect a stronger drive for muscularity. Scale scores were calculated by summing all items.

DM demonstrated good internal consistency, with α = .854 at Time 1 and improving to α = .886 at Time 2.

2.4 Materials

2.4.1 7Mind

7Mind is a mobile application that was used to enhance mindfulness practice and support mental well-being. It provides an extensive range of mindfulness-based content, including guided meditations and audio-based courses. Participants in the Experimental Group listened to one mindfulness-related audio per day, over one week. To assess engagement with the application, anonymised usage statistics were collected from each participant's profile page. These metrics included the number of days on which mindfulness exercises were completed, the cumulative minutes of practice, and the total number of meditation sessions undertaken (Mestdagh et al., 2023).

2.4.2 m-Path

m-Path is a digital platform, designed to facilitate real-time, in-daily-life assessment and intervention. The platform allowed to implement repeated short surveys using Ecological Momentary Assessment (EMA) to capture participants' mood, thoughts, and behaviors in their natural context across time, often multiple times per day. m-Path automates key aspects of the data collection process, including scheduling assessments, sending reminders, and recording participant responses.

After downloading the application, participants linked their account to the corresponding researcher´s account via an integrated search function. Once connected, we distributed questionnaires directly to participants, and the collected responses were transmitted automatically and anonymously, ensuring both data security and confidentiality.

Results

3.1 Data analysis

The analyses were performed using IBM SPSS (Version 29). Descriptive statistics were computed to summarise participants' demographics, including sex, age, familiarity with Instagram content, and frequency of viewing appearance-focused content.

A 2 × 2 mixed-design ANOVA was then carried out, to examine the effects of time (pre vs. post) and group (experimental vs. control).

Each group included 35 participants, all of whom met the inclusion criteria. With Time (pre-intervention vs. post-intervention) as the within-subjects factor and Group (mindfulness vs. control) as the between-subjects factor, this analysis examined the interaction between the Instagram detox intervention and mindfulness practice over time.
Assumptions of homogeneity of variances were tested using Levene's test.

3.1.1 Preliminary analyses

Independent t-tests confirmed baseline equivalence between the groups. A series of t-tests showed that the experimental and the control group did not differ significantly on age, $t(68) = 0.313$, $p = .705$, gender distribution, $\chi^2(1, N = 70) = 0.254$, $p = .615$ and total time spent on Instagram per week, $t(68) = 0.756$, $p = .825$.
Baseline levels were also comparable for BSQ pre-test scores, $t(59.308) = –0.015$, $p = .988$, AE pre-test scores, $t(68) = 0.380$, $p = .705$, and DM pre-test scores, $t(68) = 0.277$, $p = .783$. Therefore, random assignment to condition successfully created equivalent groups on these variables.

Table 1

Time 1 and time 2 means (and standard deviations) by condition.

Measure	*1-week break & mindfulness condition (N=35) M (SD)*	*No break Condition (N=35) M (SD)*
BSQ		
Time 1	*23.97 (6.27)*	*24.0 (9.39)*
Time 2	*17.77* (3.57)*	*23.85 (8.55)*
DM		
Time 1	*35.89 (11.66)*	*35.11 (11.67)*
Time 2	*30.71* (11.27)*	*34.77 (11.29)*
AE		
Time 1	*3.40 (0.75)*	*3.33 (0.86)*
Time 2	*3.60** (0.68)*	*3.31 (0.79)*

Note. Range of possible scores: Body Shape Questionnaire: 34-204 Drive for muscularity:15-90; Appearance evaluation: 7-35. Asterisks denote significant within - group differences in means ($p < .05$; $p < .01$) over time.

3.1.2 Body image outcomes

Hypothesis 1: A significant Time × Group interaction is expected, with the experimental group showing improvements in body image (BI) outcomes over time.

For body shape concerns, a significant Time x Group interaction emerged, $F(1, 66) = 21.325$, $p < .001$, $\eta_p^2 = .244$ (Figure 1). The experimental group showed a significant decrease in body shape concerns over time (Table 1).

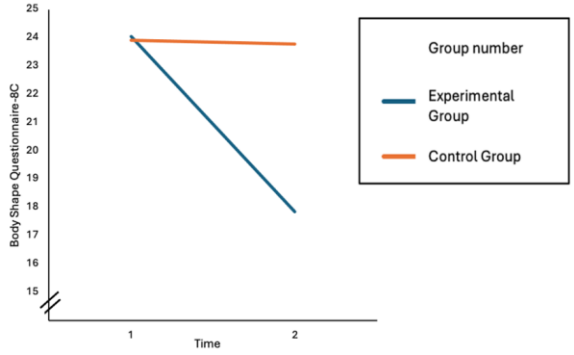


Figure 1: Total BSQ scores over time and by condition

For appearance evaluation, a significant Time x Group interaction emerged, $F(1, 66) = 5.234$, $p < .025$, $\eta_p^2 = .073$ (Figure 2). The experimental group showed a significant increase in appearance evaluation over time (Table 1).

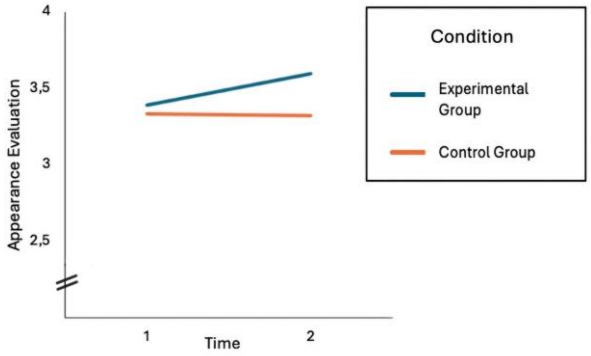


Figure 2: Total AE scores over time and by condition

For drive for muscularity, a significant Time x Group interaction emerged, $F(1, 66) = 7.682$, $p < .007$, $\eta_p^2 = .104$ (Figure 3). The experimental group showed a significant decrease in drive for muscularity over time (Table 1).

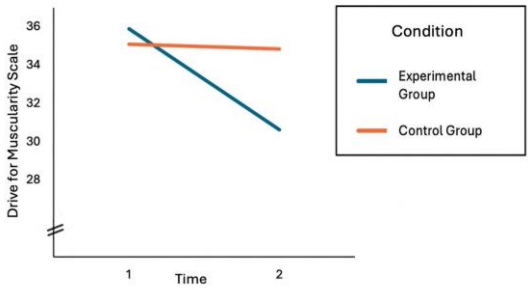


Figure 3: Total DM scores over time and by condition

Taken together, these findings provide support for H1.

Hypothesis 2: Significant between-group differences are expected at post-test, with the experimental group showing more positive BI outcomes than the control group.

Further, at post-test, the experimental group scored significantly lower than the control group for body shape concerns.
At post-test, the experimental group did not score significantly higher than the control group for appearance evaluation. At post-test, the experimental group did not score significantly lower than the control group for drive for muscularity.

These findings indicate that H2 could only be partially supported, as there was a significant difference between the experimental and control group post-intervention only in BSQ but not in AE and DM.

Hypothesis 3: No change in BI outcomes is expected for the control group over time.

No changes in body image outcomes were observed over time in the control group (Table 1). This provides support for H3.

Discussion

The current study investigates the effects of a one-week digital detox combined with mindfulness exercises on body image. Participants were randomly assigned to either take a one-week break from Instagram and engage in mindfulness exercises or continue their usual Instagram use (control condition). Consistent with our hypotheses, the findings revealed higher scores in Appearance Evaluation (AE) and lower scores in Drive for Muscularity (DM) and the Body Shape Questionnaire (BSQ) among participants in the experimental condition at posttest. The between-group differences were only significant for BSQ but not for AE or DM. No significant change could be observed in neither AE, BSQ nor DM in the control group at posttest.

In line with previous research, we could fully support our first hypothesis. The data showed a significant increase in appearance evaluation and a significant decrease in body shape concerns and drive for muscularity in the experimental group after intervention.

The positive effects observed in the experimental group may be explained by the combined and mutually reinforcing mechanisms of the digital detox and mindfulness practice. Refraining from Instagram use likely decreases exposure to idealised body standards and upward social comparisons, which are known to contribute to body dissatisfaction (Festinger, 1954; Thompson et al., 1999). In parallel, mindfulness exercises may enhance present-moment awareness and a non-judgemental attitude toward one's thoughts and bodily experiences, thereby reducing critical self-evaluations and automatic comparison processes. Previous research suggests that mindfulness buffers the negative effects of idealised social media images (Hooper et al., 2024), potentially by reducing thought suppression and fostering higher body satisfaction (Barrington & Jarry, 2019).

Importantly, mindfulness may strengthen the effects of the digital detox by reducing the internalisation of appearance-related social media ideals and increasing individuals' ability to notice and disengage from comparison-based thoughts when they arise. Meta-analytic evidence by Bonfanti et al. (2025) supports this mechanism by demonstrating a robust association between social media comparison and body image concerns. While the detox reduces external sociocultural pressure, mindfulness addresses internal cognitive and emotional responses to these pressures. Together, these processes may create a more favorable cognitive environment in which reduced comparison exposure is accompanied by increased self-acceptance promoted by non-judgment, ultimately leading to improvements in body image-related outcomes such as lower body shape concerns and drive for muscularity.

A study from 2024 (Hooper et al., 2024) showed that a 10 min. mindfulness exercise in combination with limiting exposure to appearance-focused social media content helped buffer harmful social media effects. As Miles et al. (2023) showed, mindfulness can address body image concerns even in non-formal interventions and, conversely, a healthier body image can support the practice of mindfulness. Barrington and Jarry (2019) showed that mindfulness can be a protective factor which can be integrated into prevention intervention programs for body dissatisfaction. Mindfulness, therefore, moderates the link between lower body comparison and higher body satisfaction. The relationship between reduced body comparison and increased body satisfaction is stronger at higher levels of mindfulness.

The data connects to Festinger's social comparison theory, which suggests that people evaluate their abilities and attitudes by comparing themselves to others in a process that plays a significant role in self-image and subjective well-being (Festinger, 1954). Instagram promotes upward social comparisons with idealised bodies. A detox reduces exposure to these unrealistic comparison standards forcing users to compare themselves to real bodies within their environment, resulting in a more positive self-evaluation. The Tripartite Influence Model (Thompson et al., 1999) proposes that the three main sources of influence, parents, peers and media, contribute to the development of body image problems and eating disorders. Additionally, the model suggests that the appearance of comparison and internalisation of media information mediates the relationship between influences and problems.
While an Instagram Detox reduces media influence, Mindfulness may reduce the internalisation of media information leading to less body dissatisfaction and drive for muscularity.

Our second hypothesis was only partially supported. A significant difference between the experimental and control group post- intervention could be observed for BSQ but not for AE or DM.

A possible explanation is that drive for muscularity and appearance evaluation may be associated with relatively stable personality traits, such as conscientiousness or perfectionism (Rice & Aldea, 2006; Roberts & DelVecchio, 2000). Previous research has shown that drive for muscularity is linked to stable variables including poor self-esteem and higher levels of depression specifically among boys, suggesting a degree of stability over time (McCreary & Sasse, 2000). Research also found that especially body dissatisfaction was remarkably stable across the adult life span until old age (Tiggemann, 2004). As such, appearance evaluation may be less sensitive to short-term interventions and situational influences. A longer intervention could permit the between-group differences of the control and experimental group to become significant.

Furthermore, drive for muscularity (DM) may be associated with socially accepted and desirable beliefs that link muscularity to health and fitness rather than only seeing it as an aesthetic ideal. Muscularity is often perceived as a marker of physical strength and genes that increase “offspring viability or reproductive success” (Frederick & Haselton, 2007). In line with this, some participants may strive for muscularity primarily for health-related reasons, such as maintaining physical fitness, preventing pain or injury, and preserving physical functioning in the long run, rather than for appearance-based motives alone. In this case, it is possible that DM is mostly targeted by mindfulness exercises, less though by digital detox.

Moreover, reducing only Instagram use as a form of digital detox may not be enough to significantly change AE compared to the control group, as participants remain exposed to appearance ideals through multiple other sources, including other visual social media platforms. According to Social Comparison Theory, individuals evaluate themselves through comparison with others independently of a single source of exposure (Festinger, 1954). The social comparison does not decrease solely because one medium of exposure is reduced. Instead, we can assume participants redirected their social comparison to other sources. The study did not have any influence on what sources participants used instead. Meta-analytic evidence suggests that platform modality plays a role in body satisfaction among young adult women, as video-based platforms tend to amplify the negative effects of body-ideal content. However, the difference between video- and image-based formats was not found to be significant for body-positive content (Zhang et al., 2025). If similar social media platforms, like TikTok were used instead of Instagram, or participants continued to compare themselves to idealised images in films, commercials, or magazines, the overall impact of the intervention on AE may have been limited.

Another important consideration is that the two scales already show small differences between experimental and control group at baseline. Especially for drive for muscularity, where the experimental group initially shows a stronger drive to be muscular than the control group and a lower drive after the one-week intervention, this might have contributed to the found insignificant between- group differences.

Our third hypothesis was fully supported, as there was no significant change in the control group in neither AE, BSQ nor DM.

The stable body image scores in the control group further support reported benefits of a digital detox and mindfulness exercises for one´s body image, because they show that the observed improvements do not occur without targeted intervention (Smith et al., 2024).

The Tripartite Influence Model (Thompson et al., 1999) emphasises that body image changes typically arise from sociocultural pressures like media exposure or internalisation processes. Given that the control group did not alter their Instagram use nor participate in mindfulness exercises, no changes in these influential factors were anticipated. Consequently, no changes in body image perceptions were reported.

Stable body image in the control group was especially important because it confirmed that the observed changes in the experimental group were not likely due to external factors, but more so to the changes in participants´ Instagram use and mindfulness practice.

Current findings indicate that reducing social media use may provide an effective approach to prevent low self-esteem and body image distress. Even a one-week break from social media could provide mental health benefits for young people. Overall, prevention and intervention programmes targeting body image distress should consider strategies to limit social media use to reduce its negative impact on self-esteem and body perception.

One notable limitation in the study is that participants were categorised only as female or male. This binary classification excludes gender-diverse participants, potentially reducing inclusivity of the sample and limiting the generalisability of the findings.

As digital detox and mindfulness are widely known concepts, it is likely that study participants had preconceived ideas about their expected effects. These notions may have influenced participants' responses, leading them to overestimate improvements in body image and potentially distorting the actual results. Future studies should address this issue by assessing participants' familiarity with these interventions or by introducing measures to control for expectancy bias.

In the study conducted, the effects of digital detox and mindfulness exercises were not examined separately. This complicates determining which of the influences explains change in body image. For future research, we recommend adding two more groups, one solely influenced by mindfulness exercises and one by digital detox.

It is questionable whether the one-week intervention provided sufficient time for participants to fully adapt to reduced digital exposure and for mindfulness-related effects to develop into more lasting changes in body image-related thought processes.

Furthermore, even during the planned digital detox period, some participants may have continued to use Instagram or compensated for reduced use by engaging with other visual media platforms not included in the study. This may have limited the intended impact of the intervention. For future research, implementing more comprehensive monitoring of participants' social media use or more systematically restricting access to other platforms may be advisable. One possible way to address this limitation would be to request screenshots of participants' weekly screen time summaries, which could provide a more accurate estimate of overall social media use across platforms.

Another limitation is that Levene´s test came out to be significant for BSQ at posttest, implying that the assumption of homogeneity of variance may be violated. As the Levene´s test is very sensitive to small sample sizes, a bigger sample size may resolve this issue.

Additionally, the data did not perfectly meet the assumption of normality. Although the mixed ANOVA is generally considered a robust method and can tolerate moderate violations of this assumption, the deviation from normality may still have influenced the precision of the results.

Furthermore, this study focuses on short-term effects only, leaving the long-term effects of engaging in Instagram detox and meditation exercises unexplored. Follow-up tests may permit a more comprehensive insight into the persistence and evolution of the effects of mindfulness exercises and digital detox on body image over time.

Conclusion

The current study investigated the effects of a one-week Instagram digital detox combined with mindfulness exercises on body image. The findings revealed higher scores in Appearance Evaluation (AE) and lower scores in Drive for Muscularity (DM) and the Body Shape Questionnaire (BSQ) among participants in the experimental condition at posttest. The between-group differences were only significant for BSQ but not for AE or DM. No significant change could be observed in AE, BSQ or DM in the control group at posttest.

These findings provide evidence of a causal relationship between Instagram use combined with mindfulness exercises and Body Image. They highlight the potential short-term benefits of reducing or eliminating social media use and engaging in mindfulness practices for aspects of body image.

References

American Psychological Association (2023, Februar 23). Reducing social media use significantly improves body image in teens, young adults. https://www.apa.org/news/press/releases/2023/02/social-media-body-image

Baker, Z. G., Krieger, H., & LeRoy, A. S. (2016). Fear of missing out: Relationships with depression, mindfulness, and physical symptoms. Translational Issues in Psychological Science, 2(3), 275–282. https://doi.org/10.1037/tps0000075

Barrington, J., & Jarry, J. L. (2019). Does thought suppression mediate the association between mindfulness and body satisfaction? Mindfulness, 10(4), 679–688. https://doi.org/10.1007/s12671-018-1012-2

Bonfanti, R. C., Melchiori, F., Teti, A., Albano, G., Raffard, S.Rodgers,

Cash, T. F. (1990). *The Multidimensional Body-Self Relations Questionnaire*. Unpublished manuscript, Old Dominion University, Norfolk, VA, USA.

Castellanos Silva, R., & Steins, G. (2023). Social media and body dissatisfaction in young adults: An experimental investigation of the effects of different image content and influencing constructs. Frontiers in Psychology, 14. https://doi.org/10.3389/fpsyg.2023.1037932

Charoensukmongkol, P. (2016). Mindful Facebooking: The moderating role of mindfulness on the relationship between social media use intensity at work and burnout. *Journal of Health Psychology, 21(9),* 1966–1980. https://doi.org/10.1177/1359105315569096

Cooper, P. J., Taylor, M. J., Cooper, Z., & Fairburn, C. G. (1987). The development and validation of the Body Shape Questionnaire. International Journal of Eating Disorders, *6*(4), 485–494. https://doi.org/10.1002/1098-108X(198707)6:4<485::AID-EAT2260060405>3.0.CO;2-O

Dane, A., & Bhatia, K. (2023). The social media diet: A scoping review to investigate the association between social media, body image and eating disorders amongst young people. PLOS Global Public Health, 3(3), e0001091. https://doi.org/10.1371/journal.pgph.0001091

Dittmar, H., & Howard, S. (2004). Thin-ideal internalization and social comparison tendency as moderators of media models' impact on women's body-focused anxiety. Journal of Social and Clinical Psychology, 23(6), 768–791. https://doi.org/10.1521/jscp.23.6.768.54799

Du, J., Kerkhof, P., & van Koningsbruggen, G. M. (2021). The reciprocal relationships between social media self-control failure, mindfulness and wellbeing: A longitudinal study. PLoS ONE, 16(8), e0255648. https://doi.org/10.1371/journal.pone.0255648

Evans, C. and Dolan, B. (1993), Body shape questionnaire: Derivation of shortened "alternate forms". *International Journal of Eating Disorders, 13*(3), 315–321. https://doi.org/10.1002/1098-108X(199304)13:3<315::AID-EAT2260130310>3.0.CO;2-3

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods,* 39(2), 175–191. https://doi.org/10.3758/BF03193146

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7,* 117–140. https://doi.org/10.1177/001872675400700202

Frederick, D. A., & Haselton, M. G. (2007). Why is muscularity sexy? Tests of the fitness indicator hypothesis. *Personality and Social Psychology Bulletin, 33(8),* 1167–1183. https://doi.org/10.1177/0146167207303022

Fredrickson, B. L., & Roberts, T.-A. (1997). Objectification theory: Toward understanding women's lived experiences and mental health risks. *Psychology of Women Quarterly, 21(2)*, 173-206.

Gao, Z., Zhao, J., Liu, X., Wang, H., Zhang, M., & Yuan, H. (2025). Psychometric evaluation of the Chinese Body Shape Questionnaire short form for assessing eating disorder symptoms among university students. *Frontiers in Public Health, 13,* Article 1571144. https://doi.org/10.3389/fpubh.2025.1571144

Hogue, J. V., & Mills, J. S. (2019). The effects of active social media engagement with peers on body image in young women. Body Image, 28, 1–5. https://doi.org/10.1016/j.bodyim.2018.11.002

Holland, G., & Tiggemann, M. (2016). A systematic review of the impact of the use of social networking sites on body image and disordered eating outcomes. *Body Image*, *17*, 100–110. https://doi.org/10.1016/j.bodyim.2016.02.008

Hooper, R., Guest, E., Ramsey-Wade, C., & Slater, A. (2024). A brief mindfulness meditation can ameliorate the effects of exposure to idealised social media images on self-esteem, mood, and body appreciation in young women: An online randomised controlled experiment. *Body Image*, *49*, 101702.

https://doi.org/10.1016/j.bodyim.2024.101702

Keles, B., McCrae, N., & Grealish, A. (2020). A systematic review: The influence of social media on depression, anxiety and psychological distress in adolescents. *International Journal of Adolescence and Youth*, *25*(1), 79–93. https://doi.org/10.1080/02673843.2019.1590851

Kircaburun, K., Griffiths, M. D., & Billieux, J. (2019). Trait emotional intelligence and problematic online behaviors among adolescents: The mediating role of mindfulness, rumination, and depression. *Personality and Individual Differences, 139,* 208–213. https://doi.org/10.1016/j.paid.2018.11.024

Kuss, D. J., & Griffiths, M. D. (2017). Social networking sites and addiction: Ten lessons learned. International Journal of Environmental Research *and Public Health,* 14(3), 311. https://doi.org/10.3390/ijerph14030311

Markey, C. H., & Daniels, E. A. (2022). An examination of preadolescent girls' social media use and body image: Type of engagement may matter most. *Body Image*, *42*, 145–149. *https://doi.org/10.1016/j.bodyim.2022.05.005*

McCreary, D. R., & Sasse, D. K. (2000). An exploration of the drive for muscularity in adolescent boys and girls. *Journal of American College Health*, 48(6), 297–304. https://doi.org/10.1080/07448480009596271

Mestdagh, M., Verdonck, S., Piot, M., Niemeijer, K., Tuerlinckx, F., Kuppens, P., & Dejonckheere, E. (2023). M-path: An easy-to-use and highly tailorable platform for ecological momentary assessment and intervention in behavioral research and clinical practice. *Frontiers in Digital Health*, *5*, 1182175. https://doi.org/10.3389/fdgth.2023.1182175

Miles, E., Matcham, F., Strauss, C., & Cavanagh, K. (2023). Making Mindfulness Meditation a Healthy Habit. *Mindfulness*, *14*(12), 2988–3005. https://doi.org/10.1007/s12671-023-02258-6

Neelakandan, S. (2024). Internet Addiction, Digital Detox, and Mental Health: A Review of the Evidence with Clinical Examples. *SBV Journal of Basic, Clinical and Applied Health Science*, *7*, 140–143. https://doi.org/10.4103/SBVJ.SBVJ_44_24

R., & Lo Coco, G. (2025). The association between social comparison in social media, body image concerns and eating disorder symptoms: A systematic review and meta-analysis. *Body Image, 52, 101841.* https://doi.org/10.1016/j.bodyim.2024.101841

Rice, K. G., & Aldea, M. A. (2006). State dependence and trait stability of perfectionism: A short-term longitudinal study. *Journal of Counseling Psychology, 53(2),* 205–213. https://doi.org/10.1037/0022-0167.53.2.205

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126(1),* 3–25. https://doi.org/10.1037/0033-2909.126.1.3

Saiphoo, A. N., & Vahedi, Z. (2019). A meta-analytic review of the relationship between social media use and body image disturbance. C*omputers in*

Human Behavior, 101, 259–275. https://doi.org/10.1016/j.chb.2019.07.028

Schmuck, D., Karsay, K., Matthes, J., & Stevic, A. (2019). Looking up and feeling down: The influence of mobile social networking site use on upward social comparison, self-esteem, and well-being of adult smartphone users. *Telematics and Informatics, 42,* Article 101240. https://doi.org/10.1016/j.tele.2019.101240

Sherlock, M., & Wagstaff, D. L.(2019). Exploring the relationship between frequency of Instagram use, exposure to idealized images, and psychological well-being in women. *Psychology of Popular Media Culture*, *8*(4), 482–490. https://doi.org/10.1037/ppm0000182

Smith, O. E., Mills, J. S., & Samson, L. (2024). Out of the loop: Taking a one-week break from social media leads to better self-esteem and body image among young women. *Body Image, 49*, 101715. https://doi.org/10.1016/j.bodyim.2024.101715

Stewart, T. M. (2004). Light on body image treatment. *Behavior Modification, 28(6),* 783–811. https://doi.org/10.1177/0145445503259862

Sun, Y., & Zhang, Y. (2021). A review of theories and models applied in studies of social media addiction and implications for future research. *Addictive Behaviors, 114,* 106699. https://doi.org/10.1016/j.addbeh.2020.106699

Thompson, J. K., & Harriger, J. A. (2023). Body image and social media: The fault lines are clear - We need a seismic correction. *Body Image*, *45*, 142–144. https://doi.org/10.1016/j.bodyim.2023.02.009

Thompson, J. K., & Stice, E. (2001). Thin-ideal internalization: Mounting evidence for a new risk factor for body-image disturbance and eating pathology. *Current Directions in Psychological Science*, *10*(5), 181–183. https://doi.org/10.1111/1467-8721.00144

Thompson, J. K., Heinberg, L. J., Altabe, M., & Tantleff-Dunn, S. (1999). The tripartite influence model of body image and eating disturbance: A theoretical and empirical review. In J. K. Thompson, L. J.Heinberg, M. Altabe, & S. Tantleff-Dunn (Eds.), *Exacting beauty: Theory, assessment, and treatment of body image disturbance* (pp. 109–139). American Psychological Association.https://doi.org/10.1037/10312-005

Throuvala, M. A., Griffiths, M. D., Rennoldson, M., & Kuss, D. J. (2020). Mind over matter: Testing the efficacy of an online randomized controlled trial to reduce distraction from smartphone use. *International Journal of Environmental Research and Public Health*

Tiggemann, M. (2004). Body image across the adult life span: Stability and change. *Body Image, 1*(1), 29–41. https://doi.org/10.1016/S1740-1445(03)00002-0

Topic: Social media. (n.d.). Statista. Retrieved 22 December 2025, from https://www.statista.com/topics/1164/social-networks/?srsltid=AfmBOor6527XOWkg5k0glNERepaLgJ7e1o0jQ_F4Rv_QmHDnQ5_dSHfa

Vandenbosch, L., Fardouly, J., & Tiggemann, M. (2022). Social media and body image: Recent trends and fu-

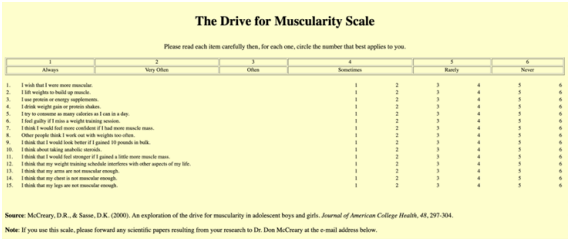
ture directions. *Current Opinion in Psychology, 45*, 101289. https://doi.org/10.1016/j.copsyc.2021.12.002

Zhang, Y., Xiong, S., Yao, N., & Chen, J. (2025). Social media and body image in young women: A meta-analysis of experimental studies on video- versus image-based platforms. *Body Image, 55,* 101991. https://doi.org/10.1016/j.bodyim.2025.101991

Appendix A

1. Scales used

A) Drive for Muscularity Scale (McCreary, D. R., & Sasse, D. K.; 2000)



The Drive for Muscularity Scale

Please read each item carefully then, for each one, circle the number that best applies to you.

