LEX

Luxemburger Experimental-praktikum Journal

Zeitschrift psychologischer Forschung
Revue de recherche en psychologie

Band 18, Jahrgang 2024

Université du Luxembourg
Bachelor scientifique en psychologie

ISSN 3093-1045

Luxemburger Experimentalpraktikum Journal
Band 18, Jahrgang 2024

Inhalt

Letting Off Steam: Testing Positive Effects of Playing Violent Video Games

Emily Krauß, Kara Karrenbrock, Sophie Silva, Inês Teixeira, Lotta Fritzen

Supervisor: M.Sc.-Psych. Gary Lee Wagener

This experimental study investigated potential cathartic effects of playing a violent video game, specifically Battlefield 1, as a means of alleviating subjective frustration and modu-lating aggressive behavior, as well as inducing physiological relaxation. Fifty-five partici-pants were randomly assigned to either the experimental group (with a frustrating cogni-tive task) or the control group (with a neutral cognitive task), before engaging in a 25-minute gameplay session. Heart rate variability and subjective frustration levels were as-sessed at baseline, after the cognitive task and after gameplay. Additionally, aggressive behavior was measured after gameplay. While participants in the experimental group showed a significant decrease in subjective frustration post-gameplay, indicating mood improvement, there was no significant reduction in aggressive behavior after the frustrat-ing task compared to the control condition. Additionally, the data showed that participants in both conditions experienced greater levels of physiological relaxation during and after gameplay. These findings suggest that engaging in violent video games may offer a cathartic release by alleviating frustration and inducing physiological relaxation, balancing out aggressive tendencies.

Introduction

In recent years, the influence of violent video games (VVGs) on aggression has been a topic of intense debate and scientific inquiry. With over half of the European population, totaling 237 million users aged between 6 and 64 years, regularly engaging in video games (European Commission, Directorate-General for Communications Networks, Content and Technology, 2023), the need for comprehensive research in this area is evident. While some studies suggest a link between VVGs and increased aggression (Anderson et al., 2010; Calvert et al., 2017), contrasting findings propose no such effect (Ferguson, 2007, 2010, 2020; Prescott et al., 2018). However, emerging research indicates potential positive effects of engaging in virtual violence. For instance, Lee et al. (2021) discovered that playing VVGs can reduce verbal and physical aggression, although the outcomes depended on the statistical model applied. Likewise, a study conducted by Tyack et al. (2020) indicated that playing a moderately violent video game (*Mark of the Ninja*[1]) after a frustration task led to a "mood repair effect" in the form of decreased frustration. Wagener & Melzer (2024) found that playing a VVG can induce physiological relaxation, particularly in individuals with elevated Machiavellianism scores.

The presented study set out to bridge existing gaps in research by empirically investigating the cathartic potential of playing a VVGs. Specifically, we aim to explore effects on alleviating frustration, mitigating aggressive behavior, and inducing physiological relaxation. Catharsis, a concept rooted in ancient Greek philosophy, refers to the release or purging of pent-up emotions and tension. According to the frustration-aggression hypothesis (Dollard et al., 1939), frustration – a negative emotional response stemming from the hindrance of goals or desires – leads to an increase in aggressive arousal. Aggressive arousal refers to a heightened state of physiological and psychological activation associated with aggressive tendencies or behaviors.

[1] Klei Entertainment. 2012. *Mark of the Ninja*. Videogame [PC]. Microsoft Studios, Washington, USA.

Geen and Quanty's (1977, p. 5) defined catharsis as "a reduction in aggressive arousal brought about through the performance of aggressive acts". In line with Mood Management Theory (Reinecke, 2016), which suggests that individuals can utilize media and entertainment to regulate their moods and emotions, engaging in virtual violence may reduce aggressive arousal, thus paving the way for mood restoration.

In this experimental study, we isolated potential catharsis effects by dividing participants into two groups. The experimental group was subjected to a frustration-inducing cognitive task, while the control group was subjected to a neutral cognitive task before playing the VVG *Battlefield 1* for 25 minutes. We combined repeated self-assessment questionnaires and repeated measures of physiological relaxation with an evaluation of aggressive behavior at the end of the study.

Three hypotheses were tested: (H1) playing a VVG after a frustrating task decreases subjective frustration and improves mood; (H2) there is a significant decrease in aggressive behavior post-frustration task; and (H3) playing a VVG induces physiological relaxation.