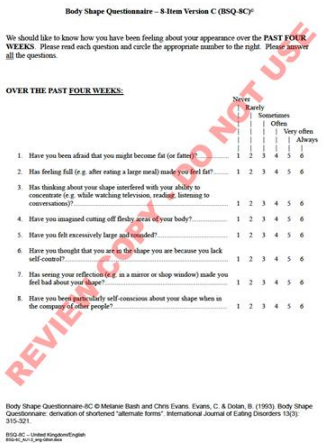
1	2	3	4	5	6
Always	Very Often	Often	Sometimes	Rarely	Never

1.	I wish that I were more muscular.	1	2	3	4	5	6
2.	I lift weights to build up muscle.	1	2	3	4	5	6
3.	I use protein or energy supplements.	1	2	3	4	5	6
4.	I drink weight gain or protein shakes.	1	2	3	4	5	6
5.	I try to consume as many calories as I can in a day.	1	2	3	4	5	6
6.	I feel guilty if I miss a weight training session.	1	2	3	4	5	6
7.	I think I would feel more confident if I had more muscle mass.	1	2	3	4	5	6
8.	Other people think I work out with weights too often.	1	2	3	4	5	6
9.	I think that I would look better if I gained 10 pounds in bulk.	1	2	3	4	5	6
10.	I think about taking anabolic steroids.	1	2	3	4	5	6
11.	I think that I would feel stronger if I gained a little more muscle mass.	1	2	3	4	5	6
12.	I think that my weight training schedule interferes with other aspects of my life.	1	2	3	4	5	6
13.	I think that my arms are not muscular enough.	1	2	3	4	5	6
14.	I think that my chest is not muscular enough.	1	2	3	4	5	6
15.	I think that my legs are not muscular enough.	1	2	3	4	5	6

Source: McCreary, D.R., & Sasse, D.K. (2000). An exploration of the drive for muscularity in adolescent boys and girls. *Journal of American College Health, 48*, 297-304.

Note: If you use this scale, please forward any scientific papers resulting from your research to Dr. Don McCreary at the e-mail address below.

B) BSQ8 (Body Shape Questionnaire) (Evans, C. and Dolan, B.;1993)



Body Shape Questionnaire – 8-Item Version C (BSQ-8C)©

We should like to know how you have been feeling about your appearance over the **PAST FOUR WEEKS**. Please read each question and circle the appropriate number to the right. Please answer all the questions.

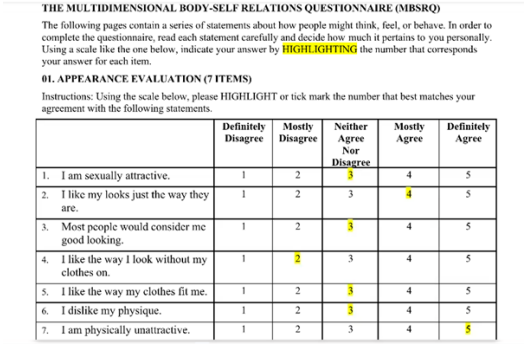
OVER THE PAST FOUR WEEKS:

	Never	Rarely	Sometimes	Often	Very often	Always
1. Have you been afraid that you might become fat (or fatter)?	1	2	3	4	5	6
2. Has feeling full (e.g. after eating a large meal) made you feel fat?	1	2	3	4	5	6
3. Has thinking about your shape interfered with your ability to concentrate (e.g. while watching television, reading, listening to conversations)?	1	2	3	4	5	6
4. Have you imagined cutting off flesh/areas of your body?	1	2	3	4	5	6
5. Have you felt excessively large and rounded?	1	2	3	4	5	6
6. Have you thought that you are the shape you are because you lack self-control?	1	2	3	4	5	6
7. Has seeing your reflection (e.g. in a mirror or shop window) made you feel bad about your shape?	1	2	3	4	5	6
8. Have you been particularly self-conscious about your shape when in the company of other people?	1	2	3	4	5	6

Body Shape Questionnaire-8C © Melanie Bash and Chris Evans. Evans, C. & Dolan, B. (1993). Body Shape Questionnaire: derivation of shortened "alternate forms". International Journal of Eating Disorders 13(3): 315-321.

BSQ-8C – United Kingdom/English

C) MBSRQ - Appearance Evaluation (Cash, T. F.; 1990)



THE MULTIDIMENSIONAL BODY-SELF RELATIONS QUESTIONNAIRE (MBSRQ)

The following pages contain a series of statements about how people might think, feel, or behave. In order to complete the questionnaire, read each statement carefully and decide how much it pertains to you personally. Using a scale like the one below, indicate your answer by HIGHLIGHTING the number that corresponds your answer for each item.

01. APPEARANCE EVALUATION (7 ITEMS)

Instructions: Using the scale below, please HIGHLIGHT or tick mark the number that best matches your agreement with the following statements.

	Definitely Disagree	Mostly Disagree	Neither Agree Nor Disagree	Mostly Agree	Definitely Agree
1. I am sexually attractive.	1	2	3	4	5
2. I like my looks just the way they are.	1	2	3	4	5
3. Most people would consider me good looking.	1	2	3	4	5
4. I like the way I look without my clothes on.	1	2	3	4	5
5. I like the way my clothes fit me.	1	2	3	4	5
6. I dislike my physique.	1	2	3	4	5
7. I am physically unattractive.	1	2	3	4	5

Analogical transfer despite irrelevant information in adults

Tatjana Botelho, Kenza Cossu, Madalena Da Rocha Lameira, Ariane Neuffer und Adelise Vieira Marques

Supervisor: Dr. Katarzyna Bobrowicz

The present study investigates how multilingual and bilingual adults perform on analogi-cal transfer tasks compared to monolinguals, with a focus on situations requiring inhibi-tion of irrelevant information and cognitive flexibility. We hypothesized that multilingual and bilingual participants would show advantages in conditions with high inhibitory de-mand and cognitive flexibility. Forty-four adults (30 females, 13, males, 1 nonbinary) completed three tasks; the A:B::C:D Analogical Task, assessing relational reasoning and analogical transfer, the Flanker Task, measuring inhibitory control; and the Flanker-Switching Task, assessing cognitive flexibility under shifting attention demands. No sig-nificant group differences were found in the A:B::C:D Analogical Task, Flanker Task or the Flanker-Switching Task. Correlation analyses showed a positive association be-tween performance on the A:B::C:D Analogical Task and the Flanker Task, but none between the A:B::C:D Analogical Task and Flanker-Switching Task. Overall, the findings do not support a general bilingual or multilingual cognitive advantage. Instead, they sug-gest that performance is shaped more by the specific cognitive demands of each task than by language background alone.

Keywords: analogical transfer, inhibitory control, cognitive flexibility, multilingualism, bilingualism, monolingualism**.**

Analogical transfer despite irrelevant information in adults

Flexible problem solving often requires people to extract the underlying relation between two concepts and apply it to a new context. This ability is known as analogical transfer and is a central cognitive function used in learning, reasoning, and adaptation. Executive functions, particularly inhibitory control, and cognitive flexibility, play a crucial role in the process of analogical transfer.

Inhibitory control is the ability to suppress irrelevant or conflicting information, whereas cognitive flexibility enables individuals to efficiently shift between rules, tasks, or perspectives. Together, these processes allow individuals to focus on relational similarities while ignoring distracting features. This means that individuals with strong inhibitory control can disregard non-relational details, such as color, size, or shape of the elements, when identifying the patterns between them. Such patterns are, for instance, spatial relations which can be identified when one object is contained within another; functional relations, are evident when one event directly causes another; and role-based analogies emerge when one object follows or pursues another, thereby highlighting the relationship between objects.

Some research suggests that managing multiple languages can enhance executive functions, such as inhibitory control and cognitive flexibility due to constant language selection and the need to suppress non-target languages (Xie et al., 2022). Furthermore, there is evidence suggesting that bilingual children may outperform monolinguals in analogy transfer when relational reasoning competes with distracting perceptual information (Shin, 2015). This suggests that similar cognitive advantages might also be observed in adults.

While some studies indicate such benefits, the overall findings about this research topic are mixed and inconsistent. To highlight this, a critical review of the literature on the bilingual advantage provided by Paap, Johnson, and Sawi (2015) argues that there is little consistent evidence to indicate bilingual advantage on cognitive functions in adults. Recent research also emphasizes that the observed effects tend to be subtle, task-dependent, influenced by age, context, and individual language experience such as individual proficiency and frequency of language use (Brunetti et al., 2019; Grundy et al., 2017; Zhou & Krott, 2020). Similarly, an empirical study by von Bastian, Souza, and Gade (2016) examined cognitive domains, wherein bilingualism was considered a continuous variable showed little evidence for cognitive advantages associated with bilingualism. Some of these studies using among others, the Flanker task, indicate that bilingual individuals tend to only outperform monolinguals under specific conditions. This emerges mainly in tasks where there are distracting stimuli that must be ignored and when the tasks require relational reasoning. (Brunetti et al., 2019; Grundy et al., 2017; Paap & Sawi, 2015; Zhou & Krott, 2020)

Therefore, we decided to investigate this relationship, examining whether multilingual and bilingual adults outperform monolinguals in analogical transfer tasks, especially when tasks require ignoring irrelevant cues and applying relations to different but similar situations.

We hypothesize that multilingual and bilingual adults will outperform monolinguals on analogical transfer tasks, particularly when inhibitory demands are high and in cognitive flexibility. This means that we expect that individuals who score high on inhibition and cognitive flexibility will also score high on analogical transfer.

Method

Participants

The study *Flexible Problem Solving in Adults* was conducted in the Cognition laboratory of the Maison des Sciences Humaines at the University of Luxemburg, Belval, as part of the course *Practical Training in Empirical Research* (BAP-B3-2B).

Forty-four healthy adult volunteers participated. Recruitment occurred via Moodle announcements (October 2nd - November 7th, 2025) and word of mouth within the university community. Testing sessions were held between October 6th and November 7th. Participation was voluntary, and individuals could withdraw from the study at any time without providing a reason. Each participant received a 10€ voucher as compensation at the end of the session.

Inclusion criteria required participants to be at least 18 years old and sufficiently proficient in one of the study languages (English, French, or German). Demographic information (age, gender, and lingual status) was collected.

Participants ranged in age from 18 to 45 years (M = 23.57, SD = 4.07). Most participants identified as female (n = 30, 68,18%), with 13 identifying as male (29,55%) and one as non-binary (2.27%). Linguistic backgrounds were diverse: 23 participants (52.27%) were multilingual, 11 (25%) monolingual, and 10 (22.73%) bilingual.

This study complies with the University of Luxembourg's ethics policy and was approved by the Ethics Review Panel (ERP 25-002 FlexPS Lab Adults).

Material

All tasks were administered on a desktop computer with a standard keyboard and mouse. Stimuli and data collection were implemented using PsychoPy.

Analogical reasoning, executive functions, and demographic characteristics were assessed through computerized cognitive tasks and a standardized questionnaire. Each participant completed a single laboratory session lasting approximately 60 minutes.

Analogical reasoning was measured using an A:B::C:D Analogical Task. Participants selected the D-item from four options

such that the relationship between C and D matched that between A and B. All items were presented simultaneously and continuously available until the participant chose the answer. Two phases were included:

Uncolored Analogical Task: Two practice trials followed by timed test trials (15 s per item). Items included both strongly and weakly associated relational pairs.

Color-Based Analogical Task: Participants judged same/different color relations with colored stimuli. Three practice trials provided corrective feedback; test trials (15 s limit) were completed without feedback.

Inhibitory control was measured using a Flanker Task adapted for shape and line-orientation discrimination. Participants responded to the central stimulus while ignoring surrounding distractions. Two example trials (one shape-based, one line-based) preceded the task. A short practice block with 5 practice trials without time limits was completed, with a conditional rule requiring 75% correct answers to proceed to the main task.

Responses were made via keyboard: “A” for triangle/left-leaning lines and “L” for circle/right-leaning lines. The main task consisted of two blocks (50 trials in total, 2 s response window). Each block included both shapes and line-orientation trials. Short breaks were allowed between blocks.

Set-shifting ability was assessed using Flanker-Switching Task in which participants alternated between paying attention to either the inside or outside part of a composite stimulus. Stimuli included spatial directions (up, down, left, right) and shapes (filled vs. empty). Cues indicated focus: red background = outside, blue background = inside, white background = word cue (“inside” or “outside”). Responses were made by pressing “A” to select the left option and “L” for the right one. The task comprised three blocks, each preceded by 5 practice trials with feedback, during which participants had to achieve a performance threshold of 75% to proceed to the main task. The main task consisted of 50 mixed switching trials in total, without receiving feedback.

Data analyses were conducted using RStudio (Version 2025.09.2+418).

Procedure

All testing took place in the Cognition Lab of the Maison du Savoir at the University of Luxembourg. Each participant completed a single 60-minute session.

Upon arrival, participants received an information sheet describing the study and signed an informed consent form, indicating their preferred language (French, German, or English). They then completed the demographic questionnaire and voucher confirmation form documenting the €10 compensation. Afterwards, participants were seated at a desktop computer and reminded of their right to withdraw at any time. Ethical and administrative materials included the participant information sheet, informed consent form, and voucher confirmation form.

The experimental session consisted of three computerized tasks administered in a pseudorandomized order, with the analogical transfer task being first:

Analogical transfer task

Participants first completed two practice trials of the uncolored A:B::C:D Analogical Task to familiarize themselves with the format. They then proceeded to timed test trials (15-s limit per item), which included both strong and weakly associated relational pairs. Next, participants completed the color-based analogy block. This began with three practice trials of same/different color judgments, which provided corrective feedback. After practice, participants solved test trials under a 15-s limit without feedback.

Flanker Task (inhibitory Control)

Before the main task, participants viewed two example trials (one shape-based, one line-orientation based). They then completed a practice block of 5 trials without time limits, which they responded with the keyboard. The main task consisted of two blocks of 50 trials in total, with a maximum response limit of 5 s. Each block included both shapes and line-orientation trials. Short breaks were allowed between blocks. Short breaks were offered between blocks to minimize fatigue.

Flanker-Switching Task (Set-Shifting)

Participants alternated between attending to either the inside or outside part of a composite stimulus, depending on cues. Each of the three blocks began with a five-trial practice phase that provided feedback, followed by 50 mixed switching trials without feedback. In the main task, each trial began with a cue presented for 2 seconds, followed by a target stimulus. Trials were always presented by a 0.2 second inter-trial interval. There were no time limits defined for this task.

After completing all the tasks, participants were provided with their €10 voucher. All data were pseudonymized, with consent forms, demographic responses, and task performance linked only via randomly generated participant codes. Only authorized research team members had access to the dataset. No follow-up with participants was required.

Statistical analysis

All analyses were conducted in R Studio (version 2025.09.2+418), using base R functions alongside the *dplyr* and *tidyr* packages for data manipulation. The raw dataset, which contained trial-level responses across all cognitive tasks as well as demographic variables, was first cleaned by converting missing values to NA and removing empty strings. A single incorrectly coded gender entry was manually corrected to ensure data integrity, and no additional imputations procedures were applied.

Participant-level overall scores were then calculated separately for each task. For each participant, the number of correct and incorrect responses was tallied, and accuracy was defined as the proportion of correct responses relative to the total number of completed trials. These participant-level variables were generated using grouped summarization functions from dplyr, resulting in three overall scores corresponding to the A:B::C:D Analogical Task, the Flanker Task, and the Flanker-Switching Task. For analyses requiring a wide data format, the dataset was reshaped using the pivot_wider function from the tidyr package (Wickham et al.,2025). Participant language information (monolingual, bilingual, multilingual) was extracted from the raw data and merged into the wide dataset by participant ID; participants with missing language information were excluded from group-level analyses.

Pairwise deletion (na.rm = TRUE) was used to handle missing data, ensuring that the maximum number of available observations contribute to each analysis without artificially altering the scores. Missing data corresponded to missing answers in trials, which were excluded from the analyses. Excluding them did not affect the results, as only fully completed trials were used to calculate each participant's score.

To address the first research question regarding group differences in analogical transfer performance, a contingency table was constructed from counts of correct versus incorrect responses for each language group. A Pearson chi-square test of independence was then conducted to examine the association between language group and A:B::C:D Analogical Task scores. All expected cell frequencies met the minimum threshold for the chi-square approximation, although R issued a warning for small sample sizes in some cells. A:B::C:D Analogical Task scores were treated as a categorical outcome variable for this analysis, consistent with the structure of the cumulative dataset.

The second and third research questions examined whether analogical transfer scores were associated with inhibitory control (Flanker Task) and cognitive flexibility (Flanker-Switching Task). Because Shapiro–Wilk tests indicated that accuracy distributions for all tasks significantly deviated from normality, Spearman rank correlations were used for all correlational analyses. Correlations were computed both across the entire sample and separately within monolingual, bilingual, and multilingual subgroups to assess whether task relationships differed by language background.

Because each research question involved four planned correlation tests (one overall, three subgroup tests), Bonferroni corrections were applied separately for each research question, resulting in an adjusted alpha level of $\alpha=0.0125$.

The analyses were fully reproducible, because each processing step was scripted using R, with one single manual demographic correction. The code is available in Appendix A.

Results

Accuracy scores for the A:B::C:D Analogical Task, Flanker Task, and Flanker-Switching Task served as outcome measures, and were calculated as each participant's overall scores.

Language group (monolingual, bilingual, multilingual) served as the independent variable for the group comparisons, the scores are treated as continuous measures in the correlation analyses. Despite being recorded, age and gender, were not included as covariates in the statistical models due to the complexity of the analyses and the focus on primary variables of interest such as language background and task performance measures.

For the first research question, which examined group differences in overall scores, a contingency table was constructed to assess correct and incorrect responses across the three language groups. Monolingual participants produced three correct and eight incorrect responses, bilingual participants produced two correct and seven incorrect responses, and multilingual participants produced eight correct and sixteen incorrect responses. A chi-squared test revealed no significant association between the groups, χ^2 (2) = 0.42, p = .81. This indicates that analogical transfer scores did not differ significantly across language backgrounds.

For the second research question, which investigated the relationship between analogical transfer and inhibitory control, the Shapiro-Wilk tests indicated that the data were not normally distributed. Therefore, Spearman's rank correlation was used. Across the entire sample, there was a significant positive correlation between performance on the analogical transfer task and the inhibitory control task, with ρ = .31 and p = .003, which remained significant after Bonferroni correction (adjusted p value = .012).

When examined by language group, monolingual participants showed a strong and significant correlation, ρ = .79, p = .004, adjusted p value = .016. Bilingual participants showed a moderate but non-significant correlation, ρ = .41, p = .27, p_adj = 1.00 and multilingual participants showed no correlation at all, ρ = –.01, p = .97, p_adj = 1.00. (see Figure 1)

For the third research question, which examined the relationship between analogical transfer and cognitive flexibility, the overall correlation was not significant, ρ = .05, p = .63. Bonferroni correction was again applied across the four correlations tested. When analyzed by language group, monolingual participants showed a non-significant correlation of ρ = .35, p = .30, p_adj = 1.00, bilingual participants showed a non-significant correlation of ρ = .15, p = .70, p_adj = 1.00 and multilingual participants showed a non-significant correlation of ρ = –.17, p = .44, p_adj = 1.00 These results suggest that analogical transfer was not associated with cognitive flexibility in this sample. (see Figure 2)

Discussion

This study examined whether language background (monolingual, bilingual, multilingual) is associated with performance in a cognitive task and whether overall scores across the three selected tasks show a positive relation. The analyses showed that analogical transfer scores did not differ between language groups. The chi-squared test revealed no significant differences, indicating that monolingual, bilingual, and multilingual participants performed similarly on this task. This finding is consistent with research, questioning the existence of a general bilingual cognitive advantage. Previous studies have reported mixed results, with some suggesting advantages for bilinguals in certain executive processes (Bialystok, Craik, & Luk, 2012), while others have found no reliable differences between bilinguals and monolinguals. Paap and Greenberg (2013) empha-

size that many reported bilingual advantages fail to replicate and that the evidence base contains numerous null findings.

The second research question investigated the relationship between analogical transfer and inhibitory control. When examined by language group, monolingual participants showed a strong and significant correlation, bilingual participants showed a moderate but non-significant correlation, and multilingual participants showed no correlation. This pattern indicates that the relationship between analogical transfer and inhibitory control may be more pronounced in monolinguals, but less consistent in bilinguals and multilinguals.

The third research question examined the relationship between analogical transfer and cognitive flexibility. No significant correlations were found, either overall or within language groups. This suggests that the cognitive flexibility task is based on different cognitive abilities than the analogical transfer one, supporting the idea that executive-function tasks measure partly independent skills rather than reflecting a single transferable ability.

Several alternative explanations may account for the observed results. Differences in the cognitive demands of the tasks may influence the extent to which they correlate with one another. Individual differences such as socioeconomic background, educational experiences, or language use patterns may also contribute to variability in performance within language groups. Additionally, variation in motivation, attention, or fatigue across tasks could influence overall scores and obscure underlying relationships, which is however unlikely due to the order of the Flanker tasks being pseudorandomized.

Several limitations must be considered when interpreting these findings. The sample size was modest, particularly within language subgroups, which reduces statistical power and may explain why some subgroup correlations did not reach high significance. Due to the uneven distribution of participants per language group, the stability of the subgroup analyses was limited further. The study also lacked detailed measures of language experience, such as age of acquisition, proficiency, and frequency of language switching, which limits the ability to draw nuanced conclusions about bilingual and multilingual experiences. In addition, by solely capturing the overall scores we run the risk of overseeing subtle performance differences that response times would reveal. Finally, using pairwise deletion for the missing data might have caused slight differences in sample sizes across our analysis, ultimately resulting in minor inconsistencies in comparisons.

Despite these limitations, the study has several strengths. The use of three distinct cognitive tasks allowed us to examine cross-task relationships rather than reliance on a single measure. By fully scripting the statistical analysis procedure, we were able to ensure transparency and reproducibility. Importantly, the inclusion of multilingual participants broadens the evidence base, as multilingual samples remain relatively underrepresented in literature.

These findings contribute to the discussion about whether bilingualism or multilingualism affects executive functioning. The lack of group differences suggests that just knowing multiple languages does not automatically improve performance, but task-specific cognitive demands may significantly contribute to said performance.

Future research should employ larger and more balanced samples to increase statistical power and stability of subgroup analyses. Measuring language use in detail, including proficiency, frequency of language switching, and contexts of language use, would allow for more precise examination of how language experience influences cognition. By measuring response times, the subtle performance difference could be detected. Researchers might also include neurocognitive instruments such as EEG or eye-trackers to further capture underlying performance during tasks. EEG could reveal the timing of detection and suppression of distracting elements, and eye-tracking could measure eye movement and how attention shifts between elements, while also revealing which strategies participants use to solve an analogical problem. Longitudinal studies could further explain how language experience influences cognitive development over time.

Figure 1: *A:B::C:D- & Flanker overall Score*

This Scatterplot figure illustrates the relationship between the overall scores of the analogical transfer task (ABCD) and the inhibitory control task (Flanker). Each point represents the aggregated score of one language group. The solid lines indicate the best-fit regression line for each language group.

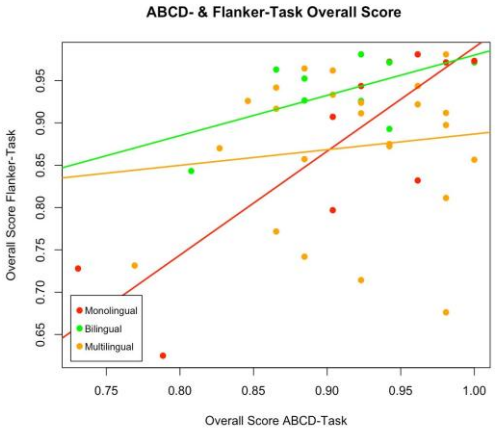
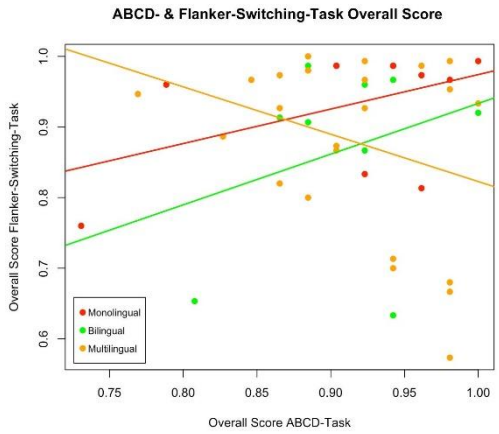


Figure 2: *A:B::C:D- & Flanker-Switching-Task Overall Score*

This Scatterplot figure illustrates the relationship between the overall scores of the analogical transfer task (ABCD) and the cognitive flexibility task (Flanker-Switching). Each point represents the aggregated score of one language group. The solid lines indicate the best-fit regression line for each language



group.

Appendix A: Questionnaire and consent form used for Participant data collection

Appendices 3-5 FINAL POST DPO (1).pdf

Appendix B: Declaration of AI use

Parts of the text in this report were reformulated and rewritten with the assistance of AI tools (e.g. ChatGPT (GPT-5.2), Microsoft Copilot (GPT-5.1)). All content has been reviewed, adapted, and verified by the authors to ensure accuracy and clarity.

References

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. Trends in Cognitive Sciences, 16(4), 240–250.

Brunetti, M., Zappasodi, F., Croce, P. & Di Matteo, R. (2019). Parsing the Flanker task to reveal behavioral and oscillatory correlates of unattended conflict interference. *Scientific Reports*, *9*(1), 13883. https://doi.org/10.1038/s41598-019-50464-x

Grundy, J. G., Anderson, J. A. & Bialystok, E. (2017). Neural correlates of cognitive processing in monolinguals and bilinguals. *Annals Of The New York Academy Of Sciences*, *1396*(1), 183–201. https://doi.org/10.1111/nyas.13333

Paap, K. R. & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*(2), 232–258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Paap, K. R., Johnson, H. A. & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265–278. https://doi.org/10.1016/j.cortex.2015.04.014

Shin, S. Y. (2015). *Is two better than one? Analogical reasoning in bilingual and monolingual children*. Works.