Methods

2.1. Participants

The required sample size of $N = 86$ was determined using G*Power 3.1.9.4 (Faul et al., 2007), calculated for F tests (ANOVA with repeated measures, between factors), considering an alpha level of 0.05, a power of .80, and an effect size of 0.25. Recruitment took place through the University of Luxembourg's mailing lists, flyers at the Campus Belval and online platforms (e.g., Facebook, WhatsApp, Moodle). Fifty-five participants took part aged between 18 and 39 years ($M = 23.21$; $SD = 4.22$), and 61.1% identifying as female and 38.2% as male. Inclusion criteria were proficiency in English and/or German and the absence of photosensitive Epilepsy and/or Reynaud's-Syndrome. Upon completing the study, participants received a 10 Euro GoGift-voucher, and psychology students earned study participation credits. Informed consent was obtained at the experiment's onset, and ethics approval was received from the Bachelor on Psychology program of the University of Luxembourg.

2.2. Violent Video Game

The 25-minute gaming session featured *Battlefield 1*, a first-person shooter game developed by *DICE* and published by *Electronic Arts* in 2016. To ensure consistent exposure to content and experiences between groups, all participants engaged in the single-player mission, *Avanti Savoia*, which follows the story of an Italian soldier within an elite military force during World War I. Participants were tasked with capturing and destroying enemy artillery positions, which involved shooting non-playable character enemy soldiers.

Avanti Savoia was chosen because it requires no prior gaming experience or knowledge about the game's story, ensuring that all participants enter the game with a similar level of understanding. As a single-player mission, it prevents the introduction of additional variables, such as social dynamics and communication. Additionally, all participants received instructions to the story and game controls before the gaming session. Participants were also informed that they do not need to complete the mission and that the other players in their team are not other humans, but non-playable characters as well as that they can continue to play even if their character dies during the gaming session.

2.2. Materials

2.2.1. Demographics

The participant sample consisted of $N = 55$ participants, with 61.1% self-identifying as female and 38.2% as male. The participants' mean age was 23.31 years ($SD = 4.216$), ranging from 18 to 39 years. Random assignment placed these 55 individuals into two conditions: the experimental group included 30 participants (54.5%), while the control group, comprised 25 participants (45.5%).

To gauge gaming experience, participants disclosed their weekly gaming involvement, averaging 3.98 hours (SD = 6.253), with a range from .00 to 24.00 hours. A one-way ANOVA revealed no significant difference between the experimental (M = 3.84; SD = 5.66) and control conditions (M = 4.16; SD = 7.01) regarding gaming experience, $F(1, 53) = 0.04$, $p = .849$, $\eta^2 < .01$.
On a scale ranging from 1 (“not at all”) to 5 (“absolutely”), participants reported a mean gamer identification score of 2.71 (SD = 1.499). Similarly to gaming experience, there was no significant difference between the experimental (M = 2.73; SD = 1.51) and control conditions (M = 2.68; SD = 1.52) regarding gamer identification, $F(1, 53) = 0.02$, $p = .897$, $\eta^2 < .01$.

2.2.2. *Cognitive Tasks*

Participants were randomly assigned to either the experimental group (with a frustrating cognitive task) or a control group (with a neutral cognitive task). Both tasks comprised six items sourced from the English version of the WAIS-IV (Wechsler Adult Intelligence Scale, Wechsler, 2012). The frustration-inducing version comprised the second-to-last or last items from the 6 subtests: matrices, vocabulary, information, figure weights, arithmetic, and visual puzzles. The neutral version comprised the easiest (first) items from the same subtest. All participants had a two-minute timeframe to complete the task. This time constraint was designed to render the frustration-inducing version nearly unsolvable, thereby inducing frustration. The material underwent pretesting to confirm that the frustration-inducing version did indeed elicit frustration. As anticipated, participants in the experimental condition reported a higher level of frustration regarding their performance in the cognitive task (M = 3.43; SD = 1.41) compared to the control condition (M = 2.04; SD = 1.10), $F(1, 53) = 16.25$, $p < .001$, $\eta^2 = .24$.