Wickham, H., Vaughan, D., & Girlich, M. (2025). *tidyr: Tidy messy data* (R package version 1.3.1) [Computer software]. https://tidyr.tidyverse.org

Xie, W., Altarriba, J. & Ng, B. C. (2022). Bilingualism, Culture, and Executive Functions: Is There a Relationship? *Languages*, *7*(4), 247. https://doi.org/10.3390/languages7040247

Zhou, B. & Krott, A. (2016). Bilingualism enhances attentional control in non-verbal conflict tasks – evidence from ex-Gaussian analyses. *Bilingualism Language And Cognition*, *21*(1), 162–180. https://doi.org/10.1017/s1366728916000869

La maturité vocationnelle et l'indécision vocationnelle chez les étudiants en psychologie

Maida Buljubasic, Thomas Flamand, Amar Zecirovic

Titulaire: Prof. Dr. Claude Houssemand, Doctorante Clémentine Robert

L'indécision vocationnelle, définie comme l'incapacité de choix scolaire ou professionnel, constitue un phénomène complexe influencé par des facteurs personnels et contextuels. Cette étude s'appuie sur le modèle du Career Decision-Making Difficulties Questionnaire (CDDQ) et sur la conception fonctionnelle de la maturité vocationnelle, envisagée comme une compétence regroupant identité, adaptabilité et résilience. La question de recherche portait sur la relation entre maturité vocationnelle et indécision vocationnelle chez des étudiants en psychologie, en comparant ceux inscrits en première année de Bachelor et en Master à l'Université du Luxembourg. Cinq hypothèses ont été formulées, notamment l'absence de différence entre les groupes et l'association négative entre maturité et indécision. L'échantillon comprenait 35 étudiants, ayant complété deux questionnaires standardisés en ligne. Les analyses montrent qu'il n'existe pas de différence significative de maturité vocationnelle entre Bachelor et Master. Contrairement aux attentes, l'identité est positivement corrélée à l'indécision, suggérant une phase d'exploration active. L'âge est négativement associé à l'indécision, mais non à la maturité vocationnelle. Une reprise d'études est liée à une identité plus affirmée sans réduction de l'indécision. Enfin, la régression indique que la maturité vocationnelle prédit significativement une baisse de l'indécision. Ces résultats soulignent la complexité des liens entre exploration identitaire et prise de décision, et invitent à développer des interventions favorisant la résilience et l'adaptabilité pour accompagner les étudiants dans un contexte professionnel incertain.

1. Introduction

De nos jours, la problématique suivante devient de plus en plus récurrente : un grand nombre d'élèves en fin d'études supérieures demeurent indécis dans le choix de leur carrière professionnelle, un phénomène largement documenté dans la littérature récente (Mata-Correas et al., 2025).

Ce phénomène intitulé l'indécision vocationnelle a été défini comme une absence de choix scolaire ou professionnel (Dosnon, 1996). L'indécision vocationnelle ne se réduit pas à un simple retard de choix, mais est un phénomène multidimensionnel et complexe. Gati (1996) rappelle que la décision de carrière est l'une des tâches les plus exigeantes auxquelles un individu doit faire face, car elle implique un grand nombre d'options, des informations souvent incertaines et des tensions entre aspirations personnelles et contraintes externes. Dans cette perspective, Dosnon (1996) propose de distinguer différentes formes d'indécision ainsi que leurs sources. Ces sources ont été systématisées dans le *Career Decision-Making Difficulties Questionnaire* (CDDQ) (Gati et al., 1996; Rossier et al., 2022), qui distingue trois catégories de difficultés : le manque de préparation, le manque d'information et les informations inconsistantes. L'indécision vocationnelle peut ainsi être comprise comme le reflet d'un déficit de ressources personnelles ou contextuelles, et non seulement comme un manque de « maturité » lié à l'âge (Forner & Dosnon, 1991; Huteau, 2007). Elle est par ailleurs associée à des variables psychologiques négatives telles que la faible estime de soi, l'anxiété et le

manque d'efficacité perçue (Forner & Dosnon, 1991).

Auparavant, la maturité vocationnelle était surtout comprise de manière normative, c'est-à-dire liée à l'âge. On supposait qu'elle augmentait automatiquement avec le développement biologique. Cette conception a été critiquée, car elle ne tenait pas compte des différences individuelles (Forner & Dosnon, 1991). Aujourd'hui, on la définit plutôt de façon fonctionnelle, comme la capacité à accomplir des tâches et à s'adapter aux exigences de la formation et du travail. Dans cette approche fonctionnelle, Ratschinski (2014) propose de remplacer la notion classique de maturité vocationnelle par celle de « *Berufswahlkompetenz* », une compétence fondée sur trois méta-compétences : l'identité professionnelle, l'adaptabilité et la résilience. Cette conception met l'accent sur des ressources évolutives permettant aux individus de s'orienter et de faire face aux exigences d'un environnement professionnel incertain, plutôt que sur un développement lié à l'âge.

Des études appuient cette définition. Par exemple, Crites (1969, cité dans Dosnon, 1996) a montré qu'environ un quart des adolescents restaient indécis face à leur choix professionnel. Il a trouvé ce résultat en comparant différents groupes de jeunes et en observant la fréquence de l'indécision. Cela montre que la maturité vocationnelle ne dépend pas seulement de l'âge, mais aussi des ressources personnelles et des capacités d'adaptation (Forner & Dosnon, 1991)

Le choix de carrière professionnelle est également de plus en plus marqué par le rythme effréné de la vie moderne. Le marché du travail est en constante évolution, les grands secteurs économiques se transforment rapidement et les nouvelles technologies, telles que l'intelligence artificielle, modifient profondément les profils professionnels. Dans cet environnement dynamique, souvent décrit comme un contexte VUCA (Volatility, Uncertainty, Complexity, Ambiguity), c'est-à-dire caractérisé par la volatilité, l'incertitude, la complexité et l'ambiguïté, il devient de plus en plus difficile pour les jeunes de prendre des décisions à long terme (Guo, 2025).

2. Question de recherche

Cette problématique nous a paru particulièrement pertinente dans le contexte des études en psychologie. La maturité vocationnelle et l'indécision vocationnelle sont des sujets particulièrement intéressants, car ils permettent de saisir comment les étudiants gèrent l'incertitude et construisent leur futur professionnel. Cela nous a conduits à formuler la question de recherche suivante : « En quoi la relation entre maturité vocationnelle et indécision vocationnelle diffère-t-elle entre les étudiants de Master en psychologie et les étudiants en première année de Bachelor à l'Université du Luxembourg ? »

3. Hypothèses

Afin de répondre à cette question de recherche, nous souhaitons examiner les cinq hypothèses suivantes.

H1= Les étudiants de Master en psychologie et les étudiants en première année de Bachelor n'ont pas de différence significative quant à la maturité.

H2 = Les étudiants disposant d'une identité plus affirmée sont également ceux qui rapportent une indécision vocationnelle moindre

H3 = Plus les étudiants montrent une identité développée, une bonne adaptabilité et une forte résilience, moins ils rencontrent de difficultés dans leur prise de décision de carrière.

H4= Les étudiants les plus âgés présentent un niveau de maturité vocationnelle plus élevé et une meilleure capacité de prise de décision.

H5= Les étudiants ayant repris leurs études sont plus matures sur le plan vocationnel et présentent un niveau d'indécision vocation-

nelle plus faible que les étudiants n'ayant jamais interrompu leur parcours académique.

4. Méthodes

4.1. Participants et recrutement

L'échantillon final se compose de 35 étudiants en psychologie (*N* = 35). L'âge moyen des participants est de *M* = 22.90 ans (ET = 6.60), avec un âge minimum de 18 ans et un âge maximum de 52 ans. Concernant le genre, 80.0 % des participants s'identifient comme femmes (*n* = 28), 14.3 % comme hommes (*n* = 5) et 5.7 % comme autres (*n* = 2). Les études de psychologie sont plutôt dominées par une population féminine, ce qui explique le déséquilibre de genre. Les participants sont inscrits en Bachelor (*n* = 19 ; 54.3 %) ou en Master (*n* = 16 ; 45.7%). De plus, 22.9 % (*n* = 8) des participants ont déclaré que leurs études en psychologie constituaient une reprise d'études, tandis que 77.1 % (*n* = 27) n'ont pas interrompu leur parcours académique.

Tableau 1 :
Caractéristiques sociodémographiques de l'échantillon (N = 35)

	n	%	*M*	ET	Min-Max
Âge (années)	-	-	22.90	6.60	18-52
Genre					
Femmes	28	80.0	-	-	-
Hommes	5	14.3	-	-	-
Autre	2	5.7	-	-	-
Cursus					
Bachelor	19	54.3	-	-	-
Master	16	45.7	-	-	-
Reprise d'études					
Oui	8	22.9	-	-	-
Non	27	77.1	-	-	-

Note. ET = écart-type.

L'échantillon se composait de deux groupes. Le premier groupe était composé d'étudiants de Bachelor du premier semestre, représentant le début du parcours universitaire. Le second groupe d'étudiants est inscrit en Master, en première ou deuxième année, correspondant à une phase plus avancée du parcours académique. Cette distinction a été établie afin de permettre une analyse comparative de la maturité vocationnelle et de l'indécision vocationnelle à différentes étapes du cursus universitaire.
Les participants ont été recrutés lors de cours de psychologie à l'Université du Luxembourg, sur le campus de Belval. Le recrutement des étudiants de Bachelor (première année) et de Master (première année) a eu lieu pendant des cours de M. Houssemand, superviseur de l'étude présente, ayant transmis le lien du questionnaire aux étudiants concernés.
Les étudiants de deuxième année de master ont été contactés par courrier électronique, lorsqu'ils ne suivaient pas de cours sur le campus au moment du recrutement. Pour pouvoir participer, les étudiants devaient être inscrits en psychologie, soit en Bachelor (première année), soit en Master et avoir une maîtrise suffisamment élevée de la langue française. La participation était volontaire et non rémunérée.

4.2 Instruments de mesure

Afin d'examiner la relation entre la maturité vocationnelle et l'indécision vocationnelle chez les étudiants en psychologie, une enquête en ligne a été menée à l'aide de deux questionnaires standardisés.

4.2.1. Berufswahlkompetenz (Ratschinski, 2018)

Le questionnaire Berufswahlkompetenz (Ratschinski, 2018) permet d'évaluer la maturité vocationnelle. C'est la capacité d'une personne à mobiliser ses ressources personnelles et son environnement pour faire des choix de carrière éclairés et adaptés. À l'origine, l'instrument a été en allemand, dans notre étude, une version française adaptée a été utilisée.

- Identité (19 items):

Cette dimension reflète la clarté de la représentation de soi, la confiance dans ses choix professionnels et la cohérence du projet de vie.

- Adaptabilité (20 items): Elle mesure la capacité à s'ajuster aux changements, à gérer l'incertitude et à envisager différentes alternatives professionnelles.

-Résilience (15 items): Celle-ci évalue la capacité à faire face aux difficultés, à garder la motivation et à persévérer malgré les obstacles rencontrés lors du choix de carrière.

Les items sont formulés sous forme d'affirmations sur une échelle de Likert en quatre points, allant de 1 (pas du tout d'accord) à 4 (tout à fait d'accord). La dimension Identité comprend la sécurité décisionnelle, le réalisme du choix professionnel, l'activité personnelle ainsi que l'engagement vis-à-vis du projet professionnel. L'Adaptabilité est composée de la préoccupation pour l'avenir, du contrôle et de la planification du parcours, de l'exploration des possibilités professionnelles et de la confiance dans sa capacité à réussir ses choix de carrière. Enfin, la Résilience regroupe la résilience générale, l'estime de soi, l'auto-efficacité liée au choix professionnel et l'auto-efficacité professionnelle.

Voici trois exemples d'items du *Berufswahlkompetenz* :

« *Je n'ai aucune idée du métier qui pourrait vraiment me convenir.* » ;

« *Choisir un métier par passion n'a pas de sens s'il n'offre pas de débouchés.* » ;

« *Mes parents savent mieux que moi quel métier je devrais choisir.* »

.

4.2.2. Career Decision-Making Difficulties Questionnaire (CDDQ ; Gati et al., 1996 ; Rossier et al., 2021)

Le *Career Decision-Making Difficulties Questionnaire* (CDDQ), développé par Gati et al. (1996) et largement utilisé dans la recherche sur l'orientation professionnelle. Permettant d'évaluer les obstacles rencontrés lors du choix d'une carrière. Il se compose de 34 affirmations, évaluées sur une échelle de Likert à neuf points, allant de 1 (« Ne me décrit pas du tout ») à 9 (« Me décrit très bien »).

Les items sont regroupés en trois dimensions principales. La première, le manque de préparation, englobe notamment une faible motivation, des croyances irrationnelles ainsi qu'une indécision persistante. La deuxième, le manque d'information, concerne l'insuffisance de connaissances sur soi, sur les options professionnelles disponibles ou sur le processus de prise de décision. Enfin, la troisième dimension, les informations inconsistantes, implique la présence de conflits internes ou externes et d'informations contradictoires.

Pour chaque sous-échelle, on calcule un score moyen ; un score total représente le degré d'indécision vocationnelle. Des valeurs plus élevées indiquent des difficultés plus importantes dans le processus décisionnel.

La version française du CDDQ a été validée par plusieurs études. Récemment, lors d'une étude de Rossier et al., (2022) une excellente consistance interne pour le score total du questionnaire (α = .93), ainsi qu'une consistance interne comprise entre .57 et .93 pour les trois sous-dimensions, a été rapportée.

4.3. Procédures

L'étude a été menée à l'aide d'un questionnaire en ligne, administré via la plateforme

LimeSurvey. Après avoir accédé au questionnaire par le lien transmis par M. Houssemand, les participants ont été informés des objectifs de l'étude, du traitement confidentiel des données et du respect de l'anonymat. Le consentement était requis pour pouvoir poursuivre le questionnaire. La participation était volontaire et les participants pouvaient se retirer de l'étude à tout moment, sans justification et sans conséquence. Les questionnaires incomplets ont été exclus des analyses.

Tout d'abord, des données sociodémographiques ont été recueillies, notamment l'âge, le genre, la nationalité et le niveau d'études. Puis, les participants ont complété le Career Decision-Making Difficulties Questionnaire (CDDQ) afin d'évaluer l'indécision vocationnelle, suivi du questionnaire Berufswahlkompetenz, utilisé pour mesurer la maturité vocationnelle.

La durée totale était estimée à environ 20 minutes. Initialement, le questionnaire devait être accessible pendant une semaine, mais la période de collecte des données a été prolongée à près de trois semaines afin d'atteindre un plus grand nombre de participants.

Les analyses statistiques ont été réalisées à l'aide du logiciel Jamovi.

5. Résultats

5.1 Analyses descriptives

L'échantillon final se compose de 35 étudiants en psychologie (*N* = 35). L'âge moyen des participants est de *M* = 22.90 ans (ET = 6.60), avec un âge minimum de 18 ans et un âge maximum de 52 ans. Concernant le genre, 80.0 % des participants s'identifient comme femmes (*n* = 28), 14.3 % comme hommes (*n* = 5) et 5.7 % comme autres (*n* = 2). Les participants sont inscrits en Bachelor (*n* = 19 ; 54.3 %) ou en Master (*n* = 16 ; 45.7%). De plus, 22.9 % (*n* = 8) des participants ont déclaré que leurs études en psychologie constituaient une reprise d'études, tandis que 77.1 % (*n* = 27) n'ont pas interrompu leur parcours académique.

Tableau 1 :
Caractéristiques sociodémographiques de l'échantillon (N = 35)

	n	%	*M*	ET	Min-Max
Âge (années)	-	-	22.90	6.60	18-52
Genre					
Femmes	28	80.0	-	-	-
Hommes	5	14.3	-	-	-
Autre	2	5.7	-	-	-
Cursus					
Bachelor	19	54.3	-	-	-
Master	16	45.7	-	-	-
Reprise d'études					
Oui	8	22.9	-	-	-
Non	27	77.1	-	-	-

Note. ET = écart-type.

5.2 Fidélité des instruments

La cohérence interne des instruments a été évaluée à l'aide de l'alpha de Cronbach (α) et de l'oméga de McDonald (ω), afin de déterminer la pertinence du calcul de scores composites.

5.2.1 Fiabilité du Career-Decision-Making Difficulties Questionnaire (CDDQ)

Les analyses de cohérence interne indiquent une fidélité acceptable à élevée pour les dimensions du CDDQ. La dimension *Readiness* présente une cohérence interne acceptable (α = .772 ; ω = .789). La dimension *Lack of Information* présente une cohérence interne élevée (α = .893 ; ω = .897). La dimension *Difficulties Related to Inconsistent Information*

présente une cohérence interne élevée (α = .851 ; ω = .861).
Le score composite global du CDDQ montre une excellente cohérence interne (α = .909 ; ω =.915).

Tableau 2 :
Statistiques descriptives et indices de fidélité du CDDQ

	M	ET	α	ω
Readiness	3.92	1.22	.772	.789
Lack of Information	4.12	1.43	.893	.897
Inconsistent Information	3.24	1.37	.851	.861
Score global	3.78	1.08	.909	.915

Note. ET = écart-type.

5.2.2 Questionnaire de maturité vocationnelle (Berufswahlkompetenz) Les analyses de cohérence interne indiquent une fidélité acceptable pour les dimensions du questionnaire de maturité vocationnelle. L'identité présente une cohérence interne acceptable (α = .710 ; ω = .751). L'adaptabilité présente une cohérence interne acceptable (α = .738 ; ω = .783). La résilience présente une cohérence interne satisfaisante (α = .766 ; ω = .792).
Tableau 3 :
Statistiques descriptives et indices de fidélité du questionnaire de maturité vocationnelle

	M	ET	α	ω
Identité	1.88	0.31	.710	.751
Adaptabilité	3.50	0.27	.738	.783
Résilience	3.29	0.32	.766	.792

Note. ET = écart-type.

5.3 Tests des hypothèses

5.3.1 Hypothèse I

Des tests t pour échantillon indépendant ont été réalisés afin d'examiner d'éventuelles différences entre les étudiants de Bachelor et de Master au sujet de la maturité vocationnelle.
Les résultats indiquent aucune différence significative entre les groupes pour l'identité, $t(33) = 1.50$, $p = .143$, $d = 0.51$, ni pour l'adaptabilité, $t(33) = 0.43$, $p = .668$, $d = 0.15$, ni pour la résilience, $t(33) = -0,07$, $p = .946$, $d = 0.02$.
Tableau 4 :
Comparaison entre Bachelor et Master sur les dimensions de la maturité vocationnelle

	Cursus	*n*	*M*	ET	*t*	*p*
I.	B.	19	1.92	0.27	1.50	.143
	M.	16	1.77	0.33	-	-
A.	B.	19	3.52	0.24	0.43	.668
	M.	16	3.48	0.31	-	-
R.	B.	19	3.28	0.31	- 0.07	.946
	M.	16	3.29	0.34	-	-

Note. ET = écart-type.

5.3.2 Hypothèse II

Une matrice de corrélation a été effectuée afin d'examiner la relation entre l'identité et l'indécision vocationnelle.
Les résultats indiquent une corrélation positive significative, $r(33) = .489$, $p < .01$.
Tableau 5 :
Matrice de corrélation entre l'identité et l'indécision vocationnelle

	Identité	Indécision vocationnelle

1. Identité	-	.489**
2. Indécision vocationnelle	-	-

Note. ***p*** < .01.

5.3.3 Hypothèse III

Une matrice de corrélation a été réalisée pour examiner les relations entre l'indécision vocationnelle et les dimensions de la maturité vocationnelle.

Les résultats indiquent une corrélation positive significative entre l'indécision vocationnelle et l'identité $r(33) = .489$, $p < 0.01$. Les résultats indiquent aucune corrélation significative entre l'indécision vocationnelle et l'adaptabilité, $r(33) = .070$, $p = .668$, ni entre l'indécision vocationnelle et la résilience, $r(33) = .003$, $p = .946$.

Tableau 6 :

Matrice de corrélation entre l'indécision vocationnelle et les dimensions de la maturité vocationnelle

	1	2	3	4
1.Indécision vocationnelle	-	.489**	.070	.003
2. Identité	-	-	-	-
3. Adaptabilité	-	-	-	-
4. Résilience	-	-	-	-

Note. ***p*** < .01.

5.3.4 Hypothèse IV

Une matrice de corrélations a été calculée pour examiner les relations entre l'âge et la maturité vocationnelle, et les relations entre l'âge et l'indécision vocationnelle.

Les résultats montrent que l'âge est négativement corrélé à l'identité, $r(33) = -.402$, $p < .05$. L'âge n'est pas significativement corrélé à l'adaptabilité, $r(33) = -.111$, $p = .524$, ni à la résilience, $r(33) = .003$, $p = .986$.

Les résultats indiquent une corrélation négative significative entre l'âge et l'indécision vocationnelle, $r(33) = -.34$, $p < .05$.

Tableau 7 :

Relation entre âge et maturité

	Indécision vocationnelle	Âge
Indécision vocationnelle	-	-.34*
Âge	-	-

Note. **p** < .05.

5.3.5 Hypothèse V

Des tests *t* pour échantillons indépendants ont été réalisés afin de comparer les participants en reprise d'études ($n = 8$) à ceux n'ayant pas interrompu leur parcours ($n = 27$).

Une différence significative est observée pour l'identité, $t(33) = 2.24$, $p < .05$, $d = 0.90$. Aucune différence significative n'est observée pour l'adaptabilité, $t(33) = 1.65$, $p = .108$, $d = 0.66$, ni pour la résilience, $t(33) = 0.21$, $p = .836$, $d = 0.08$.

L'indécision vocationnelle ne diffère pas entre les groupes, $t(33) = 0.29$, $p = .776$, $d = 0.116$.

Tableau 8 :

Comparaison selon la reprise d'études

	t	*p*
Identité	2.24	.032*
Adaptabilité	1.65	.108
Résilience	0.21	.836
Indécision vocationnelle	0.29	.776

Note. **p** < *.05.*

5.4 Analyse de régression linéaire

Une analyse de régression linéaire a été réalisée afin d'examiner dans quelle mesure la ma-

turité vocationnelle prédit l'indécision vocationnelle.

Le modèle de régression linéaire est significatif, F (1, 33) = 23.70, p < .001, et explique 41.8 % de la variance de l'indécision vocationnelle (R^2 = .418).

Tableau 9 :

Mesure de l'ajustement du modèle de régression

Modèle	*R*	R^2
1	.646	.418

Note. N = 35.

Tableau 10 :

Coefficients de la régression linéaire multiple prédisant l'indécision vocationnelle

Prédicteur	Estimation	Erreur standard	*t*	*p*
Ordonnée à l'origine	6.754	2.320	2.91	.007
Identité	1.620	0.481	3.36	.002
Adaptabilité	-0.757	0.582	-1.30	.203
Résilience	-1.051	0.491	-2.14	0.040

Note. Les estimations correspondent aux coefficients non standardisés.

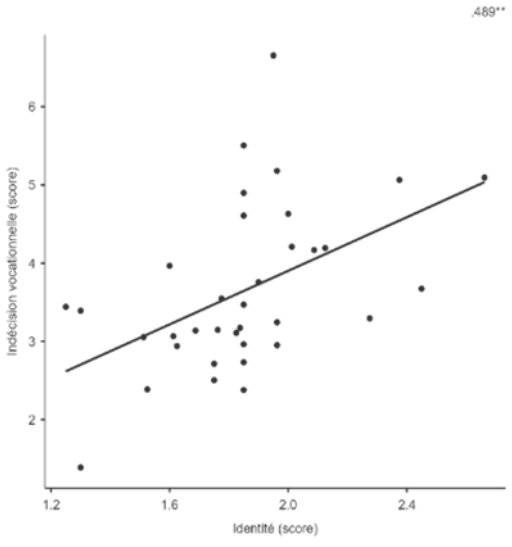


Figure 1

Relation entre l'identité et l'indécision vocationnelle

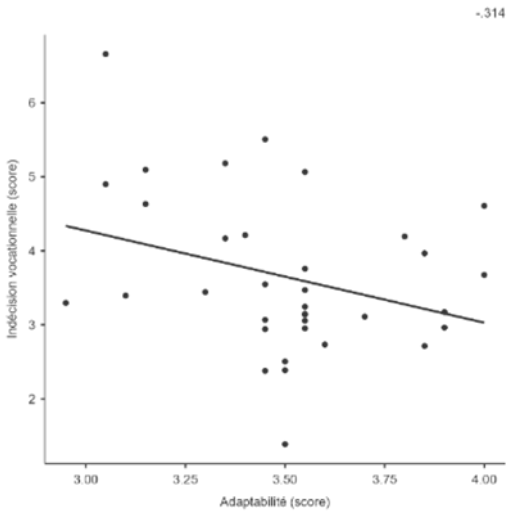


Figure 2

Relation entre l'adaptabilité et l'indécision vocationnelle

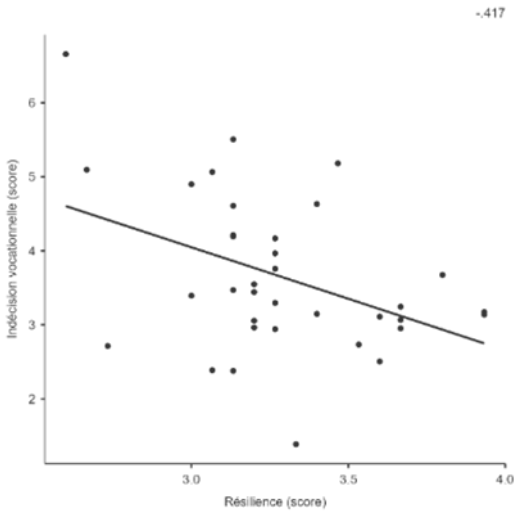


Figure 3

Relation entre la résilience et l'indécision vocationnelle

6. Discussion

6.1 Hypothèse I

La première hypothèse postulait que le fait d'être plus avancé dans le cursus universitaire entraînerait une différence quant à la maturité vocationnelle. Toutefois, les résultats ne mon-

trent aucune différence concernant l'identité, l'adaptabilité et la résilience entre les étudiants en Bachelor et en Master. Ces résultats suggèrent que la maturité vocationnelle semble être liée à un processus individuel qui se développe indépendamment du parcours académique (Ratschinski, 2014).

6.2 Hypothèse II

La seconde hypothèse affirmait qu'une identité plus forte serait associée à une indécision moindre. Les résultats de l'étude montrent qu'au contraire, l'identité vocationnelle et l'indécision vocationnelle présentent une relation positive significative.
Ces résultats montrent que l'identité renvoie davantage à une phase d'exploration identitaire active qu'à un concept cristallisé. Ainsi, un score élevé d'identité peut augmenter temporairement l'indécision (Gati et al., 1996 ; Ratschinski, 2014).

6.3 Hypothèse III

La troisième hypothèse postulait que des niveaux plus élevés d'identité, d'adaptabilité et de résilience seraient associés à une indécision vocationnelle moindre. Cette hypothèse n'a toutefois pas pu être soutenue par les résultats de l'étude.
En effet, l'identité est positivement corrélée à l'indécision vocationnelle, reflétant un engagement intense dans l'exploration, accompagné d'une phase riche en doutes concernant le choix de carrière. De plus, l'adaptabilité n'est pas significativement associée à l'indécision vocationnelle. Enfin, la résilience est associée à une indécision plus faible, suggérant qu'elle contribue à réduire les difficultés rencontrées lors de la prise de décision (Gati et al., 1996 ; Ratschinski, 2014).

6.4 Hypothèse IV

L'hypothèse suggérait que les étudiants plus âgés rencontreraient moins de difficultés dans la prise de décisions de carrière. Les résultats de l'étude confirment cette hypothèse, en montrant une diminution de l'indécision vocationnelle avec l'âge.
Ce résultat est cohérent avec le modèle du Career Decision-Making Difficulties Questionnaire (CDDQ), selon lequel l'indécision vocationnelle peut notamment être liée à un manque d'informations ou à des informations conflictuelles. Ces difficultés tendent à diminuer avec l'accumulation d'expériences personnelles et professionnelles, favorisant une meilleure connaissance de soi et des options disponibles (Gati et al., 1996).
Par ailleurs, l'hypothèse proposait également que les étudiants plus âgés présenteraient une maturité vocationnelle plus élevée. Toutefois, aucune relation significative entre l'âge et les dimensions de la maturité vocationnelle (identité, adaptabilité et résilience) n'a été mise en évidence. Cette absence de lien suggère que la maturité vocationnelle dépend davantage de la capacité à réfléchir sur soi et à mobiliser des compétences spécifiques que de l'âge chronologique en tant que tel (Ratschinski, 2014).

6.5 Hypothèse V

L'hypothèse postulait que les étudiants ayant repris leurs études présenteraient une maturité vocationnelle plus élevée et une indécision vocationnelle plus faible que ceux ayant un parcours académique linéaire.
Les résultats montrent une différence significative concernant l'identité vocationnelle. Une reprise d'études est souvent marquée par une réévaluation des objectifs professionnels et peut ainsi correspondre à une phase d'exploration identitaire, au cours de laquelle les individus s'engagent plus activement dans une réflexion sur leur avenir professionnel (Ratschinski, 2014). Dans ce contexte, la sous-dimension de l'engagement professionnel associée à l'identité peut être plus élevée sans s'accompagner d'une augmentation de

l'indécision vocationnelle. En effet, selon le modèle du Career Decision-Making Difficulties Questionnaire, l'indécision peut apparaître comme une composante fonctionnelle du processus de décision et ne reflète pas nécessairement un blocage ou une difficulté accrue (Gati et al., 1996). Ainsi, une identité vocationnelle plus élevée ne conduit pas nécessairement à une augmentation des difficultés dans la prise de décision, mais peut plutôt refléter une réflexion plus approfondie concernant l'avenir professionnel.

7. Limites

Plusieurs limites doivent être prises en compte lors de l'interprétation des résultats de cette étude. Tout d'abord, la taille réduite de l'échantillon (N = 35), réduit fortement la puissance statistique. Les résultats doivent donc être interprétés avec prudence et ne peuvent pas être généralisés.

D'autant plus que le caractère transversal de l'étude ne permet pas d'établir de relations causales entre la maturité vocationnelle et l'indécision vocationnelle. L'évolution temporelle des variables et de leur relation n'a pas été prise en compte dans cette étude.

Par ailleurs, l'utilisation de mesures auto-rapportées peut être cause de certains biais, tels que la désirabilité sociale. Les réponses des participants peuvent ainsi ne pas refléter leur situation réelle, mais plutôt leur perception subjective.

Finalement, il est important de préciser que les résultats sont liés au contexte. Donc, les résultats concernent que les étudiants en psychologie de l'Université du Luxembourg. D'autres filières académiques ou établissement universitaire pourraient présenter des résultats complétement différents à ceux présenter dans cette étude.

8. Conclusion

De manière générale, cette étude se tourne vers la complexité du choix de carrière auprès des étudiants en psychologie à l'Université du Luxembourg. Ces résultats suggèrent que l'identité vocationnelle peut refléter une phase d'exploration active associée à une augmentation temporaire de l'indécision, tandis que la résilience semble jouer un rôle de régulation.

L'association entre l'identité et l'indécision vocationnelle suggère que l'exploration active peut temporairement augmenter l'incertitude, sans pour autant souligner une vraie difficulté dans le choix de la future carrière. Par ailleurs, l'âge apparait comme un facteur protecteur face à l'indécision. Ceci est probablement due aux expériences personnelles permettant une meilleure compréhension de soi et du monde du travail.

Ainsi, cette recherche invite à repenser l'indécision vocationnelle pas comme un obstacle, mais une étape normale du développement du soi.

De plus, il est important de poursuivre les recherches avec des échantillons plus larges et à caractères longitudinaux, afin de mieux comprendre l'évolution de la maturité vocationnelle au cours du temps et son rôle dans la réduction de l'indécision vocationnelle. De futures études pourraient également intégrer des évaluations externes afin de compléter le données auto-rapportées.

9. Littérature

Dosnon, O. (1996). L'indécision face au choix scolaire ou professionnel : Concepts et mesures. *L'Orientation scolaire et professionnelle*, *25*(1), 129-168. https://doi.org/10.3406/binop.1996.1136

Forner, Y., & Dosnon, O. (1991). La maturité vocationnelle : Le processus et son évaluation. *L'Orientation scolaire et professionnelle*, *20*(2),

203-219. https://doi.org/10.3406/binop.1991.1361

Gati, I., Krausz, M., & Osipow, S. H. (1996). A taxonomy of difficulties in career decision making. *Journal of Counseling Psychology*, *43*(4), 510-526. https://doi.org/10.1037/0022-0167.43.4.510

Guo, L. (2025). Unsettled horizon : Adolescents' career expectations in the volatile, uncertain, complex, and ambiguous contexts. *Humanities and Social Sciences Communications*, *12*(1), 950. https://doi.org/10.1057/s41599-025-05298-6

Huteau, M. (2007). Maturité vocationnelle (vocational maturity, career maturity). In *Orientation et insertion professionnelle* (p. 284-287). Dunod. https://doi.org/10.3917/dunod.guich.2008.01.0284

Mata-Correas, M., Hernández-Franco, V., & Prieto-Ursúa, M. (2025). *Career indecision and career indecisiveness in high school and vocational training students : A scoping review*.

Pignault, A., Rastoder, M., & Houssemand, C. (2023). The Relationship between Self-Esteem, Self-Efficacy, and Career Decision-Making Difficulties : Psychological Flourishing as a Mediator. *European Journal of Investigation in Health, Psychology and Education*, *13*(9), 1553-1568. https://doi.org/10.3390/ejihpe13090113

Ratschinski, G. (2018). *Selbstkonzept und Berufswahl : Eine Überprüfung der Berufswahltheorie von Gottfredson an Sekundarschülern*. Waxmann Verlag.

Rossier, J., Rochat, S., Sovet, L., & Bernaud, J.-L. (2022). Validation of a French Version of the Career Decision-Making Difficulties Questionnaire : Relationships With Self-Esteem and Self-Efficacy. *Journal of Career Development*, *49*(4), 906-921. https://doi.org/10.1177/08948453211009975

Gritty Parenting: Understanding the connection between grit, parenting, and imposter phenomenon

Emily Paradowska, Elin Nilles, Giulia Rotolo, Chiara-Amélie Thomé and Sarah Koeune Thiltges

Supervisors: Dr. rer. nat. Ziwen Teuber and Dr. phil. Anke Maria Weber

Parental perseverance and self-doubt shape the way parents interact with their children. Parental grit refers to the tendency to sustain passion and perseverance toward long-term goals, while the impostor phenomenon describes an internal experience in which individuals believe they have deceived others about their abilities and achievements. This study aimed to examine how parental grit and impostor feelings are associated with autonomy-supportive and psychologically controlling parenting styles. Furthermore, the study tested possible mediation effects between these constructs, using parental impostor feelings as the mediator. A cross-sectional survey was conducted with 148 parents of children aged 5 to 18. Pearson correlations and path analyses were used to examine both direct and mediated relationships among the variables. Results showed that higher pa-rental grit was linked to more autonomy-supportive parenting. Grit was not associated with impostor feelings or with psychological control. Impostor feelings did not show a relationship with autonomy support but were modestly linked to greater psychological control. No significant indirect effects were identified. These findings suggest that grit and impostor feelings each relate to parenting behaviors independently. Further research should consider additional factors, such as parental self-efficacy and the dynamic nature of parent–child relationships, to better understand the foundations of parenting and its influencing factors.

Keywords: parental grit, impostor phenomenon, parenting styles, autonomy support, psychological control

1. Introduction

1.1 General introduction

Parenting is a challenging and stressful task (Nelson et al., 2014) that requires persistent effort and confidence under constant uncertainty on how to raise a child properly. In this context, two psychological constructs seem to play a key role in understanding parents' functioning. First, parental grit, or the persistent engagement of parents in pursuing their long-term parenting goals (Teuber et al., 2024b), and second, the impostor phenomenon, also known as the self-doubt a person might experience despite evident accomplishments (Sakulku & Alexander, 2011). While there is research examining both constructs separately, their joint role does not yet seem fully understood.

As part of our study, we created a structured survey to assess parental grit, impostor phenomenon, and parenting styles (autonomy support and psychological control) to examine possible impostor feelings parents may experience when exerting different levels of grit towards their children, with a focus on how this might influence diverse parenting styles.

1.2 The importance of grit

The concept of grit was initially defined as a non-cognitive trait linked to resilience and success. Grit refers to the tendency to sustain passion and perseverance towards the achievement of long-term goals (Duckworth et

al., 2007). This definition was later refined by subsequent studies, which added the concept of adaptability to persist in the face of difficulties (Datu et al., 2017). Furthermore, unlike intelligence or pure talent, grit reflects persistent effort and commitment despite setbacks, failures, or obstacles (Duckworth et al., 2007). Additionally, grit can be clearly distinguished from intrinsic motivation, which does not include any supportive conditions to help maintain this tendency in the face of setbacks or difficulties (Deci & Ryan, 2000). Further studies also suggest that grit is strongly linked to higher achievement (Duckworth et al., 2007; Bowman et al., 2015), better self-regulation (Wolters & Hussain, 2015), and greater persistence in the face of failure (Lucas et al., 2015).

1.3 Domain-specific grit scale: The importance of parental grit

Importantly, in addition to the previously defined general concept, grit can be conceptualized in domain-specific grit scales, such as in education (Clark & Malecki, 2019), sports (Guelmami et al., 2022; Fruchart & Rulence-Pâques, 2025), or work (Suzuki et al., 2015). This project focuses specifically on parental grit, a domain-specific grit scale referring to sustained, goal-oriented efforts directed toward everyday caregiving (Teuber et al., 2024b). It influences how parents cope with the ongoing challenges of raising their children. Evidence suggests that parents with higher grit tend to exhibit greater psychological flexibility and goal orientation, which helps them experience less emotional exhaustion throughout the parenting process (Teuber et al., 2024b). Furthermore, Du et al. (2023) report that parents experiencing higher levels of grit tend to offer more autonomy support to their children, indicating that gritty parents show greater confidence and trust in their children's behavior and decision-making. Gritty parents might also cope better with daily challenges and consequently feel less emotionally exhausted throughout the parenting process.

It is important to note that, to this date, though there exists some research about how general grit in parents influences children's well-being (e.g., Fernández-Martín et al., 2023; Lam et al., 2025), we have found little research focusing on how this specific grit influences parents' own self-perception and well-being. This study, therefore, focuses less on the child's development and more on understanding how parental grit can help parents in achieving their own parental goals.

1.4 Parents' self-doubts and impostor feelings

To properly study the influence of parental grit on parents, it is important to acknowledge how negative feelings, such as self-doubt or insecurity, might affect experienced parental grit. Vice versa, specific levels of parental grit might encourage such feelings. An important and common phenomenon that helps explain these feelings is the impostor phenomenon, also known as impostor syndrome or impostor experience. Many people, especially when faced with a task for the first time, might doubt their capabilities, even though they objectively deliver positive results. Similarly, some people feel like an "impostor" in certain areas and consequently do not consider themselves worthy of their own achievements. The impostor phenomenon describes exactly this: an internal experience in which people believe they have deceived others about their abilities and achievements (Mukti et al., 2024). This means that people experiencing this phenomenon tend not to fully acknowledge their own talents and capacities when working on a task. Huecker et al. (2023) characterize impostor feelings by self-doubt, anxiety, depression, and/or apprehension of being exposed as a "fraud" due to not living up to their idealized image. Furthermore, people with imposter feel-

ings often feel incompetent and unworthy of their accomplishments, attributing their success to luck or other external factors, disregarding their own abilities (Langford & Clance, 1993).

In this context, parents with high expectations of themselves might disregard their actual abilities and see themselves as incompetent, even when they are objectively doing well in their parental role. This mindset keeps parents in constant fear of not being able to maintain their success (Langford & Clance, 1993) and can even rob them of the ability to function with joy (Clance & OToole, 1988; Hirschfeld, 1982). In some families, especially among mothers, impostor feelings may be reinforced by specific child-rearing patterns that can leave them feeling conflicted about their own autonomy (Clance & OToole, 1988). Furthermore, encountering a challenging task, such as parenting, is likely to lead people to experience strong doubt and fear, wondering whether they will manage to succeed on this occasion (Clance & OToole, 1988). Impostor feelings in parents could therefore reduce satisfaction within parenting, trapping them in a cycle of overwork and self-doubt.

It is important to note that the impostor phenomenon, also known as imposter syndrome in lay literature (Bravata et al., 2020), has not been listed as a psychiatric disorder to date. No evidence or specific code of such a diagnosis can be found in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR) (APA, 2022), nor in the International Statistical Classification of Diseases and Related Health Problems, Eleventh Revision (ICD-11) (World Health Organization, 2022). The impostor phenomenon is therefore not to be seen as a state, but rather as a trait (Feenstra et al., 2020) or psychological experience (Mak et al., 2019).

1.5 Autonomy support and psychological control – Self-determination theory

The self-determination theory is an empirical theory that focuses on the "human motivation and personality in social contexts that differentiates motivation in terms of being autonomous and controlled" (Deci & Ryan, 2000, p. 416) The theory analyses various aspects, for example, the effects of social environments on human qualities such as motivation, values, and attitudes. It assumes that humans are naturally active, motivated, and "oriented toward developing naturally through integrative processes" (Deci & Ryan, 2000, p. 417) These qualities "are inherent in human nature" (Deci & Ryan, 2000, p. 417) but can develop and be influenced by the environment.

Within the framework of self-determination theory, several mini theories were developed, such as the "basic psychological needs theory" Ryan et al. (1996, as cited in Deci & Ryan, 2000, p. 423), according to which specific psychological and biological nutriments must be fulfilled for "intrinsic motivation and integration to operate effectively toward healthy development and psychological well-being" Ryan (1995, as cited in Deci & Ryan, 2000, p. 417). When these needs are disturbed, these processes will be diminished, which can lead to inadequate experiences, development, and behaviors (Deci & Ryan, 2000). More specifically, this theory focuses on the psychological nutriments and assumes that there are three universal needs: autonomy, relatedness, and competence (Deci & Ryan, 2000). These needs are considered highly important for human functioning, development, and well-being (Deci & Ryan, 2000). The theory assumes that these universal needs are at the core of how people organize their behaviors, and that they can be individually adapted (Deci & Ryan,

2000). Furthermore, various studies have shown that intrinsic motivation, satisfaction, and well-being rise when these universal nutriments are fulfilled (Deci & Ryan, 2000).

Additionally, people are more likely to be active and to persist (Grolnick, 2015) when they feel autonomous rather than controlled in their activities. It has also been shown that people who feel more in control of their motivation tend to show more determination and engagement (Grolnick, 2015). Autonomous motivation during activities is therefore mainly associated with positive feelings, such as enjoyment. Controlled motivation, however, is associated with negative feelings such as pressure, suggesting that parents experience more positive feelings during activities that they engage in themselves Ryan (1982, as cited in Grolnick, 2015).

In the context of parenting, the need for autonomy seems to play an important role.

Parental behavior can influence children's experience of independence. Controlling behavior, characterized by dominance and pressure, can reduce the child's feelings of independence and competence (Dumont et al., 2014), whereas autonomy-supportive behavior (Dumont et al., 2014) can strengthen these feelings. Studies suggest that higher autonomy support from parents increases children's autonomous motivation, meaning that they offer their children choices, consider their perspectives and thereby encourage them, for instance, to take the initiative to solve everyday problems on their own or together Deci & Ryan (1985, as cited in Lerner et al., 2022) ; Deci & Ryan (2017, as cited in Lerner et al., 2022). On the contrary, parental control increases children's controlled motivation, meaning that such parents tend to control their children's behavior and decision-making. Psychological control primarily influences the child's emotional and psychological development (Dumont et al., 2014). Furthermore, the tendency to exercise a high degree of psychological control can, for example, arise from self-doubt (Deci & Ryan, 2000), suggesting an association between this construct and impostor feelings.

1.6 Hypotheses

H1: Higher levels of parental grit predict lower levels of parental imposter feelings, which in turn predict higher levels of autonomy support.

H2: Higher levels of parental grit predict lower levels of parental imposter feelings, which in turn predict lower levels of psychological control.

2. Methodology

2.1 Sample and Data Collection

This study employs a quantitative, cross-sectional, correlational design and an online questionnaire to examine the relationships between parental grit, parenting style, and parental imposter feelings. The study is quantitative because participants responded to items measured on Likert scales, enabling the calculation of descriptive statistics and the examination of correlations among variables. This study is correlational because the aim was to investigate associations between variables rather than to establish a causal relationship.

The intended target sample was 300 participants. However, due to time constraints, a total of 148 valid questionnaires were available for analysis. Although more than 150 individuals started or completed the questionnaire, a

subset of responses could not be included due to invalid or incomplete data. The target group consisted of parents of school-aged children (ages 5-18).

Participants were recruited through flyers and social media platforms, such as Facebook and Instagram. The flyers included three QR codes that linked directly to the online questionnaire, available in English, French, and German. Flyers were distributed both online and physically in public places.

After scanning the QR code, participants were redirected to a Google Forms questionnaire containing 29 questions. The questionnaire assessed parental grit, imposter phenomenon, and parenting style, specifically autonomy support and psychological control. Completing the survey took approximately 10 to 15 minutes.

Data were collected over a two-week period, and each participant completed the questionnaire only once and voluntarily. Before the participation, all participants provided their informed consent. No directly identifying personal information was required to take part in the study.

Complete anonymity cannot be fully guaranteed when using Google Forms, as the platform may temporarily store limited metadata (e.g., IP addresses or technical information) on its secure servers. However, the researchers did not have any access to any such metadata. Participants were informed about this potential limitation before participating in the study. An email address was collected only voluntarily for the purpose of entering a voucher raffle.

All questionnaire responses were treated as strictly confidential and were accessed exclusively by the research team.

All data handling procedures complied with the ethical standards approved by the University of Luxembourg Ethics Review Panel.

2.2 Measures

To conduct our study, we used a questionnaire made in "Google Forms" available in three languages. We translated the original German version of Teuber et al. (2024) into English and French using software "DeepL" and guidance and feedback from our supervisors to ensure the wording was accurate.

Before beginning with the survey, participants had to agree to the private policy. The questionnaire included 29 questions and 95 items, which took about ten minutes to complete. The survey consisted of multiple-choice, short-answer, and Likert-scale questions. Questions based on the Likert scale varied from 1 to 5; from “strongly disagree” to “strongly agree” and from “not like me at all” to “very much like me.” There was a single question using a Likert scale ranging from 1 to 7, from “never” to “every day.”

For the questionnaire we specifically referred to the four following scales: parental grit based on twelve items (Cronbach’s α= .93); for example “As a parent, I persist to achieve long-term ambitions, even after frustrating mo-

ments.", parental imposter based on five items (Cronbach's α= .89); for instance "I'm afraid people may find out that I'm not as capable as they think I am in parenting.", autonomy support based on five items (Cronbach's α= .82); for example "I ask my youngest child how I can help them.", and psychological control based on four items (Cronbach's α= .79); for instance "I scold my youngest child and request that they learn more.".

The scales we have not analyzed, but were part of the survey, concerned questions about demographics, exhaustion in the parental role, parental self-efficacy, perceptions of the child's emotional and social inclusion, PIQ-Parent version, General Grit, conscientiousness, individualism, and collectivism.

2.3 Data Analysis

The primary method we used to test our hypotheses was path analysis. This method allowed us to test for both direct and mediated pathways–with Parental Impostor as the mediator–between Parental Grit and Autonomy Support, and between Parental Grit and Psychological Control. We conducted all analyses in R using RStudio. With the help of our supervisors, we wrote a code that first tested the reliability of our scales (Cronbach's α), computed the descriptive statistics, then examined the correlations between our four main constructs (Parental Grit, Parental Imposter, Autonomy Support, and Psychological Control), and finally estimated the path model. The main packages we used were psych (Revelle, 2007) for descriptive statistics and correlations, lavaan (Rosseel, 2012) for path analysis, semPlot (Epskamp, 2013) for model plotting, and MBESS (Kelley, 2006) for testing the significance of the indirect effects of Parental Grit on Autonomy Support and Psychological Control. We based all analyses on complete cases (N = 148) after excluding cases with missing data using listwise deletion.

3. Results

3.1 Abbreviations

The following abbreviations, present in the tables with results, were used to describe the main constructs:

ParGrit = Parental Grit

ParImp = Parental Impostor

AuSup = Autonomy Support

PsyCon = Psychological Control

3.2 Correlations

To test for relationships among our variables, we conducted Pearson correlations (Table 1). First, these correlations showed that grittier parents tended to provide more autonomy support (r = .46, p < .001), which is consistent with the existing literature (e.g., Du et al., 2023). Parental grit, however, did not show any significant relationship with psychological control (r = -0.04, p = .645) or impostor syndrome (r = .10, p = .910). Similarly, parental impostor was not significantly related to autonomy support (r = –.03, p = .742) but showed a slight association to psychological control (r = .17, p = .036). Autonomy support and psychological control were, as expected, not correlated (r = -.01, p = .896).

Table 1

Pearson correlations

	ParGrit	ParImp	AuSup
ParImp	0.10		
AuSup	0.46**	-0-03	

PsyCon	-0.04	0.17*	-0.01

Note. ** p < 0.001; * p < 0.05

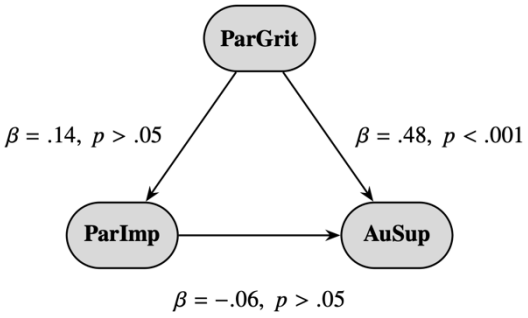
3.3 Path Analysis

We conducted two path analyses, with autonomy support as the outcome in the first (Table 2) and psychological control as the outcome in the second (Table 3). Both models had parental grit as the predictor and parenting impostor as the mediator.

The first model showed that Parental grit indeed had a direct effect on autonomy support (β = .48, p < .001), but it did not significantly predict impostor feelings (β = .14, p = .222). Furthermore, parental impostor did not predict autonomy support (β = –.06, p = .313). Finally, the indirect effect of grit on autonomy support through impostor feelings was small and non-significant (β = –.008, p = .437; 95% bootstrap CI [-.041, .019]). The model also explained about 22% of the variance in autonomy support (R^2 = .22).

Table 2

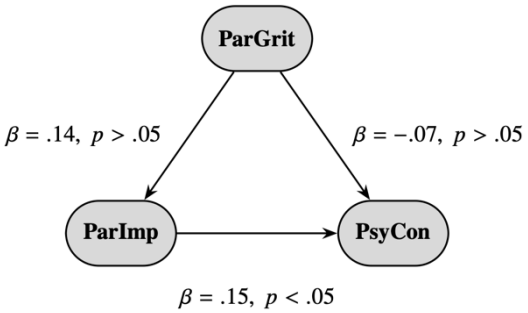
Path analysis 1: Effect of parental grit on autonomy support through impostor feelings



The second model we tested showed similar results. Parental grit did not significantly predict psychological control (β = –.07, p = .491), and it again showed no significant association with impostor feelings (β = .14, p = .222). However, higher parental impostor significantly predicted more psychological control (β = .15, p = .028). Additionally, the indirect effect of grit on psychological control through impostor feelings was small and non-significant (β = .02, p = .286; 95% bootstrap CI [-.010, .072]). Finally, this model explained about 3% of the variance in psychological control (R^2 = .03).

Table 3

Path analysis 2: Effect of parental grit on psychological control through impostor feelings



4. Discussion

4.1 Main findings

Our study aimed to analyze the association between parental grit, parental imposter feelings, autonomy support, and psychological control. In addition, the study sought to explore how these factors affect parents' self-concept and perseverance. Through our analysis we encountered various results.

First, the Pearson correlations showed that gritty parents generally provide more autonomy support. However, there was no significant correlation between parental grit and imposter syndrome or psychological control. Also, pa-

rental imposter was not significantly correlated with autonomy support but was positively correlated with psychological control. The data also showed that autonomy support and psychological control are not correlated.

Secondly, the path analysis showed that parental grit directly influences autonomy support but does not predict imposter feelings. It also showed that parental imposter does not predict autonomy support. Hereby, the indirect effect of grit on autonomy support via imposter syndrome was small and non-significant. Furthermore, we also found no mediation effect of parental grit on psychological control through parental impostor feelings.

In conclusion, our initial hypotheses could not be confirmed, as our analysis showed no indirect effects of parental grit on autonomy support, respectively on psychological control, via impostor feelings. Yet, our study showed that parental grit was positively correlated with autonomy support, meaning that parents with higher grit often provide stronger support for their children's autonomy. However, it showed that parental grit is not significantly associated with imposter feelings. Nevertheless, imposter feelings were related to higher psychological control.

4.2 Limitations

Several factors limit the generalizability of this study. Firstly, the gender of the participants is not evenly balanced, with 98 mothers, 50 fathers, and one non-binary person participating. The results are predominantly influenced by the mothers' perspective, with the fathers' perspective being underrepresented, limiting the possibility of generalizing the results equally for both parents. In this study, we did not use control variables. Still, we would suggest, for instance, using gender or socio-economic background as control variables in future studies to test whether results may vary depending on different factors.

Secondly, the survey is a self-report questionnaire. This makes the data susceptible to self-report bias, social desirability bias (Bornstein et al., 2015), recall bias (Coughlin, 1990), and common method bias (Aguirre-Urreta & Hu, 2019). It is therefore possible that the data collected does not reflect actual behavior and feelings. This limits the transferability of the results to real-life behavior.