2.2.3. *Manipulation Check & Game Perception*

To assess potential influences of cognitive task and game characteristics on participants' emotional responses, we employed 14 items related to game perception and cognitive task perception. Examples include statements like "I was already familiar with the game", "I found the game frustrating", "The game was stressful", and "I was frustrated about my performance in the cognitive task." Participants rated their agreement with each statement on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). This not only facilitated control for potential influences of game characteristics but also allowed us to confirm the distinction in perceived complexity between the neutral task and the frustration task.

2.2.4. *Self-Reported Emotional State*

To follow the course of participants emotional state, participants indicated their perceived emotional state on a Visual Analogue Scale (VAS − Aitken, 1969; Bond & Lader, 1974; Gündoğdu et al., 2021; Liu et al., 2019; Luria, 1975).
Before and after the cognitive task, as well as after having played the VVG, perceived frustration, aggression and relaxation were examined. Additionally, participants indicated their responses on buffer-items (e.g., perceived happiness, annoyance, stress, and feelings of guilt). Participants were to place a cross on a straight horizontal line of fixed length (100 mm). Scores were documented with one decimal and ranged from 0 (e.g., “not frustrated at all” to 100 (e.g., “extremely frustrated”).

2.2.5. *Positive and Negative Affect*

To assess changes in positive and negative affect before and after engaging with the VVG, participants completed the Positive and Negative Affect Schedule (PANAS) twice (Breyer & Bluemke, 2016; Watson et al., 1988). The PANAS included a scale for positive affect with 10 items (e.g., enthusiastic) and a scale for negative affect with 10 items (e.g., guilty).

Participants rated each item on a 5-point Likert scale, ranging from 1 ("not at all") to 5 ("very much"). Internal consistency for both PANAS scales was very good across both measurement occasions (see Table 1).

Table 1

Internal Consistencies for Positive and Negative Affect at Baseline and after VVG

	Cronbach's α	*N* of Items	*N*
PA at Baseline	.89	10	55
PA after VVG	.83	10	55
NA at Baseline	.92	10	55
NA after VVG	.93	10	55

Note. PA = Positive Affect, NA = Negative Affect

2.2.6. Aggressive Behavior

Prior to engaging in the cognitive task, participants performed a Cold Pressor Task (CPT – Schwabe & Schächinger, 2018). During this task, participants were instructed to immerse their non-dominant hand in icy water for a duration of 45 seconds, with the option to remove their hand earlier if needed. The cold water induces a painful sensation due to the activation of temperature-sensitive pain receptors. After playing the VVG, participants were tasked with assigning a time duration (ranging from 0 to 600 seconds) for the Cold Pressor Task to the subsequent participants. The allocated timeframe served as an indicator for aggressive behavior, enabling the assessment of mean aggression levels.

In reality, all participants underwent the CPT for a uniform 45-second duration. However, specific duration details were withheld until after the study to prevent participants from using their own experiences as reference points or "anchors" when assigning a time frame to the subsequent participant. At the end of the study, participants were informed that the CPT served as a measure of aggression, and it was clarified that all participants experience the CPT for a duration of 45 seconds.

2.2.7. Pain Sensitivity

Pain sensitivity varies among individuals, and differences between conditions could influence the duration of the timeframe allocated for the CPT. To gain insight into participant's subjective sensitivity to pain and to control its potential influences, participants filled out a Pain Sensitivity Questionnaire (PSQ − Ruscheweyh et al., 2009). Participants were asked to imagine themselves in six imaginary situations related to temperature. For each situation, they indicated the level of discomfort or pain they would feel on a 10-point Likert scale (ranging from 0 = “no pain” to 10 = “the most severe pain imaginable or considered possible”). Reliability analysis showed that the internal consistency of the six PSQ-Items was high, with a Cronbach's Alpha coefficient of .86.

2.2.8. Heart Rate Variability

To investigate physiological reactions that indicate stress or relaxation, we continuously assessed participants' heart rate variability (HRV) using a heart rate monitor, specifically the *BIOPAC MP36* hardware (*Biopac Systems*, Inc.; California, USA). Data recording involved the use of Acq*Knowledge*® 5.0.8.1 (*Biopac Systems*, Inc.; California, USA), while the subsequent analysis was done using *Physiodata Toolbox* (version 0.6.3)[2]. Our focus was on examining alterations in mean RMSSD (root mean square of successive differences), mean HF-Power (mean high-frequency power), which are both indicative of parasympathetic activity, and LF/HF-Ratio (low frequency to high frequency power ratio), which reflects the interplay between sympathetic and parasympathetic activity within the autonomic nervous system. An increase in parasympathetic activity typically occurs during periods of relaxation and recovery, leading

[2] Sjak-Shie, E. E. (2022). PhysioData Toolbox (Version 0.6.3) [Computer software]. Retrieved from https://PhysioDataToolbox.leidenuniv.nl

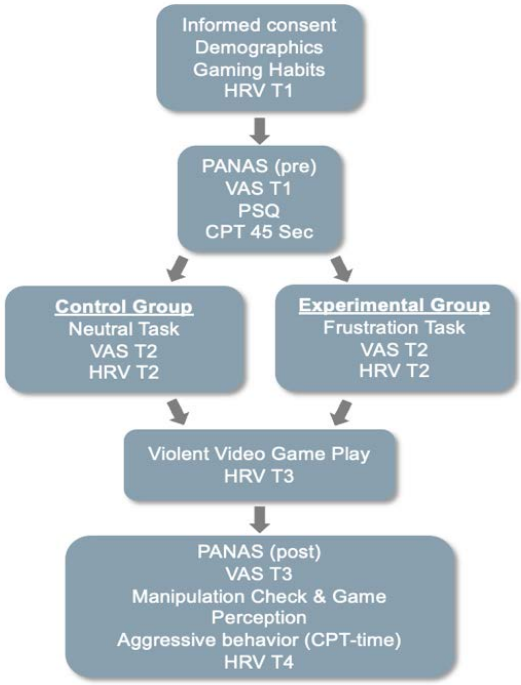
to an increase in RMSSD and HF-Power. Similarly, a lower LF/HF-Ratio indicates a potential dominance of the rest and digest or relaxation response, through heightened parasympathetic activity and reduced sympathetic activity.
The monitoring of HRV involved four key measurements: during participants' completion of pre-test self-assessment questionnaires (T1); while participants engaged in the neutral vs. frustration-inducing cognitive task (T2); during gameplay (T3); and during participants' completion of post-test self-assessment questionnaires (T4).

2.3. Procedure

First, participants received information about the study and gave informed consent. Participants were assured of privacy and confidentiality, with collected data saved under a pseudo-anonymized code, ensuring non-identification, and the option to withdraw or request deletion during or after the experiment. After attaching ECG electrodes, baseline heart rate variability (HRV T1) was measured while participants answered demographic and gaming habit questions on a PC (see Figure 1). Baseline emotional states were evaluated using PANAS (pre) and a Visual Analog Scale (VAS1). The Pain Sensitivity Questionnaire (PSQ) was completed, followed by the Cold Pressor Task (CPT) where participants immersed their non-dominant hand in cold water for 45 seconds.
Subsequently, participants engaged in a neutral or frustration-inducing cognitive task with a two-minute time limit, during which their heart rate variability was measured again (HRV T2).
Post-cognitive task, emotional states were assessed (VAS2). Participants then received game instructions and played the mission *Avanti Savoia* from the VVG *Battlefield 1*. Heart rate variability was assessed during this 25-minute gaming session (HRV T3).
After gameplay, participants again reported their emotional state (VAS3). A final heart rate variability measurement was taken (HRV T4) while participants indicated their positive/negative affect (PANAS post), answered manipulation check questions, and allocated of a timeframe for the CPT for the next participant. After debriefing, participants removed the ECG cables and received compensation. They were then informed about the study's purpose and the standardized CPT duration.

Figure 1

Overall Study Design



Results

3.1. Subjective Frustration Levels

To test the first hypothesis, firstly, a one-way ANOVA showed that there was no difference between the experimental (M = 20.41; SD = 24.01) and the control condition (M = 19.64; SD = 25.21) for their baseline subjective frustration level, $F(1, 52) = 0.01$, $p = .909$, $\eta^2 < .01$.
After the cognitive task, there was a significant difference between the experimental (M = 45.66; SD = 24.82) and the control condition (M = 15.92; SD = 15.88) for subjective frustra-

tion, Welch's $F(1, 48.23) = 28.22$, $p < .001$, $\eta^2 = .34$.[3]

Lastly, after gameplay, there was again no difference between the experimental ($M = 33.83$; $SD = 25.52$) and the control condition ($M = 27.92$; $SD = 21.75$) for subjective frustration, $F(1, 52) = 0.82$, $p = .368$, $\eta^2 = .02$.