Thirdly, the study has a cross-sectional design. Since the data are collected at a single point in time, no temporal or causal conclusions can be drawn, limiting the possibility of testing with certainty, as the results cannot be generalized to changes over time.

Lastly, the questionnaire is available in three languages, which increases accessibility but leaves room for distortion of interpretation due to cultural norms or translation differences.

4.3 Future directions

We acknowledge that several steps could strengthen future research building on our study.

First, to further elaborate on the study, a larger, more balanced sample should be aimed for to improve statistical power.

Additionally, using multi-method and multi-informant designs could reduce method bias by combining self-report questionnaires with additional sources such as partner reports, behavioral observations, and qualitative interviews. Similar findings across different methods would support higher construct validity, while incorporating multiple perspectives would offer a better understanding of the constructs and improve the generalizability of the results.

Moreover, including longitudinal designs would allow researchers to examine change over time and to draw more precise temporal conclusions about the effects. The use of experience-sampling designs could also further reduce recall bias and enable dynamic process analysis.

Furthermore, adding cross-cultural comparisons could further improve the generalization of the findings and help identify potential cultural influences.

Finally, future research could benefit from analysing more homogeneous subsamples to allow for more targeted analyses and clearer interpretations of the effects.

In addition, it is important to take a closer look at alternative mechanisms that may explain the observed association, such as stress (Ping et al., 2023), parental self-efficacy (Glatz et al., 2024), and bidirectional parent-child dynamics (Pardini, 2008) .

4.4 Conclusion

The present study aimed to examine the relationship between parental grit, parental imposter feelings, and parenting styles, with a particular focus on autonomy support and psychological control. By drawing on concepts from grit research (Duckworth et al., 2007; Teuber et al., 2024b) and self-determination theory (Deci & Ryan, 2000), this study provides a deeper understanding of how parents' perseverance and self-doubt relate to everyday parenting behaviors.

Overall, our findings indicate that parental grit is positively associated with autonomy-supportive parenting, suggesting that parents who report higher levels of grit tend to support their children's autonomy more. This result is theoretically consistent with self-determination theory, which emphasizes the role of perceived competence and confidence in autonomy supportive behavior (Deci & Ryan, 2000). Parents with higher grit may feel more capable of coping with parenting challenges, which could increase trust in their children's abilities and reduce the need for controlling strategies (Duckworth et al., 2007; Teuber et al., 2024b).

In contrast to our initial expectations, parental grit was not significantly related to parenting impostor feelings, and imposter feelings did not mediate the relationship between grit and autonomy support, nor did they mediate the relationship between grit and psychological control. One possible explanation could be that parental grit and imposter feelings represent distinct psychological processes (Clance & Imes, 1978). Parents may therefore remain behaviorally persistent despite the internal self-doubt.

However, higher levels of parenting imposter feelings were associated with higher psychological control, indicating that parental self-doubt may be linked to less supportive parenting practices. This finding also aligns with self-determination theory, which suggests that feelings of insecurity and low perceived competence can increase controlling behaviors as a means of regulating uncertainty (Deci & Ryan, 2000). The impostor phenomenon has been shown to be closely associated with anxiety and self-doubt, which may contribute to controlling parenting strategies (Clance & Imes, 1978).

These results suggest that parental grit may function as a protective factor for positive parenting behaviors, whereas imposter feelings may represent a potential factor for controlling parenting, independent of grit. These findings highlight the importance of considering parents' internal psychological experiences when examining parenting styles.

5. Appendix

5.1 Study information and Data Consent

Top of Form

Parental Grit Scale: English version

In this study, we want to understand how parent's perseverance and resilience, also known as "grit", influence the process of raising children in different cultures. You are invited to complete a questionnaire that examines how parents deal with challenges and support their children. The questionnaire takes about 10 minutes to complete. Your answers will help us to assess the relevance of this scale in different cultural and linguistic contexts. Thank you for your interest in our research.

This survey contains 29 questions.

Participation in this study is voluntary and you have the right to withdraw at any time after the study has begun.

Participants' responses will be collected via Google Forms. Complete anonymity cannot be fully guaranteed, as Google Forms may store limited metadata (e.g., IP addresses, E-mail address) as well as response data on its secure servers.

The researchers have access to the responses but not to any personal metadata. No identifying information will be collected unless participants voluntarily provide their email address to enter the prize draw. All data will be treated as strictly confidential, stored securely, and used solely for research purposes.

An anonymized version of the data from this study may be made publicly available, for example via the Open Science Framework (osf.io), without requiring additional written consent. The anonymized data may be used for further analysis, but also for additional analysis by the same or other researchers. The purpose and scope of this secondary use cannot be predicted. All personal data that could directly identify you will be removed before the data and results are published. All personal data collected about you will be stored separately from all other data. If you are

unclear about anything in the study or about participating in it, or if you would like to report a problem related to the conduct of the research, you can contact Dr Ziwen Teuber (ziwen.teuber@uni.lu). This study has been approved by the Ethics Committee of the University of Luxembourg.

* Indicates required question

Please enter your email address. If you win a 10 euro voucher we will send it to the email address you provide. Your email address will not be used for any other purpose. Bottom of Form.

5.2 Use of AI

For this report “DeepL” was used for translation. The translation was evaluated by the authors.

For this report “Grammarly” was used for grammar revision. The revisions were evaluated by the authors.

6. Literature

Aguirre-Urreta, M. I., & Hu, J. (2019). Detecting Common Method Bias : Performance of the Harman’s Single-Factor Test. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *50*(2), 45-70. https://doi.org/10.1145/3330472.3330477

APA. (2022). *DSM-5-TR*. https://www.psychiatry.org:443/psychiatrists/practice/dsm

Bornstein, M. H., Putnick, D. L., Lansford, J. E., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Maria Uribe Tirado, L., Zelli, A., Peña Alampay, L., Al-Hassan, S. M., Bacchini, D., Silvia Bombi, A., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Malone, P. S., & Oburu, P. (2015). Mother and Father Socially Desirable Responding in Nine Countries : Two Kinds of Agreement and Relations to Parenting Self-Reports. *International journal of psychology : Journal international de psychologie*, *50*(3), 174-185. https://doi.org/10.1002/ijop.12084

Bowman, N. A., Hill, P. L., Denson, N., & Bronkema, R. (2015). Keep on Truckin’ or Stay the Course? Exploring Grit Dimensions as Differential Predictors of Educational Achievement, Satisfaction, and Intentions. *Social Psychological and Personality Science*, *6*(6), 639-645. https://doi.org/10.1177/1948550615574300

Bravata, D. M., Watts, S. A., Keefer, A. L., Madhusudhan, D. K., Taylor, K. T., Clark, D. M., Nelson, R. S., Cokley, K. O., & Hagg, H. K. (2020). Prevalence, Predictors, and Treatment of Impostor Syndrome : A Systematic Review. *Journal of General Internal Medicine*, *35*(4), 1252-1275. https://doi.org/10.1007/s11606-019-05364-1

Clance, P. R., & Imes, S. A. (1978). The imposter phenomenon in high achieving women : Dynamics and therapeutic intervention. *Psychotherapy: Theory, Research & Practice*, *15*(3), 241-247. https://doi.org/10.1037/h0086006

Clance, P. R., & OToole, M. A. (1988). The Imposter Phenomenon : An Internal Barrier to Empowerment and Achievement. *Women & Therapy*, *6*(3), 51-64. https://doi.org/10.1300/J015V06N03_05

Clark, K. N., & Malecki, C. K. (2019). Academic Grit Scale : Psychometric properties and associations with achievement and life satisfaction. *Journal of School Psychology*, *72*, 49-66. https://doi.org/10.1016/j.jsp.2018.12.001

Coughlin, S. S. (1990). Recall bias in epidemiologic studies. *Journal of Clinical Epidemiology*, *43*(1), 87-91. https://doi.org/10.1016/0895-4356(90)90060-3

Datu, J. A. D., Yuen, M., & Chen, G. (2017). Development and validation of the *Triarchic Model of Grit Scale* (TMGS) : Evidence from Filipino undergraduate students. *Personality and Individual Differences*, *114*, 198-205. https://doi.org/10.1016/j.paid.2017.04.012

Deci, E. L., & Ryan, R. M. (2000). Self-Determination Theory. In *Handbook of Theories of Social Psychology* (p. 416-437).

Du, W., Li, Z., Xu, Y., & Chen, C. (2023). The Effect of Parental Autonomy Support on Grit : The Mediating Role of Basic Psychological Needs and the Moderating Role of Achievement Motivation. *Psychology Research and Behavior Management*, *16*, 939-948. https://doi.org/10.2147/PRBM.S401667

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

Dumont, H., Trautwein, U., Nagy, G., & Nagengast, B. (2014). Quality of parental homework involvement : Predictors and reciprocal relations with academic functioning in the reading domain. *Journal of Educational Psychology*, *106*(1), 144-161. https://doi.org/10.1037/a0034100

Epskamp, S. (2013). *semPlot : Path Diagrams and Visual Analysis of Various SEM Packages' Output* (p. 1.1.7) [Jeu de données]. https://doi.org/10.32614/CRAN.package.semPlot

Feenstra, S., Begeny, C. T., Ryan, M. K., Rink, F. A., Stoker, J. I., & Jordan, J. (2020). Contextualizing the Impostor "Syndrome". *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.575024

Fernández-Martín, F. D., Arco-Tirado, J. L., Mitrea, E.-C., & Littvay, L. (2023). The role of parenting behavior's on the intergenerational covariation of grit. *Current Psychology*, *42*(24), 20872-20882. https://doi.org/10.1007/s12144-022-03185-w

Fruchart, E., & Rulence-Pâques, P. (2025). Development and validation of the Sports Grit Scale. *International Journal of Sport and Exercise Psychology*, *23*(4), 658-671. https://doi.org/10.1080/1612197X.2024.2342029

Glatz, T., Lippold, M., Chung, G., & Jensen, T. M. (2024). A Systematic Review of Parental Self-efficacy Among Parents of School-Age Children and Adolescents. *Adolescent Research Review*, *9*(1), 75-91. https://doi.org/10.1007/s40894-023-00216-w

Grolnick, W. S. (2015). Mothers' motivation for involvement in their children's schooling : Mechanisms and outcomes. *Motivation and Emotion*, *39*(1), 63-73. https://doi.org/10.1007/s11031-014-9423-4

Guelmami, N., Chalghaf, N., Tannoubi, A., Puce, L., Azaiez, F., & Bragazzi, N. L. (2022). Initial Development and Psychometric Evidence of Physical Education Grit Scale (PE-Grit). *Frontiers in Public Health*, *10*. https://doi.org/10.3389/fpubh.2022.818749

Hirschfeld, M. M. (1982). The Impostor Phenomenon in Successful Career Women. *ETD Collection for Fordham University*, 1-173.

Huecker, M. R., Shreffler, J., McKeny, P. T., & Davis, D. (2023). Imposter Phenomenon. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK585058/

Kelley, K. (2006). *MBESS : The MBESS R Package* (p. 4.9.41) [Jeu de données]. https://doi.org/10.32614/CRAN.package.MBESS

Lam, K. K. L., Zhao, S., Du, H., Huang, L., & Chi, P. (2025). Longitudinal Impact of Parent-

ing on Grit : Serial Mediating Effect of Gratitude and Life Satisfaction. *International Journal of Psychology*, *60*(4), e70074. https://doi.org/10.1002/ijop.70074

Langford, J., & Clance, P. R. (1993). The imposter phenomenon : Recent research findings regarding dynamics, personality and family patterns and their implications for treatment. *Psychotherapy: Theory, Research, Practice, Training*, *30*(3), 495-501. https://doi.org/10.1037/0033-3204.30.3.495

Lerner, R. E., Grolnick, W. S., Caruso, A. J., & Levitt, M. R. (2022). Parental involvement and children's academics : The roles of autonomy support and parents' motivation for involvement. *Contemporary Educational Psychology*, *68*, 102039. https://doi.org/10.1016/j.cedpsych.2021.102039

Lucas, G. M., Gratch, J., Cheng, L., & Marsella, S. (2015). When the going gets tough : Grit predicts costly perseverance. *Journal of Research in Personality*, *59*, 15-22. https://doi.org/10.1016/j.jrp.2015.08.004

Mak, K. K. L., Kleitman, S., & Abbott, M. J. (2019). Impostor Phenomenon Measurement Scales : A Systematic Review. *Frontiers in Psychology*, *10*, 671. https://doi.org/10.3389/fpsyg.2019.00671

Mukti, D. A., Grafiyana, G. A., Na'imah, T., & Wijaya, D. A. P. (2024). *Parental autonomy support and the impostor phenomenon : The mediation role of self-esteem in psychology students*. https://doi.org/10.5281/ZENODO.13881328

Nelson, S. K., Kushlev, K., & Lyubomirsky, S. (2014). The pains and pleasures of parenting : When, why, and how is parenthood associated with more or less well-being? *Psychological Bulletin*, *140*(3), 846-895. https://doi.org/10.1037/a0035444

Pardini, D. A. (2008). Novel Insights into Longstanding Theories of Bidirectional Parent–Child Influences : Introduction to the Special Section. *Journal of Abnormal Child Psychology*, *36*(5), 627-631. https://doi.org/10.1007/s10802-008-9231-y

Ping, Y., Wang, W., Li, Y., & Li, Y. (2023). Fathers' parenting stress, parenting styles and children's problem behavior : The mediating role of parental burnout. *Current Psychology*, *42*(29), 25683-25695. https://doi.org/10.1007/s12144-022-03667-x

Revelle, W. (2007). *psych : Procedures for Psychological, Psychometric, and Personality Research* (p. 2.5.6) [Jeu de données]. https://doi.org/10.32614/CRAN.package.psych

Rosseel, Y. (2012). **lavaan** : An *R* Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Sakulku, J., & Alexander, J. (2011). The Impostor Phenomenon. *International Journal of Behavioral Science*, *6*(1), 73-92.

Suzuki, Y., Tamesue, D., Asahi, K., & Ishikawa, Y. (2015). Grit and work engagement : A cross-sectional study. *PLoS ONE*, *10*(9). https://doi.org/10.1371/journal.pone.0137501

Teuber, Z., Datu, J. A., Botes, E., Dicke, T., Jordan, G., Lan, X., Iliescu, D., Guo, J., & Greiff, S. (2024a). *Gritty Parenting : The Development and Validation of the Parental Grit Scale*. https://doi.org/10.31234/osf.io/gnyex

Teuber, Z., Datu, J. A. D., Botes, E., Dicke, T., Jordan, G., Lan, X., Iliescu, D., Guo, J., & Greiff, S. (2024b). Gritty Parenting : The Development and Validation of the Parental Grit Scale. *Assessment*, *32*(7), 1120-1141. https://doi.org/10.1177/10731911241289242

Wolters, C. A., & Hussain, M. (2015). Investigating grit and its relations with college students' self-regulated learning and academic achievement. *Metacognition and Learning*, *10*(3), 293-311. https://doi.org/10.1007/s11409-014-9128-9

World Health Organization. (2022). *ICD-11*. https://icd.who.int/en/

7. Acknowledgments

The authors express their sincere gratitude towards their supervisors Dr. Ziwen Teuber and Dr. Anke Weber for their support throughout the whole process of this research project. This work was funded by the University of Luxembourg, to which we are very grateful.

How does home language influence number processing?

Adnan Akram, Madeline Lehnertz, Alexandra Somesan, Aimée Thein, Leandra Vogel

Supervisor: Dr. Angélique Volfart

Numerical cognition is closely related to language as number processing relies on verbal, semantic, and structural properties of number words. Previous research has shown that the language of mathematics instruction as well as differences in number word structure can influence arithmetic and transcoding performance, particularly in bilinguals. In multi-lingual contexts such as Luxembourg, individuals grow up with different home languages and experience a switch in the language of mathematics instruction from German to French during schooling. The present study investigated how home language affects number transcoding performance in these two instructional languages. An online verbal-visual number matching task was administered in German and French. Participants were required to match spoken two-digit number words to Arabic numerals under three experimental conditions: Unit first, Decade first and simultaneous. The sample consisted of 140 young adults aged 18 to 25 who had spent at least ten years in the Luxembour-gish public school system. Five home language groups were included: Luxembourgish, German, French, Portuguese and South Slavic. Correct response times and error rates were analyzed using repeated measures ANOVAs. The results revealed significant in-teractions between task language and home language indicating that performance dif-fered across groups depending on the language used. In addition, presentation format influenced performance with faster and more accurate responses when the order of digit presentation matched the verbal structure of the task language. No overall main effect of home language was found but group specific patterns emerged in both response times and error rates. These findings suggest that number transcoding is shaped by both lin-guistic background and language specific number word structures. The results highlight the importance of considering language related factors in numerical processing and may help to better understand individual differences in mathematics performance in multilin-gual educational settings.

Introduction

What is numerical cognition?

Numerical cognition describes the mental and neural processes that help people understand and use numbers in everyday life (Knops, 2019). It includes core abilities such as estimating quantities, recognizing symbolic numbers, and learning basic arithmetic facts and procedures. These skills begin to develop early in childhood and depend on general cognitive functions like attention and memory (Knops, 2019). Research shows that language plays an important role in numerical cognition and in some cases, it can make number learning easier, while in others it can make it more difficult (Bahnmueller et al., 2021).

How do linguistic factors affect numerical cognition?

Language plays an important role in how people learn and process numbers. Research shows that bilinguals perform better on complex additions in the language in which they first learned mathematics (Van Rinsveld et al., 2017). Besides the language of instruction, other components of language, such as lexical, syntactic, phonological, visuo-spatial orthographic and semantic features, influence how people learn, perceive, represent, and use numerical and mathematical concepts (Bahnmueller et al., 2021). Language can influence numerical cognition in many ways, because differences in language exposure, number-word structure, and furthermore abilities like phonological skills, grammar, vocabulary, and verbal reasoning affect how children learn, understand and use basic or advanced mathematical concepts (Schiltz et al., 2024).

The effect of language of math instruction

Research shows that the language in which students first learn mathematics has an important impact on their arithmetic performance. In German–French bilinguals for example arithmetic, especially simple additions are usually solved more easily in the first language of math learning (LM1) even when students later become very good in the second instructional language (Van Rinsveld et al., 2015). This study also found that participants were slower and made more mistakes in complex additions when these were presented in their second language of instruction. This shows that early math learning creates strong links between arithmetic and the language used to learn it. Previous research suggests that mathematical knowledge can be "closely tied to the specific language of instruction and/or learning," which leads to cognitive costs when the language of learning and the language of retrieval differ (Saalbach et al., 2013, p. 37). Accordingly, Hahn (2019) shows that solving mathematical tasks in a language different from the language of instruction can lead to cognitive costs, known as language-switching costs. Furthermore, brain imaging studies show, that highly proficient bilingual adults show different brain activity depending on whether they solve additions in LM1 or LM2 (Van Rinsveld et al., 2017). Lachelin et al. (2022) showed that transcoding numbers was harder in LM2 and participants made more errors in the reading of numbers in LM2. In summary, these findings demonstrate that the language used to teach math has long-lasting effects on how bilinguals work with numbers.

The effect of home language

In Luxembourg pupils' home language, which is often linked to migration background is an important source of inequality in school (Lenz & Schiltz, 2025). About two-thirds of pupils are taught in a school language that is different from the language they use at home which makes it hard for many of them to deal with the high language demands of the Luxembourgish system. They also show weak skills in both Luxembourgish and German (Lenz & Schiltz, 2025). Large-scale data from PISA and ÉpStan, „Épreuves standardisés", Luxembourg's national school monitoring program, show that these L2 students who do not speak the language of instruction at home score low-

er in mathematics than L1 students, even when social background is taken into account (Lenz & Schiltz, 2025). Using ÉpStan data for Grade 3, Greisen et al. found that German reading comprehension can predict mathematics scores and that lower math's performance of pupils whose home language is more distant from German is explained through their weaker German reading skills (Greisen et al., 2021). These studies show that when the language spoken at home matches the language used in school, pupils have better chances in education.

The effect of each language's verbal number word structure

Languages differ in their representations of two-digit numbers, and this can influence how people process numbers. The German "inversion" structure where numbers like 23 are said as "three-and-twenty" can slow people down and produce more difficulties when converting number words (Lachelin et al., 2022). Recent work shows that inverted number-word systems are linked to poorer performance in mathematics. When tens and units are named in reverse order, place value is less clear and more information must be held in mind, which affects tasks such as transcoding, magnitude comparison, and arithmetic (Woo & Schiller, 2025). Furthermore, number-word structure has lasting effects on the processing of two-digit numbers (Steiner et al., 2021). Overall, the way number words are built in each language plays an important role in how numbers are understood and processed.

The context of multilingualism in Luxembourg

Luxembourg is a multilingual country, and this is clearly visible in its school system, including the way mathematics is taught. In Luxembourg's traditional school system, children grow up with several languages at school. In Cycle 1 (preschool), the main language used in class is Luxembourgish. From Cycle 2 to 4 (1st grade to 6th grade), children learn to read and write in German which becomes the main language for most subjects including mathematics. French is first taught as a spoken language in 2nd grade and reading and writing in French begins in 3rd grade. In secondary school, German continues to be the teaching language for most subjects at first except for mathematics which is taught in French from the 7th grade on. Later, English is added as a foreign language, and some students can also learn other languages like Spanish or Italian (Ministry of Education, 2025). This system means that students often switch the language used in math class during their school years and they also speak different languages at home. This creates many different language backgrounds, which is important for our study. Furthermore, Luxembourg is highly linguistically diverse: only 48.9% of residents report Luxembourgish as their main language. Portuguese (15.4%) and French (14.9%) are the next most common main languages, while another 20.8% report another language as their main language, as the census identifies over 50 additional language communities (STATEC, 2023). This diversity explains why our study includes many different home language groups.

This study

Through this study we aim to examine how home language influences performance in a number transcoding task, where participants match auditory number words to their Arabic digit format. The task is presented either in LM1 German or LM2 French in different home language groups: Luxembourgish, German, French, Portuguese and South Slavic. In this study, the independent variables were experimental condition (ten-first vs. unit-first vs simultaneous), home language (as listed above), and task language (German = LM1 vs. French = LM2). The goal was to examine not only the main effects of experimental condition, home language, and task language, but also to investigate whether there is a potential interaction effect between these factors.

Hypotheses

To examine the influence of home language on number processing we established several hypotheses concerning the performance of the difference home language groups on the three experimental conditions and the two task languages.

Hypotheses concerning the task languages

H1: Participants of the Luxembourgish home language group will perform better/faster in the German task language than in the French task language (LM1>LM2).
H2: Participants of the German home language group will perform better/faster in the German task language than in the French task language (LM1>LM2) due to the home language being consistent with the LM1.
H3: Participants of the French home language group will perform better/faster in the French task language than in the German task language (LM1<LM2) due to the home language being consistent with the LM2.
H4: Participants of the Portuguese home language group will perform better/faster in the French task language than in the German task language (LM1<LM2) due to the home language being linguistically closer to French than German (both follow the same decade-unit verbal structure).
H5: Participants of the South Slavic home language group will perform better/faster in the German task language than in the French task language (LM1>LM2). Because their home language is neither linguistically close to German nor French, the effect of the LM1 should be predominate in this group.

Hypothesis concerning the experimental conditions

H6: Participants will perform better in the unit-first condition than in the decade-first condition when the task language is German, due to the unit-first format matching the verbal structure of German.
H7: Participants will perform better in the decade-first condition than in the unit-first condition when the task language is French, due to the decade-first format matching the verbal structure of French.

Methods

Participants

To examine the relationship between an individual's home language and their understanding of numbers within Luxembourg's multilingual school context, this study focused on participants with the following home languages: Luxembourgish, German, French, Portuguese and South Slavic languages. The goal was to recruit 100-150 participants, with approximately 20-30 individuals per home language group. These datasets consist of newly collected and already existing datasets (collected under equivalent conditions). A previous study (Lachelin, 2024) used the same experimental task as our study, but with a different research question: They aimed to compare Luxembourgish bilinguals with French or German monolinguals. Providing a Statistical Power Analysis, they calculated that a sample size of about 40 participants per group would be appropriate for their purpose. Considering our limited timeframe for recruiting in this project, we aimed at recruiting 20-30 participants per group instead.
To fulfil the requirements for this study, the participants needed to be between 18 and 25 years old and have spent at least 10 years in the Luxembourgish public school system. These requirements need to be fulfilled because our main interest is how home language may interact with languages of math instruction, which in the Luxembourgish public school system involves a switch of math instruction language from German (LM1) to French (LM2) after 6th grade. Selecting participants who were schooled in the Luxembourgish system for a sufficient duration thus ensured that all participants experienced this switch in language of math instruction.
The participants were recruited through posters (digital and print), social media, the university mailing portal and word-of-mouth.

All participants received detailed information about the study and provided informed consent. They were rewarded with a 15€ voucher afterwards.
After data cleaning (removing incomplete datasets and those not fitting eligibility criteria, trimming RT +/- 3 individual standard deviations from each individual mean), the final sample consisted of N=140 bilingual participants whose demographic data can be found in Table 1. Further, N=73304 trials were included in the analysis.

Table 1
Demographic data

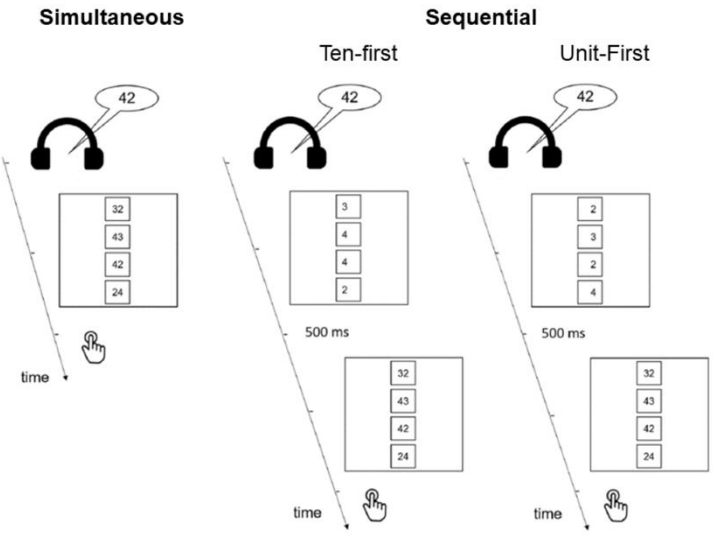
Sample size	Gender ratio (M/F)	Mean age (SD)	Mean TTR (SD)	Mean N spoken languages (SD)
40	15/25	21.98 (1.81)	112.18 (16.07)	4.08 (0.47)
18	7/11	21.61 (1.95)	112.89 (21.25)	4.06 (0.52)
20	10/10	21.30 (2.10)	114.20 (20.82)	3.80 (1.08)
46	19/27	21.91 (1.91)	107.09 (20.39)	5.00 (0.00)
16	5/11	21.69 (1.93)	96.19 (13.74)	5.00 (0.00)

Note. HL = home language; M/F = Male/Female; SD = standard deviation; TTR = Tempo Test Rekenen; LUX = Luxembourgish; GE = German; FR = French; POR = Portuguese; SLA = South Slavic

We found no significant differences between the home language groups, except for the two following: Firstly, when analyzing the influence of home language on the TTR score, Tukey post-hoc tests revealed that the South Slavic group had a significantly lower score than both the French and the Luxembourgish group ($p<.05$). It is important to consider that this could be due to the small sample size for the South Slavic group that didn't meet our recruitment goal and could have influenced our statistical power which was also taken into account for our discussion.
Moreover, the Portuguese and the South Slavic group showed a significant higher mean number of spoken languages across all groups ($p<.05$).

Stimuli and Procedure

The experiment was conducted using Labvanced, an online experiment platform. For the single online session that participants completed on their own device, we asked them to complete the session in a quiet environment to minimize distraction.
Firstly, the participants had to fill out a demographic and linguistic questionnaire to collect information and exclude participants outside the target population. To determine their home language, participants were asked in the questionnaire which language they learned first.
Moreover, we asked them about how many languages they speak, how often they speak those languages and finally how the participants rank the fluency level of these languages. Additional demographic information about the participants that was collected was age, gender and declarations concerning any neurodevelopmental difficulties for example dyslexia, ADHD or dyscalculia. This was nec-



essary, to exclude individuals that did not align with the frame of the study and which may have affected the results with confounding variables.
The main experimental task was a verbal-visual number matching task which was used in previous studies (e.g. Lachelin et al., 2022) to investigate the influence of the language of mathematical instruction on number processing (see figure 1).
Figure 1: Illustration of the auditory-visual number matching task, in which numbers were presented orally in German or in French to the participants.

In the task, which was completed in two languages (German and French), participants heard a two-digit number which was the target stimulus. Afterwards, they saw four Arabic numbers on their screen, from which they had to pick the heard target number. The three other numbers were distractors. One of the three distractors was the reversed version of the target number, one had the same decade as the target but the unit ±1 and the third number had the same unit as the target, but the decade is ±1. For instance, if the target was 35, the distractors would be 53 (reversed version), 25/45 (decade ±1) and 34/36 (unit ±1). The stimuli were restricted to 2-digit numbers between 23 and 69, excluding ties, round tens, and numbers with 1 or 9 as unit digits. Numbers above 69 were excluded as well because they follow a different phonological structure in French as in German which would cause longer reaction times (Lachelin et al., 2022). We counterbalanced if the participant started with the German or French task to avoid order effects, learning effects and fatigue, whereas the order of conditions, trials and stimuli (placement of target and distractor stimuli on the screen) was randomized.
In the simultaneous condition the units and tens of the target number and the distractors were presented at the same time. In the ten-first condition, the tens of the target and the three distractors appeared, which mimics the verbal structure of number words in French in which the decade is named before the unit (e.g. "trente-deux"). After 500ms, the units appeared. In the unit-first condition, the units of the target and the three distractors were shown first, which corresponds to the verbal structure of number words in German in which the unit is named before the decade (e.g. "zweiunddreißig"). Here, the tens then appeared also after 500ms.
In each condition, the stimuli remained on the screen until the participant provided their response. The dependent variables were correct response times in ms and error rate. The response time is measured either from the first appearance of the last digit (for the sequential conditions) or from the entire number (for the simultaneous condition) until the response is submitted. Reaction times in the sequential and simultaneous conditions are not directly comparable, as participants in the simultaneous condition must process both digits concurrently, whereas in the sequential condition they have already processed the first digit when the second one appears.
After the main task, participants completed the Tempo-Test-Rekenen (TTR) - a short arithmetic fluency test. It contains 40 additions/subtractions/multiplications/divisions from which the participants had to solve as many as possible in one minute. This was used as a control variable to ensure that the home language groups did not differ in their general mathematical abilities.
Overall, the experiment did not extend the length of approximately forty minutes.

Results

Data Preparation and Procedure

Correct response Time (RT in milliseconds) data and accuracy data (error rates in percent) were analysed using a repeated-measure ANOVA. All analyses were conducted in SPSS using a 2(Task Language: German=LM1, French = LM2) and 3(Condition: Decade-first, Unit-first and simultaneous) within-subjects design, with home language group (Luxembourgish, German, French, Portuguese and South Slavic) as a between-subjects factor.

Response Time

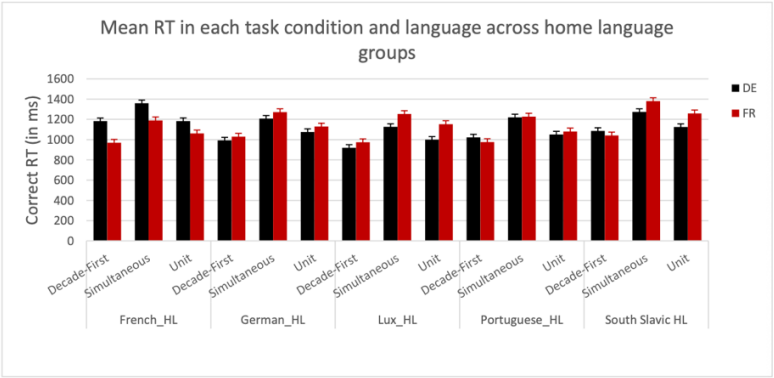


Figure 2: Mean response times (ms) across task conditions and all home language groups for German (DE) and French (FR) task language

Main effects

TASK LANGUAGE: LM1 VS. LM2. Response time did not differ significantly between the German and French languages *F* (1,135) = 0.73, p = .395, partial η² = .005.

CONDITION. There was a strong and significant effect of Condition, *F* (2,270) =508.059, *p* = .001, partial η² = .790. Post-hoc tests (Bonferroni) showed that reaction times in the Decade-first condition (M = 1019 ± 19 ms) were significantly faster than in the Unit-first condition (M **=** 1112 ±19ms; p < .001) In addition, responses in the Simultaneous condition (M = 1251 ±20ms) were also significantly slower than in the Unit-first condition (p < .001). The comparison between Simultaneous and Decade-first revealed that Decade-first processing was significantly faster than Simultaneous (p < .001). Therefore, participants processed number words most efficiently in the Decade-first condition, followed by the Unit-first condition, while the Simultaneous condition resulted in the slowest responses. A plausible explanation is that the sequential presentation (Decade-first and Unit-first) allows participants to process the first digit immediately and can already start processing the number before the second digit appears which can lead to faster responses. In contrast, in the Simultaneous condition, both digits appear at the same time, so participant need to process both digits at the same time which results in slower reaction times.

HOME LANGUAGE GROUP. The main effect of home language group was not significant, *F* (4,135) = 1.357, *p* = .252, partial η² = .039.

Interaction Effects

TASK LANGUAGE AND HOME LANGUAGE. Interaction between Task language and home language was significant, *F* (4,135) = 13.252, *p* < .001, partial η² = .282. Bonferroni post-hoc tests were conducted to compare performance in LM1 (German) and LM2 (French) within each home language group.
Post-hoc analyses showed that the Luxembourgish group showed a significant difference between LM1 and LM2 (p <.001). Participants were faster in German than in French as expected in *H1*.
For the German speakers, there was no significant difference found between LM1 and LM2 (*p* = .134). Therefore, the hypothesis *H2* that German participants perform better in LM1(German) is not supported.
The French home language group responded significantly faster in LM2 in than in LM1 (*p <.001).* This result confirms the hypothesis *H3* that predicted that French participants would perform better in their home language.
The Portuguese group showed no significant difference between LM1 and LM2 (*p* =.880). This result means that the home language of Portuguese speakers did not significantly influence processing speed in the two LMs.
For the South Slavic group, the comparison between LM1 and LM2 did not reach significance (*p* = .075). However, the estimated marginal indicated numerically faster RTs in LM1(M = 1161.68 ms) than in LM2(M = 1227.07 ms) although the effect was not statistically significant. It is important to note that the South Slavic group had fewer participants than the other home language groups, which may have reduced statistical power and contribute to the lack of significance. Therefore, the hypothesis *H5* that the South Slavic group would respond faster in German was not supported.

CONDITION AND HOME LANGUAGE. This interaction was significant, F (8,270) = 3.201, p = .002, partial η^2 = .087. This result indicates that strength of the condition effects is different between the home language groups.
Post-hoc comparisons showed that all home language groups responded fastest in the Decade-first condition with shorter response time than Unit-first condition and Simultaneous condition (p < .001).

TASK LANGUAGE AND CONDITION. The interaction between Task language and Condition reaches significance F (2,270) =23.950, p < .001, partial η^2 = .151, indicating that the difference between LM1 (German) and LM2 (French) is different across the three conditions. The post-hoc tests revealed that, in the Unit-first condition, participants responded significantly faster in LM1 than in LM2 (1087 vs 1137; p= .001) which supports *H6*. In the decade first condition, participants responded slower in German than in French (1040 vs 998; p = .007) which supports *H7*. For the simultaneous condition, there was no significant difference between German and French (1237 vs 1264; p = .119).

TASK LANGUAGE X CONDITION X HOME LANGUAGE. The three-way interaction of Task language, Condition and home language was not significant, F (8,270) = 1.611, p = .122, partial η^2 = .0415.

Error Rates

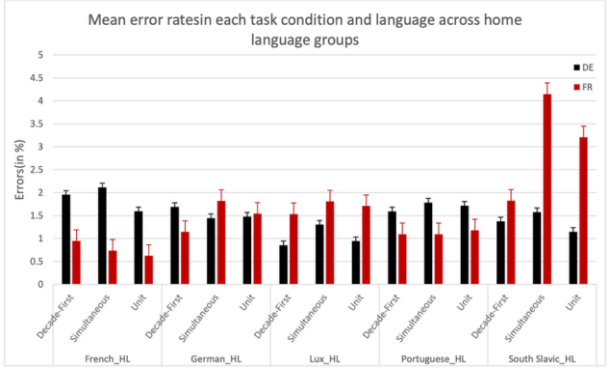


Figure 3: Mean error scores across task conditions (decade-first, simultaneous, and unit-first) and home language groups for German (DE) and French (FR) task languages.

Main effects:

There was no significant main effect for the Task Language, F (1,135) = 0.42, p = .518, η^2 = .003.
However, the Condition significantly affected error rates, F (2,270) = 4.661, p = .010, η^2 = .033. Post-hoc tests showed that participants made less errors in the decade-first condition than in the simultaneous condition(p=.020), although the comparison between
Simultaneous and Unit-first condition was not significant (p = .159), as well as the comparison between Unit-first and Decade-first condition (p = .809).
There was no significant main effect of home language, F (4,135) = 1.42, p = .231, partial η^2 = 0.40.

Interaction effects

TASK LANGUAGE AND HOME LANGUAGE. The interaction of Task language and home language group was significant, F (4,135) = 7.07, p < .001, partial η^2 = .173.
For the Luxembourgish group, participants also made fewer number of errors in LM1 than in LM2 (p = .035). This result confirms the hypothesis that Luxembourgish speakers would perform better in LM1 than LM2 and it is consistent with the RT result, which also showed better performance in LM1.
For the German home language group, participants made a similar number of errors between LM1 and LM2 (p = .938), this result is compatible with the RT results but the hypothesis (*H2*) that the German speakers perform better in their home language is not supported.
The post-hoc tests revealed that the French speakers made significantly fewer errors in LM2(French) than in LM1(German) (p = .011), which confirms the hypothesis (*H3*) that French participants perform better in their home language.
The participants in the Portuguese HL group made more errors in the LM1 (German) than LM2 (French), which differed from their ab-

sence of response between the two LMs in RT. Hence, our hypothesis (*H4*) was partly confirmed.
The South Slavic speakers made significantly fewer errors in LM1 than in LM2 ($p<.001$) which strongly supports the hypothesis (*H5*) that the South Slavic group would perform better in German.

CONDITION AND TASK LANGUAGE. The interaction between Task Language and Condition was not significant, F (2,270) = 2.41, p = .092, partial η^2 = .018.

HOME LANGUAGE AND CONDITION. The interaction between Condition and Home Language Group did not reach significance, F (8,270) = 1,32, p = .233, partial η^2= .038.

CONDITION X HOME LANGUAGE X TASK LANGUAGE. The three-way interaction of Condition x home language x Task language was not significant, F (8,270) = 1.45, p = .177, partial η^2 = .041.

Discussion

The main goal of this study was to investigate the influence of home language on number transcoding performance in the two languages of math instruction (LM1 German and LM2 French) in Luxembourgish multilinguals. Our results found significant interaction effects between the task language and the home language group, suggesting different patterns of performance across groups. We also found a significant interaction between condition and task languages, suggesting that the performance can be facilitated if the verbal structure of number words is consistent with the order of appearance of decade or units first.

Interaction of task language and home language

As Bermejo et al. (2021) have shown, the influence of the task language on mathematical processes is evidenced. Although their paper indicates that age and task type have an influence, the performance in mathematics is overall still lower when language of task instruction and mother tongue (comparable to our definition of home language) are not matched. Following, it seems interesting to look into our findings:
As predicted by our hypothesis (*H3*), we found that French speakers were slower in LM1 German than LM2 French, both for response time and for error rates. This aligns with Lachelin's (2022) findings, that the difference in the German and French language structure influences transcoding performance.
Contradictory to our hypothesis (*H2*), we found that participants in the German home language group performed equally in the LM2 French task condition and in the LM1 German task condition, for error rates as well as for response time.
Regarding the Luxembourgish group we found a significant difference between the LM1 German and the LM2 French condition for both response times and error rates. That confirms our hypothesis (*H1*) that, they would perform better in LM1 German. Our findings are consistent with Lachelin's (2024) results, who only focused on one home language group (Luxembourgish). Furthermore, Van Rinsveld et al. (2017) indicate similar results, as they could prove that Luxembourgish participants required more time and made more errors for complex mathematic additions in French compared to German.
In participants speaking Portuguese as home language, we only found a better performance in French compared to German for error rates, whereas in response time, no difference between LM1 German and LM2 French was found. Following, our hypothesis (*H4*) can be confirmed, as we thought they would perform better in the French task.
Regarding our hypothesis (*H5*) that South Slavic participants would perform better in the German task: we found that for error rates the performance in LM1 German was indeed better than in LM2 French. The results on response time did not meet statistical significance. Following, we can partly confirm our hypothesis. Future investigations in this case would be possible.

When we compare our findings to the study of Greisen et al. (2021), we see similar results: Luxembourgish speaking pupils (third graders) had an advantage in comparison to French, Portuguese or South Slavic home language groups in a mathematic task written in German (LM1). Only in comparison to the South Slavic group, our findings are more differentiated, as their error rates performance was better in LM1 German than in LM2 French. Nevertheless, our findings also indicate that task language could have an impact on mathematic performance and achievement.
Following, our findings can be used as guidelines for educational implications. As French speakers are slower in processing mathematic tasks in German, the Luxembourgish LM1 system can be a disadvantage to them, because they are first confronted with their „slower" language, German. They might need longer to understand certain tasks in mathematics and thus may have more
problems to follow the lessons. Knowing this can be an essential key, as personal support by the teachers can be provided specifically for French speakers in the LM1 system. Concerning the German home language group, our findings indicate that they do not need special care in LM1 or LM2, as they perform the same in German and French. The Luxembourgish home language group may receive extra care when transitioning from the LM1 to the LM2 system, as they performed overall better in the German task than in French and for this reason might need help with the new language setting. Concerning the educational implications for the Portuguese speakers, it would be plausible to provide additional help in the LM1 German math's lessons, as participants from the Portuguese home language group had higher error rates in the German tasks than in the French ones.
Similarly, the South Slavic speakers could also be supported in the LM1 system, as they made more errors in German than in French.

Interaction of task language and condition

Concerning the Unit-first condition, participants responded faster in German than in French, which is in line with our hypothesis (*H6*).
Following, the analysis of the Decade-first condition showed that participants responded slower in German than in French. Similarly, in a slightly different research design, Van Rinsveld et al. (2015) found that participants systematically produced more errors on the "ten digit" (compared to the Decade-first condition), when they calculated in German and more errors on the "unit digit" (compared to the Unit-first condition), when they calculated in French. Additionally, Klein et al. (2013) highlighted the strong differences between number word systems. In particular, their findings indicate that place value integration of numbers seem harder for inverted number words, like in German. Those findings are quite interesting, because they indicate that seeing the Unit-first is an advantage when performing the task in German, which fits with German number words being pronounced with the Unit-first.
The opposite pattern occurs in French, with better performance in French for the Decade-first condition, which also fits with the linguistic structure of French that pronounces the decade first.
Overall, it means that the linguistic structure of the language has an impact on how the auditory number word is processed and that showing either the decade or the unit alone as a prime, before showing the full two-digit number can facilitate the transcoding performance in French or in German, respectively.

Limitations

RESEARCH DESIGN. The present study investigates how home language, task language, and presentation format influence number processing performance.
When reflecting on the research design, several factors need to be considered that may had an impact on our results, with home language as one of them. To determine the home language of the
participants, they were asked in the questionnaire which language they have learned first. Although this procedure may appear as the best way to identify someone's home lan-

guage (often corresponding to the native language/mother tongue), it is important to acknowledge that this assumption may not always hold. In bilingual or multilingual families, the initially acquired language may later be overshadowed by another due to internal factors (e.g., increased interaction with the parent who speaks the later-learned language) or external factors (e.g., changes in kindergarten, peer group language, relocation). These considerations should be taken into account when interpreting our findings.
Furthermore, reaction times in the sequential and simultaneous conditions are not directly comparable, as participants in the simultaneous condition must process both digits concurrently,
whereas in the sequential condition they have already processed the first digit when the second one
appears.

ONLINE EXPERIMENT. A complete internal validity is not given as the experiment was conducted online. In this design it is nearly impossible to control if the participants carried out the experiment with environmental distractions, such as auditory (e.g. noises from roommates), physical (e.g. personal belongings) or general discomfort due to exposure to heat or cold. Not being able to control sources of error variance may reduce the power of online experiments, as also stated by Lacherez (2008). In a laboratory, more equal conditions can be ensured. However, the advantage of doing the experiment online is that the participants can conduct the study at their normal environment and thus the performances could be more accurate, rather than in the arbitrary laboratory setting (Clifford & Jerit, 2014).
Nevertheless, the possibility of extraneous variables should be considered.

SOUTH SLAVIC HOME LANGUAGE. There was a trend for a response time difference between German and French in the South Slavic group that did not reach significance, which is likely to be due to the fewer number of participants in this group that may have led to reduced statistical power.
In addition, this group was also significantly different from other home language groups on the TTR scores, which calls for further study with more participants in this group to better understand their different profile of performance.

Potential Advancements for the Future

As explained above, the choice of defining home language as the language first learned can be questioned. Considering that many other linguistic variables were collected for the study, it would be interesting to see how the definition of a participant's home language (i.e., rated by the participant as its most fluent language, or most frequently spoken language) may potentially affect the pattern of performance in our number transcoding task.
Another possible way to conduct further investigations, would be to have a look at home language groups from other language backgrounds, such as English or Italian as those language groups
have a rising impact in Luxembourg.

References

Bahnmueller, J., Lopes-Silva, J., Haase, V. G., Moura, R., & Moeller, K. (2021). *Ties of math and language: A cognitive developmental perspective* (pp. 101–131). https://doi.org/10.1515/9783110661941-006

Bermejo, V., Ester, P., & Morales, I. (2021). How the Language of Instruction Influences Mathematical Thinking Development in the First Years of Bilingual Schoolers. *Frontiers in Psychology*, *12*, 533141. https://doi.org/10.3389/fpsyg.2021.533141

Census. (2025, August 18). http://statistiques.public.lu/en/recensement.html

Clifford, S., & Jerit, J. (2014). Is There a Cost to Convenience? An Experimental Comparison of Data Quality in Laboratory and Online Studies. *Journal of Experimental Political Science*, *1*(2), 120–131. https://doi.org/10.1017/xps.2014.5

Greisen, M., Georges, C., Hornung, C., Sonnleitner, P., & Schiltz, C. (2021). Learning mathematics with shackles: How lower reading comprehension in the language of mathematics instruction accounts for lower mathematics achievement in speakers of different home languages. *Acta Psychologica*, *221*, 103456. https://doi.org/10.1016/j.actpsy.2021.103456

Hahn, C. (2019). *Language-Switching Costs in Bilingual Mathematics Learning* [Georg-August-University Göttingen]. https://doi.org/10.53846/goediss-7530

Klein, E., Bahnmuellller, J., Mann, A., Pixner, S., Kaufmann, L., Nuerk, H.-C., & Moeller, K. (2013). Language influences on numerical development—Inversion effects on multi-digit number processing. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00480

Knops, A. (2019). *Numerical Cognition: The Basics* (1st ed.). Routledge. https://doi.org/10.4324/9781351124805

Lachelin, R. (2024). *Bilingual lexical and semantic representation of numbers*. https://orbilu.uni.lu/handle/10993/62052

Lachelin, R., Van Rinsveld, A., Poncin, A., & Schiltz, C. (2022). Number transcoding in bilinguals—A transversal developmental study. *PLOS ONE*, *17*(8), e0273391. https://doi.org/10.1371/journal.pone.0273391

Lacherez, P. (2008). The internal validity of web-based studies. *Empirical Musicology Review*, *3*(3), 161–162. https://doi.org/10.18061/1811/34107

Languages in Luxembourg schools. (2024, August 9). http://men.public.lu/en/systeme-educatif/langues-ecole-luxembourgeoise.html

Lenz, T., & Schiltz, C. (2025). *Different schools for different pupils? What are the advantages and problems of Luxembourg's highly differentiated and stratified school system?* https://orbilu.uni.lu/handle/10993/64652

Saalbach, H., Eckstein, D., Andri, N., Hobi, R., & Grabner, R. H. (2013). When language of instruction and language of application differ: Cognitive costs of bilingual mathematics learning. *Learning and Instruction*, *26*, 36–44. https://doi.org/10.1016/j.learninstruc.2013.01.002

Schiltz, C., Lachelin, R., Hilger, V., & Marinova, M. (2024). Thinking about numbers in different tongues: An overview of the influences of multilingualism on numerical and mathematical competencies. *Psychological Research*, *88*(8), 2416–2431. https://doi.org/10.1007/s00426-024-01997-y

Steiner, A. F., Banfi, C., Finke, S., Kemény, F., Clayton, F. J., Göbel, S. M., & Landerl, K. (2021). Twenty-four or four-and-twenty: Language modulates cross-modal matching for multidigit numbers in children and adults. *Journal of Experimental Child Psychology*, *202*, 104970. https://doi.org/10.1016/j.jecp.2020.104970

Van Rinsveld, A., Brunner, M., Landerl, K., Schiltz, C., & Ugen, S. (2015). The relation between language and arithmetic in bilinguals: Insights from different stages of language acquisition. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00265

Van Rinsveld, A., Dricot, L., Guillaume, M., Rossion, B., & Schiltz, C. (2017). Mental arithmetic in the bilingual brain: Language matters. *Neuropsychologia*, *101*, 17–29. https://doi.org/10.1016/j.neuropsychologia.2017.05.009

Woo, V., & Schiller, L. K. (2025). How the inversion property in number-naming impedes mathematical performance, potential interventions to minimize its impact. *Psychological Research*, *89*(5), 147. https://doi.org/10.1007/s00426-025-02177-2

Welche Sprachen spricht man in Luxemburg? (o. J.) Public.lu Abgerufen 14. Juni 2025, von https://luxembourg.public.lu/de/gesellschaft-und-kultur/sprachen/welche-sprachen-luxemburg.html

Zaripova, R. R., Salekhova, L. L., Grigorieva, K. S., &Azrou, N. (2019). Potential influence of bilingualism on the development of abstract thinking. *Journal of Computational and Theoretical Nanoscience*, *16*(11), 4546–4549. https://doi.org/10.1166/jctn.2019.8351

Daily Mood and Audiovisual Behaviour: Ambulatory Assessment with Facial Landmarks and Acoustic Features?

Elias Holzgartner, Carla Haasenritter, Audrey Haeck, Jil Bocquet, Sébastien Gorecki

Supervisor: Prof. Dipl.-Psych. Philip Santangelo, David Tekampe

Ambulatory assessment captures affect in daily life, yet affect–facial–vocal behaviour associations are dominated by cross-sectional work. In a remote, privacy-preserving web-based protocol, we paired self-reports with a prompted recording to extract multimodal affect features in real-time. Adults in Luxembourg (N = 38) completed a 7-day protocol (1/day) providing self-reported valence and tense-arousal, and a 90-second recording comprising a standardised sentence followed by speech. Facial behaviour was quantified via MediaPipe blendshape-derived composites of positive (smile/cheek-squint) and negative expressions (frown/brow-down/nose-sneer). Voice-derived valence and arousal were quantified using the devAIce Multi-modal Expression module (acoustic/linguistic). Between-person analyses used Pearson correlations, hierarchical regression for incremental validity of daily significant-event ratings beyond self-reported affect, and literature-informed moderation models (sleep, personality). Convergent validity was strongest for valence (voice-derived valence with self-reported valence, $r = .60$, $p < .001$). Criterion validity was supported by significant associations between voice features and event ratings ($r = .60$, $p < .001$), whereas facial composites showed smaller but significant links with valence and event ratings ($|r| \approx .36–.45$). Sleep was positively related to valence and moderated the voice–valence association, and neuroticism was negatively related to valence.Preliminary results suggest voice-derived features are more informative than facial composites for augmenting self-reported valence and daily event appraisal. Overall, this pilot supports privacy-preserving multimodal ambulatory assessment; future work should increase sampling density and apply multilevel (within-person) models with dimensionality reduction.

1. Theoretical Backround

This quantitative study examines the objectivity and subjectivity of mood in everyday life. This is highly relevant, as mood also influences our perception, our actions and thinking, and our memory (Dreisbach, 2008). Moreover, if it is possible to objectively measure mood using facial landmarks and voice features, then these programs could also be used, for example, to better evaluate statements in forensic psychology. This highlights the importance of accurate mood measurement for a comprehensive understanding of mood.

2. Introduction

Emotions play a crucial role in our everyday lives and affect our psychological well-being, decision-making and social interaction. In psychology, emotions are commonly separated into valence and arousal, describing how pleasant or unpleasant emotional states are and how intense they are. This framework helps us study emotional experiences along these dimensions.

Traditionally, emotions have been assessed using self-report measures, as they are easy to do and offer direct access to individuals' perceptions of their subjective experiences. However, there are certain limitations to self-reports, including recall bias, social desirability, and limited understanding of one's own emotions (Mauss & Robinson, 2009). These limitations lead researchers to consider objective indicators for the assessment of emotion, such as facial expressions and vocal characteristics, as a complementary source of information.

There is a lot of research supporting that emotional states are reflected in facial behavior and vocal features (Ekman, 1993)(Scherer, 2003). The association between facial expression and emotional processes has long been made, with specific patterns of facial muscle activation linked to different affective states (Ekman, 1993). Similarly, vocal characteristics such as pitch, intensity, and prosody have been shown to vary depending on emotional valence and arousal (Scherer, 2003). Technological advances enabling the automatic analysis of emotional information from facial landmarks and acoustic speech features have made objective emotion assessment increasingly accessible.

Despite these advances, much of the existing evidence on objective emotion measurement relies on cross-sectional or laboratory-based designs. Laboratory studies have certain limitations such as artificial settings, controlled stimuli, and low external validity. In cross-sectional approaches, within-person variability of emotional states is neglected. As emotions fluctuate within individuals, there is a growing consensus that repeated, within-person assessments in real-world contexts are necessary to adequately study emotional experiences (Trull & Ebner-Priemer, 2020).

Ambulatory assessment methods address these limitations by enabling repeated measurements of emotions in participants' natural environments. Using mobile technology, ambulatory assessment allows for the collection of momentary self-reports alongside objective behavioral data, thereby increasing external validity and reducing retrospective bias (Trull & Ebner-Priemer, 2020). This approach is really useful when investigating the correspondence between objective emotional indicators and subjective mood reports as they occur in daily life.

In addition to emotional indicators, the relationship between objective and subjective measures of emotion may be influenced by individual differences. Sleep for example plays an important role in emotional regulation, with insufficient sleep being associated with a more negative mood and altered emotional processing (Walker, 2009). Additionally, personality traits, such as neuroticism, are linked to differences in emotional self-reporting (Costa & McCrae, 1992). These factors may therefore moderate the degree to which objectively measured emotional signals align with individuals' subjective emotional experiences.

The present study aimed to investigate the relationship between objective audiovisual emo-

tion indicators and momentary self-reported mood using a fully remote, ambulatory assessment design. Specifically, we examined whether facial and vocal features correspond with self-reported emotional valence and arousal across repeated daily measurements. Furthermore, we evaluated the validity of these objective measures and explored whether sleep and personality traits moderate the association between objective and subjective emotional assessments.

2.1 Comparing subjective and objective perceptions of mood

As already mentioned, both objective and subjective measurements of mood are taken in the study. This serves to provide as accurate as possible a picture of mood, as it is examined from both sides.

2.1.1 Subjective measurement of mood

Mood refers to an individual's emotional condition and mindset. This shapes the way they perceive their experiences and aspirations and may fluctuate in response to their physical or psychological state (Mackenthun, 2004).
The subjective perception of mood is measured using the six-item short scale Momentary mood (primary outcome) by (Wilhelm & Schoebi, 2007). This has been proven useful in a field study because it is easy and quick to complete. By only using single items, problems with the reliability may occur. For that reason, using short scales is advisable. The study shows that the three factors (calmness, valence and energetic arousal) are well suited to ensuring reliability. Therefore, using the questionnaire is very useful due to the good psychometric properties (Wilhelm & Schoebi, 2007). Direct subjective measurement methods offer the personality of flexible and user-friendly data collection. There are some situations in which subjective measurements are necessary, if the object of research is subjective in nature (e.g., personality) (Weigand, 2018).
However, the subjective measurement can also lead to data falsification. Conscious falsification of data occurs, for example, because of the "fake good" or "fake bad" effect. This means that the test subjects intentionally try to portray themselves positively or negatively in relation to the subject being researched. The falsification does not always have to be conscious, but it is necessary to pay attention to it (Weigand, 2018).
A study by (Johnson, 1981) addressed the topic of self-disclosure. Among other things, the study sought to examine the extent to which self-disclosure captures factual information or merely reveals how a person wants to be seen by society. The results of the study show a slight correlation between self-presentation concepts (median r = 0.21). This means that a self-report tends to represent how a person would like to be seen and is not a true "self-disclosure". However, since only a slight correlation was found, the results are of little significance (Johnson, 1981). Given this situation, it makes sense to conduct both an objective and a subjective measurement to obtain a picture of mood as accurate as possible.

2.1.2 Objective measurement of mood

In order to obtain more accurate data, this study also uses indirect objective measurements, as this type of data collection does not involve any of the aforementioned distortions (Weigand, 2018).
Emotions are the subjective experience of stimuli and influence, for example, the automatic nervous system. The physiological response of the body to emotions makes it possible to measure them objectively (Mackenthun, 2004).
The affect-as-information hypothesis refers to the fact that people can tell how someone feels or what they think about something from their voice or facial expressions. For example, if a teacher notices that a student has a shaky voice during a presentation and concludes that the student is nervous, they are using the voice as information. But the student themself can also hear their voice. Therefore, the person themself knows that they are becoming nervous, and hearing their own shaky voice confirms this once again. So, the student uses their expression as confirmation of their feelings. However, if the student does not feel nervous but still has a shaky voice, it would

lead to discrepancy between an affective belief and an embodied affect (Schwarz, 1987).
This results in poorer memory of the situation and complicated decision-making (requiring more time) and shows that people want proof of what they feel (Clore et al., 2012). The hypothesis therefore confirms the objective measurement approach of the present study, as it assumes that facial expressions and voice are carriers of information about a person's mood.

2.1.2.1 Objective measurement through affective prosody

Even with children as young as three or four years old, people use a special kind of language called "baby talk". This involves a distinctive intonation, a higher pitch, the use of nicknames, and the formation of short and simple sentences. Baby talk not only leads to a better understanding of language by children towards adults, but also conveys affection and care to children (Zebrowitz et al., 1992). Language therefore also serves as a nonverbal expression of emotions, as prosodic features convey a person's mood and state of mind. This is called affective prosody and can thus be used to measure mood (Ackermann et al., 2004).

2.1.2.2 Objective measurement through facial expressions

The phenomenon of emotional communication is already shown in a person's facial expressions early on. An experiment with one-year-old children showed that children can understand the emotional facial expressions (happy, fearful, sad, angry) of other people (in this case, their mother). To do this, children were lured onto a plexiglass panel with a toy and different facial expressions. The plexiglass panel looks like a cliff to the children, as they already have depth perception and can see that there is an "abyss" below the plexiglass panel. When the caregiver's facial expression was happy, most of the children crawled onto the plexiglass panel. When the caregiver's facial expression was fearful, no children crawled on the plexiglass panel.
This phenomenon is called "social reassurance". When a child feels uncertain in a new situation, it seeks information from the facial expression of its caregiver. If the person smiles (positive information), the child feels safe. If the information is negative, the child avoids the situation. Accordingly, emotional information is conveyed through facial expressions. Thus, as with affective prosody, emotion can be gleaned from a person's face as an objective perception of their mood (Bolten & Schneider, 2010).
In this study, facial features are measures using the Facial Action Coding System (FACS). FACS is a system to scientifically measure facial features and was developed based on knowledge of muscle movement and the associated facial features (Waller et al., 2020). These muscle movements are called action units (AU) (Ekman et al., 2002). Ekman postulates that emotions have a universal form of expression. However, their interpretation and intensity are dependent on culture, personality and situation.
Falsifications occur when a person attempts to convey an emotion through facial features even though that emotion is not present, or when attempting to conceal a facial expression that would reflect the actual emotion. Accordingly, emotions can be determined from the face, but possible distortions must be taken into account (Ekman, 1993). We could conclude that there are interindividual differences, but the similarities outweigh them. Thus, FACS can be used to make an objective measurement of mood (Ekman et al., 2002).

2.2 Variables influencing mood

In order to determine a person's mood more accurately, variables that influence the mood of the test subjects are also examined. In this case, the influencing variables are sleep and personality.

2.2.1 Sleep and mood

"The quality and amount of sleep (…) influences the way we react to these events and may be an important determinant in general well-being" (Vandekerckhove & Cluydts, 2010, p.220). It is assumed that a person's mood is influenced by REM sleep, whereby reduced sleep leads to greater sensitivity to emotional

events during the day (Vandekerckhove & Cluydts, 2010). However, a person's sleep can also be influenced by their mood (Hold, 2014). In this context, a negative mood can be equated with stress (Dreisbach, 2008). One study measured the relationship between sleep and stress. Hold (2014) postulates that experiencing stress has an influence on subjective sleep quality. The more negative the mood, the poorer the subjective sleep quality. Accordingly, the relationship between mood and sleep is a bidirectional causal connection and is therefore relevant to the present study, because it is vital that a person's subjective sleep should be taken into account when assessing their mood.

2.2.2 Personality and mood

Personality is considered as a stable trait over time (Kumpf, 2016). The construct of personality could be summarized by factor analysis into five facets of personality, the five-factor-model. The dimensions are extraversion, agreeableness, neuroticism, conscientiousness and openness. The model is widely accepted by researchers. It is very relevant to take personality into account, as it describes the most important individual differences in humans (Goldberg, 1990).
Personality traits influence a person's way of thinking, behaviour, and coping with situations in response to external stimuli. Depending on the expression of a personality trait, people react differently to positive and negative moods in their environment (Kumpf, 2016). A study by Larsen & Ketelaar (1991) postulates that a stronger expression of neuroticism shows a higher sensitivity to positive external stimuli (Larsen & Ketelaar, 1991). A more recent study once again demonstrated the connection between extraversion and neuroticism and positive and negative moods (Canli et al., 2004). Based on multiple confirmations, it can therefore be assumed that these effects actually exist. Accordingly, it can be assumed that personality traits have an influence on mood and are therefore extremely relevant to the present study.

2.3 Aim of the study

With regard to the theoretical background mentioned above, the study aims to examine the relationship between objective and subjective measurements and aims to determine the extent to which the objective measurements are valid. In addition, demographic data should be taken into account with regard to moderation.

This results in the following research question:
1. To what extent do objective audiovisual features correlate with momentary self-reported mood?
2. Are the objective measures psychometrically / criterion-valid (convergent and incremental validity vs. self-report)?
3. Do demographics (age, sex/gender, education, language fluency, sleep, personality traits) moderate these associations? (exploratory research question)

This gives rise to the following hypothesis:
H1: There is a correlation between the objective audiovisual features and the momentary subjective self-reported mood.
H2: The objective measures show psychometric validity (incremental/criterion validity) in predicting subjective mood.
H3: Sleep moderates this association.
H4: Personality moderates this association.

3. Methods

In a 7-day, within-subjects study (Luxembourg), participants complete the validated six-item mood short scale (~1–2 min) and a 90-second prompted video daily. We compute MediaPipe facial landmarks and openSMILE acoustic features.

3.1 Sample

The sample consisted of English-speaking adults (≥18 years) (N = 38, 7 male, 27 female and 1 divers) recruited via posters and online notices at the University of Luxembourg. Participation was anonymous and voluntary. The

recruitment emphasised inclusivity and diversity in participant demographics to provide a heterogeneous sample. To maximise protocol adherence while preserving voluntariness, all participants will receive a 10€ voucher irrespective of completion. We will emphasise the value of complete data and support adherence via brief reminders and progress feedback. Compensation remains unconditional. The age ranged from 18 to 39 (M = 23.8, SD = 4.19).

3.2 Procedure and Measures

After registration, participants receive an information sheet and provide informed consent. Afterwards some demographics were measured once via a Qualtrics questionnaire. Here the participants had to indicate some demographic data (see Appendix A). A short Personality assessment was also included, the BFI-10. The BFI-10 from Rammstedt & John (2007) is a 10-item short form of the Big Five (5-point Likert). brief but widely used and provides acceptable reliability for research with minimal burden.
For the daily assessment a link was sent to the participants via e-mail. Every day they got a reminder. In the daily assessment, the participants had to fill out a questionnaire about their Momentary mood. The six-item short scale from Wilhelm & Schoebi (2007) was used. This is an empirically validated EMA instrument assessing valence, calmness, and energetic arousal (two bipolar items per dimension; 7-point response). Items are anchored to “right now”; subscales are item means (higher = more positive valence, greater calmness, higher energetic arousal).
Afterwards a 90-second self-report video with audio responding to a standardised prompt about the day and feelings had to be done. The participants had to read a basic sentence and then speak freely. To avoid self-correction behaviour regarding their facial expressions, the screen blacks out as soon as the recording starts.
During the recording Facial landmarks are extracted in real time with MediaPipe (frame-wise landmark coordinates). Raw video is not stored; only derived landmark time-series are retained. No face recognition is performed. Acoustic features were extracted with openSMILE (audEERING) from the daily audio track (e.g., prosodic, spectral, and voice-quality descriptors). To minimise identifiability, we will store derived features only (no raw audio).

3.3 Statistical Analysis

For all analyses, continuous recordings were reduced to a single row per participant and EMA session. Segment-level acoustic features from audEERING were first extracted for each speech segment and then summarised at the session level using duration-weighted means, with segment length as the weight. Facial action unit (AU) intensities (0–5) and AU-on indicators (0/1) were treated as framewise time series and aggregated per session using time-weighted means, where weights corresponded to the interval between successive frames to approximate their represented time. MediaPipe blendshape coefficients (continuous deformation weights for specific facial regions such as lips, jaw, or brows) were processed in the same way, yielding per-session time-averaged blendshape intensities. These session-level summaries of voice features, AUs, and blendshapes were then merged with the corresponding self-report EMA ratings and questionnaire data.
Then a Pearson correlation was done to quantify how strongly self-reported mood is associated with objective voice and face features. A correlation analysis was also used with the different personality scales to look for any effects.
For the criterion and incremental validity, a hierarchical regression was conducted. The hierarchical regression was used to test whether objective voice and facial features explain additional variance in event ratings after accounting for self-reported mood. Finally, a moderation regression was done with sleep factors and personality. Therefore, the variables were centred to be more easily interpreted. All

statistical analyses were conducted using SPSS.

4. Results

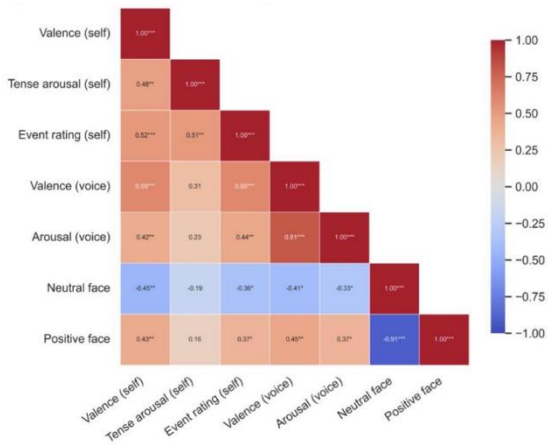


Figure 1: Correlation Matrix of Self-report, Voice and face Features

Hypothesis 1: There is a correlation between the objective audiovisual features and the momentary subjective self-reported mood.
A correlation analysis of self-report mood and voice and face variables was conducted to assess whether there is a significant connection between subjective and objective measurements of mood. The results of this correlation analysis (Pearson) revealed a significant and moderate relationship between subjective and objective measurements of mood. Voice features show a significant and moderate correlation, whereas face features display a significant and moderate correlation (see figure 1).

Hypothesis 2: The objective measures show psychometric validity (incremental/criterion validity) in predicting subjective mood.
Convergent validity was assessed via a Pearson correlation and strongest for valence (voice derived valence with self-reported valence, r = .60, p < .001).
A hierarchical regression was conducted to investigate whether objective voice and face features explain additional variance in an outcome (event rating) after accounting for self-reported mood. This directly addresses incremental validity. In step 1, self-report arousal mean, and self-report valence mean predicted a significant portion of the variance in event rating, R^2 = .355, adjusted R^2=.318, F(2, 35) = 9.638, p < .001.
Adding objective measures (voice mean, arousal mean, voice mean valence mean, face neutral blend shape mean, and face positive expression blend shape mean) in step two increased explained variance to R^2 = .489 (ΔR^2 = .134) and adjusted R^2 = .390. This change was, however, not statistically significant, F(4, 31) = 2.026, p = .115.
The standardised coefficients for self-report valence mean reveal a moderate positive influence on self-report event rating $|\beta|$ = .354, p = .028. Self-report tense arousal mean shows a similar moderate positive influence on event rating $|\beta|$ = .338, p = .036.
The standardised coefficients for objective measures reveal a moderate positive influence of voice mean valence mean on event rating, which is, however, not statistically significant $|\beta|$ = .486, p = .063. Other objective measures showed a weak influence -.1 < $|\beta|$ < .1 with no statistical significance p > .6.
Criterion validity was overall supported by the significant association between voice features and event rating (r =.60, p < .001).
Across analyses voice-derived features consistently demonstrated larger effect sizes than those of facial composite measures, although facial composites displayed smaller but significant associations with valence and event ratings ($|r| \approx .36–.45$).
Hypothesis 3: Sleep moderates the association between momentary mood and objective audiovisual features.
A moderated regression has been conducted to examine whether the strength or direction of the association between objective indicators (voice and facial features) and self-reported mood differs with regards to the variable “sleep”. Sleep showed a significant positive association with self-reported valence β = .41, p = .004. The moderation between sleep and voice on self-reported valence was also significant and moderate β = .31, p = .030.
Hypothesis 4: Personality moderates the association between momentary mood and objective audiovisual features.

A further moderated regression was conducted to assess the role of personality variables. Neuroticism showed a significant negative association with self-reported valence $\beta = -.382$, $p = .016$. Other personality variables did not display any significant influence on self-reported valence.

5. Discussion

The purpose of this study was to gain better understanding of the quantification of the associations between momentary mood and objective audiovisual features within a person and the evaluation of convergent/incremental validity of features compared to self-reported mood.

Hypothesis 1: There is a correlation between the objective audiovisual features and the momentary subjective self-reported mood.
Our data provides evidence that there is a significant and moderate correlation between objective and subjective measurements of mood. Voice features and facial features both show moderate agreement with self-reported mood, although voice features show a stronger relationship with the self-reported mood. These results are consistent with the claim that the face expresses emotions to a similar extent as the voice (Bolten & Schneider, 2010). This phenomenon is known as affective prosody and can explain a person's mood and state of mind through the voice (Ackermann et al., 2004). Furthermore, when talking about the affect-as-information hypothesis, which assumes that facial expressions and voice features give away information about a person's mood, we need to take into consideration that this does not always occur (Schwarz, 1987).

Hypothesis 2: The objective measures show psychometric validity (incremental/criterion validity) in predicting subjective mood.
A hierarchical regression indicated that self-report significantly predicts much of the variance in the self-reported event rating, whereas objective measures increase the variance without statistical significance. These results indicate that self-report arousal and self-report valence reveal a moderate positive influence on the self-report event rating and only voice valence revealed a moderate positive influence on self-reported event rating, which is however not statistically significant. The hypothesis that objective measures show psychometric validity in predicting subjective mood was therefore not supported, which could be, because objective features only have a small influence on self-reported emotions. Although it is also important to consider other reasons that explain these results. For instance, it might be due to the small sample size (n=38). This small sample size makes it harder to find significance, and thus suggests a larger sample size so significance can be reached.
Another contributing factor could be that objective features might be harder to measure in comparison to subjective features by using MediaPipe facial landmarks and openSMILE acoustic features, which are a reliable tool for objective measurements. Therefore, it remains important to consider that technological tools often still lack the ability to detect and interpret human emotions correctly.

Hypothesis 3: Sleep moderates the association between momentary mood and objective audiovisual features.
A moderate regression analysis reveals that sleep moderates the relationship between objective (voice and face) and self-reported mood, indicating that the direction or strength of this relationship varies depending on sleep. Sleep showed a significant positive association with self-reported valence. In other words, people who have a greater quality and amount of sleep tend to report feeling more positive. The present results are consistent with (Vandekerckhove & Cluydts, 2010) work that deals with the assumption that a person's mood is influenced by REM sleep and that reduced sleep makes a person more strongly sensitive to emotional events.
In addition, the moderation between sleep and voice self-reported valence was also significant. This means that sleep quality and duration influence how well the voice reflects emotional state.

Hypothesis 4: Personality moderates the association between momentary mood and objective audiovisual features.
A moderate regression indicated that there is a negative association between neuroticism and self-reported valence. These results imply that higher levels of the personality trait neuroticism tend to report feeling more negative. These results are consistent with findings from previous studies that reported an association between neuroticism and negative emotional state (Larsen & Ketelaar, 1991). Although these studies also reported a significant relationship between extraversion and positive emotional state, our study did not find this association to be statistically significant.

5.1 Study Limitations

While the findings in this study provide valuable insights, certain limitations must be considered that may influence the interpretations of the results. One significant limitation in our study is the sample size, which may limit the statistical power and the generalisability. Moreover, was the subjective data collected through programmes. Although these have proven to be very reliable, they may capture emotions differently than how they were originally expressed. As a final consideration, measurement may show limitations because subjective data was collected through self-report measures. These measures are more vulnerable to bias and social desirability and may not reflect participants actual behaviors. In this study, we mainly constricted ourselves on detecting general effects in objective and subjective measures, rather than also considering the emotional intensity in the self-report. The actual content of the self-report was not analysed. This could have shown other important effects and additional differences in the subjective and objective measures. Despite these limitations, the study has helped us to understand the relationship between objective and subjective measures and can provide as a support for future research.
5.2 Practical application
Our findings suggest that face and especially voice, reliably reflect a person's emotional state. These findings show some practical implications in areas where understanding emotional states is relevant such as clinical or forensic settings. However, this phenomenon should be further analysed with larger sample sizes before such methods can be applied.

5.3 Future research

Regarding future research, it would be useful to expand the current findings by examining the content of the self-reports and analyse whether these may show a stronger effect. Furthermore, the present data could be more expanded by including not only demographics, such as sleep or personality, but also measures of physical fitness, which were collected in this study but not included in the current analyses.

6. Note on the use of assistive devices

DeepL was used to translate individual text passages. All text were checked and improved where necessary.

Literatur

Ackermann, H., Hertrich, I., Grodd, W., & Wildgruber, D. (2004). „Das Hören von Gefühlen": Funktionell-neuroanatomische Grundlagen der Verarbeitung affektiver Prosodie. Aktuelle Neurologie, 31(09), 449–460. https://doi.org/10.1055/s-2004-828377

Bolten, M., & Schneider, S. (2010). Wie Babys vom Gesichtsausdruck der Mutter lernen. Kindheit und Entwicklung, 19(1), 4–11. https://doi.org/10.1026/0942-5403/a000002

Canli, T., Amin, Z., Haas, B., Omura, K., & Constable, R. T. (2004). A double dissociation between mood states and personality traits in the anterior cingulate. Behavioral Neuroscience, 118(5), 897–904. https://doi.org/10.1037/0735-7044.118.5.897

Clore, G. L., Gasper, K., & Garvin, E. (2012). Affect as information. In Handbook of affect and social cognition (S. 122–145). Psychology Press. https://api.taylorfrancis.com/content/chapters/edit/download?identifierName=doi&identifierValue=10.4324/9781410606181-8&type=chapterpdf

Costa, P. T., & McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. Journal of Personality Disorders, 6(4), 343–359. https://doi.org/10.1521/pedi.1992.6.4.343

Dreisbach, G. (2008). Wie Stimmungen unser Denken beeinflussen.

Ekman, P. (1993). Facial expression and emotion. American Psychologist, 48(4), 384–392. https://doi.org/10.1037/0003-066X.48.4.384

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). Facial Action Coding System: Facial action coding system : the manual : on CD-ROM. Research Nexus.

Goldberg, L. R. (1990). An alternative „description of personality“: The big-five factor structure. Journal of Personality and Social Psychology, 59(6), 1216–1229. https://doi.org/10.1037//0022-3514.59.6.1216

Hold, A. (2014). Stress und Schlaf: Der Einfluss von Stress in einer akademischen Leistungssituation auf die Schlafqualität und die kardiovaskuläre Reaktivität während des Schlafes / Annelisa Hold. http://unipub.uni-graz.at/obvugrhs/239951

Johnson, J. A. (1981). The „self-disclosure“ and „self-presentation“ views of item response dynamics and personality scale validity. Journal of Personality and Social Psychology, 40(4), 761–769. https://doi.org/10.1037/0022-3514.40.4.761

Kumpf, U. (2016). Der Einfluss von Persönlichkeitsfaktoren auf die Stimmungsänderung gesunder Probanden nach repetitiver transkranieller Magnetstimulation (rTMS) [Text.PhDThesis, Ludwig-Maximilians-Universität München]. https://edoc.ub.uni-muenchen.de/19239/

Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. Journal of Personality and Social Psychology, 61(1), 132–140. https://doi.org/10.1037/0022-3514.61.1.132

Mackenthun, G. (2004). Stimmungen und Emotionen. Oktober. https://www.geraldmackenthun.de/app/download/5798541467/Stimmungen+und+Emotionen+(1998).pdf

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. Cognition and Emotion, 23(2), 209–237. https://doi.org/10.1080/02699930802204677

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of Research in Personality, 41(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. Speech Communication, 40(1–2), 227–256. https://doi.org/10.1016/S0167-6393(02)00084-5

Schwarz, N. (1987). Stimmung als Information: Zum Einfluß von Stimmungen und Emotionen auf evaluative Urteile (Bd. 1987/12). Zentrum für Umfragen, Methoden und Analysen -ZUMA-.

Trull, T. J., & Ebner-Priemer, U. W. (2020). Ambulatory assessment in psychopathology research: A review of recommended reporting guidelines and current practices. Journal of Abnormal Psychology, 129(1), 56–63. https://doi.org/10.1037/abn0000473

Vandekerckhove, M., & Cluydts, R. (2010). The emotional brain and sleep: An intimate relationship. Sleep Medicine Reviews, 14(4), 219–226. https://doi.org/10.1016/j.smrv.2010.01.002

Walker, M. P. (2009). The role of sleep in cognition and emotion. Annals of the New York Academy of Sciences, 1156, 168–197. https://doi.org/10.1111/j.1749-6632.2009.04416.x

Waller, B. M., Julle-Daniere, E., & Micheletta, J. (2020). Measuring the evolution of facial „expression“ using multi-species FACS. Neuroscience and Biobehavioral Reviews, 113, 1–11. https://doi.org/10.1016/j.neubiorev.2020.02.031

Weigand, D. (2018). Objektive Tests: Empirische Untersuchung zu Integration und Kontrast direkter und indirekter Messverfahren [Phd]. https://doi.org/10.5283/epub.37863

Wilhelm, P., & Schoebi, D. (2007). Assessing mood in daily life: Structural validity, sensitivity to change, and reliability of a short-scale to measure three basic dimensions of mood. European Journal of Psychological Assessment, 23(4), 258–267. https://doi.org/10.1027/1015-5759.23.4.258

Zebrowitz, L. A., Brownlow, S., & Olson, K. (1992). Baby talk to the babyfaced. Journal of Nonverbal Behavior, 16(3), 143–158. https://doi.org/10.1007/BF00988031

Appendix

Appendix A: Demographics questionnaire

Demographics questionnaire

Start of Block: Informed consent

I have read the information sheet and have been informed about the nature, potential conse quences, and risks of the research project "Objective and subjective assessment of emotional states in everyday life." I have had sufficient opportunity to ask questions and to consider my participation in this study. I understand that my data will be collected and used in connection with this research project and to enable publication of the research results. I have been informed that I am entitled to withdraw my consent to participate in the research project at any time without giving a reason and without negative consequences to myself. Furthermore, I acknowledge that due to the anonymised storage of data, it will not be possible to identify or delete my specific data after it has been collected. Consent Options (Please tick the appropriate boxes in the table below): I consent to the collection and use of my data in relation to the Research Project.

- ○ Yes (1)
- ○ No (4)

Consent Options (Please tick the appropriate boxes in the table below): I consent to the collection and use of my data in relation to the Research Project.

- ○ Yes (1)
- ○ No (2)

I am happy to be contacted after this Research Project to ask whether I would be interested in taking part in a follow-up study.

- ○ Yes (1)
- ○ No (2)

I consent to my data, as described in the information sheet, being processed for the purposes of the Research Project.

- ○ Yes (1)
- ○ No (2)

End of Block: Informed consent

Start of Block: Demographic information

Q1 Please enter your e-mail address (university email in code form if possible. e.g. 0247……@uni.lu). You will be contacted via this email address for the next part of the experiment.

_____

_____

_____

_____

_____

Q2 What is your assigned sex at birth?

- ○ Male (1)
- ○ Female (2)

Q3 Which gender do you identify with?

- ○ Male (1)
- ○ Female (2)
- ○ Non-binary / third gender (3)
- ○ Prefer not to say (4)

Q4 What is your age?

_____

_____

_____

_____

_____

Q5 Which country were you born in?

_____

_____

_____

_____

_____

Q6 Which language is your mother tongue?

_____

_____

_____

_____

Q7 Which of the following best describes your ethnic group?

- ○ White (1)
- ○ Asian (2)
- ○ Black (3)
- ○ Arab (4)
- ○ Mixed or multiple (5)
- ○ Other (please specify): (6) _____
- ○ Prefer not to say (7)

Q8 Which country do you currently live in?

_____

_____

_____

_____

_____

Q9 How many years have you lived in your current country?

_____

_____

_____

_____

_____

Q10 How would you describe your cultural background? (e.g. heritage, community, culture(s) you identify with)

_____

_____

_____

_____

_____

Q11 Which describes your current menstrual/ reproductive status? (for participants with uterus)

- ○ Naturally cycling (1)
- ○ Using hormonal contraception (2)
- ○ Pregnant (3)
- ○ Breastfeeding (lactating) (4)
- ○ Post-meno pause (no period ≥ 12 months) (5)
- ○ Post-hysterectomy (6)
- ○ Post-oophorectomy (both ovaries removed) (7)
- ○ Prefer not to answer (8)
- ○ Does not apply to me (9)

Q12 Are you currently taking any of the following medication? (tick all that apply)

- ☐ Antidepressant (1)
- ☐ Antipsychotic (2)
- ☐ Mood stabiliser (3)
- ☐ Beta-blocker (4)
- ☐ None of the above (5)
- ☐ Prefer not to answer (6)

Q13 How much do you sleep per night on average?

- ○ less than 5 hours (1)
- ○ 5-6 hours (2)
- ○ 6-7 hours (3)
- ○ 7-8 hours (4)
- ○ more than 8 hours (5)

Q14 How much do you exercise per week on average?

- ○ Never (1)
- ○ Once a week (2)
- ○ 2-3 times a week (3)
- ○ 4-6 times a week (4)
- ○ Daily (5)

End of Block: Demographic information

Start of Block: Vision and potential occlusions relevant to facial landmarks

Q15 Do you usually wear:

- ○ Glasses (1)
- ○ Contact lenses (2)
- ○ Neither (3)

Q16 Do you expect to wear glasses during most recordings?

- ○ Yes (1)
- ○ No (2)
- ○ Unsure (3)
- ○ Not applicable (4)

Q17 Typical facial hair over the next 7 days

- ○ None (1)
- ○ Mustache only (2)
- ○ Beard-short (3)
- ○ Beard-medium (4)
- ○ Beard-full (5)

Q18 Typical make-up during recordings

- ☐ None (1)
- ☐ Light (2)
- ☐ Heavy (3)

Q19 Is lipstick typically used?

- ○ Yes (1)
- ○ No (2)

Q20 Do you have facial piercings that cover or occlude the lips or nose?

- ○ Yes (1)
- ○ No (2)

Q21 Do you consider yourself to be:

	Definitely a morning person (1)	More a morning person (6)	Do not know (7)	More an evening person than a morning person (8)	Definitely an evening person (9)	Prefer not to answer (10)
Do you consider yourself to be: (2)	○	○	○	○	○	○

End of Block: Vision and potential occlusions relevant to facial landmarks

Start of Block: Personality assessment

Q22 I see myself as someone that ...

	Disagree strongly (1)	Disagree a little (2)	Neither agree nor disagree (3)	Agree a little (4)	Agree strongly (5)
... is reserved (1)	○	○	○	○	○
... is generally trusting (2)	○	○	○	○	○
... tends to be lazy (3)	○	○	○	○	○
... is relaxed, handles stress well (4)	○	○	○	○	○
... has few artistic interests (5)	○	○	○	○	○
... is outgoing, socialble (6)	○	○	○	○	○
... tends to find fault with others (7)	○	○	○	○	○

... does a thorough job (8)	○	○	○	○	○
... gets nervous easily (9)	○	○	○	○	○
... has an active imagination (10)	○	○	○	○	○

End of Block: Personality assessment

Interpersonal self-control attributions

Jana Bellia, Kim Lenners, Nina Ständebach, Aysun Weicker and Leo Wohl

Supervisor: Dr. Axel Grund

This preregistered online experiment examined the social function of self-control attributions in an educational context, drawing on attribution theory and prior work highlighting self-control as a key predictor of academic and broader life out-comes. Specifically, we tested whether preservice teachers attribute student aca-demic failure more strongly to low self-control when their professional identity is threatened by external blame. German-speaking preservice teachers (final sample: N = 101; 84.2% female; age 18–27, M = 21.9, SD = 3.88) completed an online sur-vey after reading either a parent letter blaming teachers for student failure (ex-perimental condition) or a neutral text about health as a school subject (control condition). Participants then answered manipulation-check evaluations, attribu-tion-related controllability judgments, emotional reactions, and intended peda-gogical responses; group differences were tested using MANOVAs across outcome domains. Results indicated a successful manipulation: the blaming letter elicited markedly more negative evaluations (e.g., anger). However, the manipulation did not produce significant differences between conditions in perceived controllabil-ity of self-control as a cause for academic failure, emotional reactions (sympathy, anger), or behavioural intentions (e.g., support) related to self-control. Overall, the findings suggest that while identity threat can shift immediate affective evalu-ations of criticism, it may not readily alter underlying attributional beliefs or downstream responses toward students in this preservice teacher sample.

1.Theoretical background

1.1 Self-control and the big promise

Self-control is widely defined in educational psychology as the "self-initiated regulation of thoughts, feelings, and actions when enduringly valued goals conflict with momentarily more gratifying goals" (Duckworth et al., 2019). In simpler words, it means being able to control your own behaviour when immediate wishes clash with long-term goals. Within the academic context, this ability is believed to be essential, as students often struggle between pursuing long-term educational goals and giving in to immediate distractions. Duckworth et al. (2019) point out that almost all students experience this tension, highlighting the central role of self-control in meeting educational requirements.

Research suggests that self-control isn't just a helpful skill but also a strong predictor of academic success. Duckworth and Seligman (2005) found that self-control consistently predicts academic outcomes such as grades and, in some cases, even exceeds intelligence as a predictor of performance. This underlines that being disciplined and able to resist short-term temptations can matter more for academic success than raw cognitive ability alone.

However, the importance of self-control goes far beyond the classroom. Moffitt et al. (2011) demonstrated that childhood self-control predicts a wide range of major life outcomes, including physical health, personal finances, and public safety. In other words, children who learn to regulate themselves early are more likely to grow into adults who manage their health better, handle money more responsibly, and avoid risky or harmful behaviour.

1.2 Attribution and perceptions of self-control

Based on attribution theory (Weiner, 1985), attributing academic failure to specific causes has a strong influence on emotional and beha-

vioural responses. When failure is attributed to a lack of self-control, specific perceptions may arise: Self-control is usually seen as an internal and unstable factor. Consequently, it is often judged as something the student can change or improve, making it highly controllable (Gaier, 2015).
This perception of controllability is crucial. If teachers believe that a student had control over their failure, the emotional reaction tends to be anger rather than empathy. Reyna and Weiner (2001) showed that such attributions reduce the willingness of teachers to provide support and encouragement.

1.3 Aim of the study

The present study investigates the social function of self-control attributions. Specifically, we examine whether preservice teachers change their opinions on self-control when their professional identity is under threat. In other words, we examine if teachers shift the way they explain student failure depending on whether they are blamed.
To test this, we manipulate the professional responsibility of teachers for academic success and failure. In one condition, teachers' professional identity is threatened by blaming them for the academic failure of the student. In the control condition, no such blame is assigned. By comparing these two situations, we aim to analyse whether this threat leads to a change in the perceived controllability of self-control as a potential cause for academic failure. This way, preservice teachers could reject the blame attributed to them.

1.4 Hypotheses

Based on the outlined theory, we expect that teachers' attribution of academic failure to a lack of self-control will be more pronounced when their professional role is questioned. In situations where teachers feel blamed for a student's poor performance, they should be more likely to judge self-control as the main cause of failure compared to a control condition. Furthermore, in the experimental condition, self-control is expected to be perceived as something that is particularly controllable by the student but not by the teacher, reinforcing the idea that responsibility lies with the learner. As a consequence of these perceptions, teachers under threat should experience stronger feelings of anger and reduced sympathy toward failing students, which in turn makes them less willing to provide support or help.

2.Methods

2.1 Sample

The sample consisted of pre-service teachers who were fluent in German and participated in the online survey on a voluntary basis. Recruitment took place via university mailing lists and social networks.
Participants were recruited from various universities and colleges across several German-speaking countries, as well as through family and friends of the researchers.
The online survey was conducted using SoSci Survey (Leiner, 2025); the data was collected between 27.10.2025 and 20.11.2025 at https://www.soscisurvey.de/Schulerfolgsattribution/.
On average, completing the questionnaire took approximately 958 seconds.
Participants were excluded if their completion time was excessively short (2 *SD* below the mean). Another exclusion criterion was the self-reported seriousness of participation. Individuals who scored below 5 on a 6-point Likert scale were excluded from the analyses.
In the approximately 3 weeks of running the study, *N* = *118* participants were gathered. After applying exclusion criteria, *N* = *101* participants were used for all analyses conducted, thereby slightly missing the target of 114 subjects, based on the assumption of a medium-sized effect (f^2 = .15), alpha = .05, and power = .80.
The corrected sample (*N* = 101) consisted of 85 female (84,2 %) and 16 male (15,8 %) participants with an age ranging from 18-27 (*M* = 21.9; *SD* = 3.88).
In terms of language, 78,2 % of participants considered their German to be on a mother tongue level.

2.2 Design

The study employed an experimental between-subjects design with two conditions:

· Experimental group: Participants read a parent letter that blamed teachers for students' academic failure ("Poor teaching jeopardizes students' futures").
· Control group: Participants read a neutral parent letter on the topic of "Health as a school subject".

The manipulation aimed to vary participants' perceived threat to teachers' professional identity and, consequently, to induce a shift in attribution patterns for academic failure among those in the experimental group. The study was preregistered at https://aspredicted.org/ prior to the initiation of data collection.

2.3 Instruments

The instruments recorded key assessments by prospective teachers regarding various potential causes of academic failure, in particular low self-control, lack of intelligence and low extrinsic motivation, of which we will focus on self-control in the following.

In addition, assessments of the frequency of characteristics exhibited by pupils in everyday school life (e.g. lack of self-control, lack of intelligence or low extrinsic motivation) were collected on a percentage scale. The participant's socio-demographic data and other personal assessments not relevant to the present study were also collected using a 6-point Likert scale, from 1 ("does not apply at all") to 6 ("applies exactly"). In terms of content, the questionnaire can be divided into four main categories: (a) manipulation check, (b) attribution dimensions (e.g., controllability), c) emotional reactions and pedagogical behavioural intentions, and (d) to what extend self-control is seen as the cause of academic failure in the perspective of the preservice teachers (prevalences).

2.4 Manipulation check

To check whether the negative bogus letter from parents had an influence on the participants, both the control and experimental groups evaluated several statements after reading the letter they had been assigned, which were designed to capture their emotional assessment and critical attitude, but also their agreement with the content of the letter (e.g. "I was annoyed by the letter"). The assessment was carried out on a 6-point Likert scale (1 = "does not apply at all" to 6 = "applies exactly"). These items were used to check whether the two texts differed in their emotional impact and whether the experimental manipulation was therefore successful.

2.5 Controllability (Attribution dimensions)

The perception of the causes of academic failure was assessed using various attribution dimensions. For each of the three potential causes examined, seven items were surveyed that reflect central constructs of attribution theory (Weiner, 1985). Based on the items, participants assessed the stability of the cause over time, the location of the cause (whether external or internal), and, most importantly for the present purpose, the controllability by the teacher and by the student as well as the pedagogical influenceability by the teacher and the trainability, i.e. long-term change (e.g. "Do you consider the cause attributable to the students to be currently controllable or uncontrollable (that is, whether it is voluntarily influenceable by the student at certain moments)?") All assessments were made on a 7-point Likert scale ranging from -3 ("not at all …") to +3 ("totally …"). These items allow for the collection of subjective assessments by prospective teachers regarding the extent to which self-control is perceived as changeable and influenceable by students and teachers. They also capture the central mechanism through which the experimental condition is intended to influence expectations regarding self-control attributions.

2.6 Emotion and behaviour

During the study, emotional responses to hypothetical pupils were assessed in both the experimental
and control groups. The participants were asked to imagine a student who had received a poor grade due to the cause described (e.g. poor self-control). On this basis, they indicated various emotional and behavioural reactions. First, key emotional reactions were recorded, including the extent of anger towards the student and feelings of compassion (e.g., "To what extent do you think that the student deserved the failure in the exam?"). Further, participants assessed the extent to which they at-

tributed blame for failure to the student or to themselves as teachers (e.g., “To what extent is it your fault as a teacher, that your student failed in the exam?”).
In addition, pedagogical behavioural intentions were surveyed. (e.g., “To what extent would you criticize/ admonish/ punish the students?”). Participants indicated the extent to which they would support the affected student and how strongly they would tend to react with punishment or support. All emotional reactions and behavioural intentions were recorded on a single item on a Likert scale from -3 = “not at all” to +3 = “very strongly” and built on Reyna and Weiner (2001).

2.7 Prevalences

To measure to what extend blamed “teachers” judge self-control as a more likely cause for academic failure, we formulated two items (e.g.” Many students you will have in the future, will experience academic failure. What do you estimate: How many of those failures are due to a lack of self-control from the students?”). The two Items were rated with percentages (estimated percentage of students that fail because of a lack of self-control), ranging from under 9 % to over 90 %.

2.8 Analysis

We applied four MANOVA’s to test whether both groups differ in regard to the dependent variables. Accordingly, we examined four outcome domains: manipulation check, controllability, emotion and behaviour, and prevalence. For all demographic data in the sample description, descriptive statistics (*M, SD, Frequencies, Percentages*) were used.

3.Results

3.1 Manipulation check

A multivariate analysis of variance (MANOVA) tested whether the manipulation influenced participants’ evaluative responses to the constructed reader letter.
The multivariate effect was significant, Wilks’ Λ = .57, $F(5, 95) = 14.1$, $p < .001$, indicating large differences between conditions.
Participants in the experimental condition evaluated the letter more negatively than those in the control condition.
Descriptive statistics (Table 1) show that participants in the experimental condition reported substantially higher anger and a more critical attitude toward the letter than those in the control group.

Table 1
Descriptive Statistics for Manipulation Check Items

	Control *M (SD)*	Experimental *M (SD)*
Anger	1.65 (.87)	3.16 (1.30)
Critical Attitude	2.45 (1.05)	4.04 (1.20)
Persuasiveness	4.63 (.98)	3.66 (1.10)

Note. Only example items are shown per category.

Univariate ANOVAs confirmed significant condition differences across all five manipulation-check items (Table 2), including anger, agreement, critical stance, persuasiveness, and comprehensibility. Together, these results indicate that the manipulation was successful in activating and threatening participants’ identity as a preservice teacher.

Table 2
Univariate Tests for Manipulation Check

	F(1, 99)	*p*
Anger	47.53	< .001
Critical Attitude	50.67	< .001
Persuasiveness	21.84	< .001

Note. Only example items are shown per category.

3.2 Controllability

To examine whether the manipulation influenced participants’ beliefs about the nature of self-control as a cause of academic failure, a MANOVA was conducted on student controllability, student trainability, and teacher controllability.
The multivariate effect was not significant, Wilks’ Λ = .97, $F(3, 97) = 1.09$, $p = .359$, indicating no systematic differences between groups (Table 3).
Univariate analyses likewise revealed no significant group differences (all *p*s > .14).

Table 3
Descriptive Statistics and Univariate Tests for Controllability, Emotional and Behavioural Variables

	Control *M (SD)*	Experimental *M (SD)*	*F*(1, 99)	*p*
Controllability				
Student Controllability	3.76 (1.35)	4.18 (1.50)	2.17	.144
Student Trainability	5.57 (1.19)	5.62 (.90)	.06	.807
Teacher Controllability	4.39 (1.51)	4.70 (1.28)	1.22	.273
Emotion and Behaviour				
Student Fault	4.47 (1.38)	4.40 (1.30)	.07	.791
Teacher Fault	3.80 (1.33)	4.10 (1.34)	1.24	.268
Deservingness	2.90 (1.38)	2.56 (1.42)	1.52	.221
Support	6.12 (.84)	6.28 (0.70)	1.11	.295
Punishment	2.20 (1.18)	1.96 (1.00)	1.16	.284
Sympathy	5.57 (1.08)	5.48 (1.33)	.14	.714
Anger	2.78 (1.50)	2.84 (1.63)	.03	.859

3.3 Emotional and behavioural reactions

As attributional beliefs are thought to trigger emotional responses (Weiner, 1985), a second MANOVA assessed whether the manipulation affected emotional reactions as well as behavioural responses toward the student.

The analysis showed no significant multivariate effect, Wilks' $\Lambda = .95$, $F(7, 93) = .695$, $p = .676$.

Univariate analyses likewise revealed no significant group differences (all ps > .22).

This pattern suggests that although the manipulation changed direct evaluations of the letter itself, it did not influence participants' intentions to support, sympathize with, or criticize the student. It also indicates that participants' emotional attributions remained stable across conditions.

4.Discussion

4.1 Summary of the results

The present study examined whether and how threats to the professional identity of prospective teachers influence the attribution of academic failure, with a particular focus on attributions to low self-control. Drawing on attribution theory (Weiner, 1985), we assumed that an identity threat induced by a critical letter from parents would lead prospective teachers to attribute students' poor performance more strongly to internal and controllable causes, such as a lack of self-control. In line with the previous research, these attributions were expected to be associated with increased anger, reduced empathy, and a lower willingness to provide support (Reyna & Weiner, 2001).

Contrary to these hypotheses, the results revealed no significant differences between the experimental and control groups with regard to causal attributions, emotional reactions, or intended pedagogical behaviour toward students suffering from poor academic performance. Although the manipulation was successful – participants in the experimental condition reported higher levels of perceived criticism and anger in response to the offensive letter – this reaction did not generalise to their judgements about students' responsibility for academic failure.

This finding is particularly noteworthy in regard to our expectations concerning self-control. Based on previous research, self-control is commonly perceived as an internal and highly controllable factor (Duckworth et al., 2019; Gaier, 2015), which typically elicits stronger blame-related responses. We initially assumed that a threat to professional identity would therefore increase the tendency to view students' poor performance as intentional and controllable. The absence of such effects suggests that threats to professional identity did not translate into more blaming or less supportive responses toward students.

4.2 Theoretical classification

The absence of significant effects may indicate that attributions to self-control are perceived by prospective teachers as relatively stable and are therefore not easily influenced by situational identity threats. As outlined in the theoretical background, self-control is strongly embedded in broader cultural and educational understandings of academic success and failure (Duckworth et al., 2019). These entrenched beliefs may render attributional judgements less susceptible to short-term experimental manipulations.

At the same time, prospective teachers may already possess emerging professional norms or role-related standards acquired through internships and early teaching experiences. These developing professional schemas may function as a protective mechanism, preventing them from directly attributing blame to students even when their own professional identity is threatened. From this perspective, the present findings do not necessarily contradict assumptions of attribution theory (Weiner, 1985) but rather suggest that attributional pro-

cesses in teacher education are embedded within a broader professional identity framework that already shows a degree of stability (Kelchtermans, 2005).

This interpretation, however, stands somewhat in contrast to the idealism typically observed in prospective teachers prior to extensive practical experience. At early stages of their studies, many believe that teachers have a substantial influence on their students' learning and success and therefore see themselves as largely responsible for academic outcomes. This developmental pattern has been described in the context of the so-called “Konstanzer Wanne” (Dann et al., 1978).

The concept of the “Konstanzer Wanne” (Dann et al., 1978) refers to the practical shock experienced by newly qualified teachers experience after their first encounters with the practical field. While prospective teachers often begin their studies with rather traditional views on education, these are replaced during their studies by liberal, student-oriented, and idealistic attitudes. These orientations remain relatively stable until the first prolonged encounters with everyday school practice, where a “practical shock” may occur. Following this phase, more conservative attitudes – emphasising clear rules, discipline, hierarchy, and control – often re-emerge as adaptive responses to institutional demands and workload.

As the participants in the present study had not yet experienced this practical shock, they were likely still in the phase characterised by pedagogical idealism. Consequently, they may have been inclined to assume responsibility for students’ academic success themselves, rather than attributing failure to students’ lack of self-control. Although no statistically significant effects emerged, the absence of stronger student-blaming tendencies is in line with the assumption that preservice teachers may still predominantly assume responsibility for students’ academic outcomes.

Taken together, these considerations suggest that stronger attributional effects of identity threat may emerge in samples of fully qualified teachers. As professional experience accumulates, initial idealism may diminish and be replaced by a more differentiated and realistic professional self-concept, potentially leading to different attributional patterns in response to external criticism and identity threats.

4.3 Practical implications

Although the hypothesis could not be confirmed, the findings offer several important implications for teacher education. First, the results indicate that prospective teachers do not readily adopt negative or blaming attributions toward students, even when their professional identity is threatened. This suggests a notable degree of stability in their sense of pedagogical responsibility and professional ethics.

Second, the findings imply that attributions related to self-control may be less flexible than initially assumed. From a practical standpoint, this suggests that interventions aimed at fostering reflective attributional processes may need to be more intensive, repeated and explicitly integrated into teacher education curricula in order to produce lasting changes.

From a practical perspective, this indicates that interventions aimed at fostering reflective attributional processes may need to be more comprehensive and implemented more consistently in order to produce lasting effects. In addition, teacher training programs should explicitly address how future teachers can respond constructively to external criticism without allowing negative emotions, such as anger and frustration, to influence their attitudes or behaviour toward students.

4.4 Limitations

Several limitations of the present study should be considered. First, the use of hypothetical scenarios limits ecological validity, as such vignettes can only approximate real classroom situations and may encourage socially desirable responding. Consequently, participants’ reactions may differ from those observed in real-life teaching contexts.

Second, the sample consisted exclusively of prospective teachers, which restricts the generalisability of the findings to fully qualified and experienced teachers. Moreover, the length and partial repetition of questionnaire sections may have contributed to participant fatigue or automated response patterns.

Finally, the online format of the study limited control over participants’ attention, engage-

ment and working conditions, potentially introducing additional variability into the data.

4.5 Prospects for future research

Future research could benefit from incorporating process-oriented measures such as reaction times, implicit judgements or more fine-grained physiological indicators to gain deeper insight into attributional dynamics. Additionally, examining the moderating role of individual belief systems – such as a belief in a just world (Lerner, 1980) or work ethic – may help clarify under which conditions identity threats influence teachers' attributions.

Extending this line of research to real educational settings would be highly valuable. Investigating qualified or experienced teachers could provide a more comprehensive understanding of how attributions are formed in concrete classroom situations and how factors such as stress, perceived responsibility and institutional constraints shape these processes.

4.6 Conclusion

Overall, the study indicates that although an experimentally induced identity threat elicits short-term emotional reactions, it does not affect attributional processes related to self-control or intended pedagogical responses toward students, respectively. This suggests that self-control is perceived as a relatively stable cause of academic failure and is not easily altered by contextual influencers. Furthermore, the findings highlight the importance of examining identity-related and attributional processes in prospective teachers in a nuanced and developmentally sensitive manner.

5.References

Dann, H.-D., Humpert, W., Krause, F., & Tennstädt, K.-C. (1978). *Lehrerbiographien und Lehrerrolle: Zur Sozialisation des Lehrers*. Urban & Schwarzenberg.

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents. *Psychological Science*, *16*(12), 939–944. https://doi.org/10.1111/j.1467-9280.2005.01641.x

Duckworth, A. L., Taxer, J. L., Eskreis-Winkler, L., Galla, B. M., & Gross, J. J. (2019). Self-Control and Academic Achievement. *Annual Review of Psychology*, *70*(1), 373–399. https://doi.org/10.1146/annurev-psych-010418-103230

Gaier, S. E. (2015). *Understanding Why Students Do What They Do: Using Attribution Theory to Help Students Succeed Academically*. 6–19.

Kelchtermans, G. (2005). Teachers' emotions in educational reforms: Self-understanding, vulnerable commitment and micropolitical literacy. *Teaching and Teacher Education*, *21*(8), 995–1006. https://doi.org/10.1016/j.tate.2005.06.009

Leiner, D. J. (2025). *SoSci Survey* (Version 3.8.00) [Software]. SoSci Survey GmbH. https://www.soscisurvey.de

Lerner, M. J. (1980). *The Belief in a Just World*. Springer US. https://doi.org/10.1007/978-1-4899-0448-5

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Thomson, W. M., & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*(7), 2693–2698. https://doi.org/10.1073/pnas.1010076108

Reyna, C., & Weiner, B. (2001). Justice and utility in the classroom: An attributional analysis of the goals of teachers' punishment and intervention strategies. *Journal of Educational Psychology*, *93*(2), 309–319. https://doi.org/10.1037/0022-0663.93.2.309

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*(4), 548–573. https://doi.org/10.1037/0033-295X.92.4.548

Imprint
Luxemburger Experimentalpraktikum Journal
Vol. 19, Jahrgang 2025

Manuscript Guidelines
Submissions must follow the "Publication Manual of the American Psychological Association" (2010) or the "Guidelines for the Preparation of Term Papers and Theses" (2004) of the German Psychological Society (DGPs).

Frequency of Publication
Once to twice a year

Subscription / Contact
Assoc. Prof, Heiko Schütt, Université du Luxembourg, Maison Sciences Humaines 11, Porte des Sciences, L-4366 ESCH

Available online at
https://bap.uni.lu

ISSN 3093-1045