A mixed ANOVA was calculated with condition as between-subjects factor (experimental condition vs. control condition) and measurement occasion of subjective frustration as within-subjects factor (baseline vs. after cognitive task vs. after gameplay). The significant within-subjects effect, $F(2, 104) = 6.57$, $p = .002$, $\eta^2_p = .11$, was qualified by a significant interaction effect, $F(2, 104) = 10.07$, $p < .001$, $\eta^2_p = .16$.

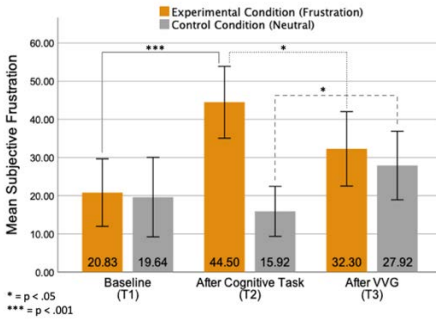
Contrasts showed that there was a significant interaction from baseline to after the cognitive task, $F(1, 52) = 17.26$, $p < .001$, $\eta^2_p = .25$ as well as from after the cognitive task to after gameplay, $F(1, 52) = 13.66$, $p = .001$, $\eta^2_p = .21$. For participants in the experimental condition, there was a significant increase in frustration from baseline ($M = 20.41$; $SD = 24.01$) to after the cognitive task ($M = 45.66$; $SD = 24.82$), $F(1, 28) = 20.33$, $p < .001$, $\eta^2_p = .42$, and a significant decrease from after the cognitive to after gameplay task ($M = 33.83$; $SD = 25.52$), $F(1, 28) = 6.07$, $p = .020$, $\eta^2_p = .18$ (see Figure 2).

For participants in the control condition, there was no change from baseline ($M = 19.64$; $SD = 25.21$) to after the cognitive task ($M = 15.92$; $SD = 15.88$), $F(1, 24) = 0.98$, $p = .333$, $\eta^2p = .04$. However, they experienced a significant increase in subjective frustration from after the cognitive task to after gameplay ($M = 27.92$; $SD = 21.75$), $F(1, 24) = 8.38$, $p = .008$, $\eta^2_p = .26$.

[3] Levene's test of homogeneity of variances was significant ($p = .002$) and therefore, Welch's test was used.

Figure 2

Mean Subjective Frustration in Experimental Group and Control Group



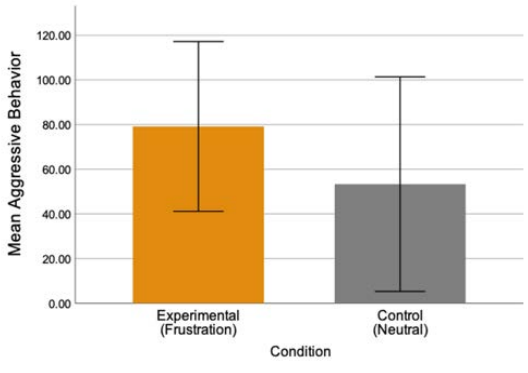
Note. Error Bars: 95% CI

3.2. Aggressive Behavior

A one-way ANOVA did not reveal a significant difference for aggressive behavior (i.e., time in the CPT allocated to next participant) between the experimental condition ($M = 79.52$; $SD = 103.63$) and the control condition ($M = 53.36$; $SD = 116.41$), $F(1, 52) = 0.76$, $p = .386$, $\eta^2 = .01$ (see Figure 3).

Figure 3

Mean Aggressive Behavior in Experimental Group and Control Group



Note. Error Bars: 95% CI

3.3. Physiological Relaxation

Mean RMSSD. A Mixed ANOVA with measurement occasion (RMSSD at T1 vs. RMSSD

at T2 vs. RMSSD at T3 vs. RMSSD at T4) as repeated measure and condition as between-subjects factor (experimental vs. control condition) was calculated. The results revealed that there was only a significant within-subjects effect, $F(3, 144) = 5.48$, $p = .001$, $\eta^2_p = .10$. Bonferroni corrected pairwise-comparisons showed a significantly higher mean RMSSD in ms at T4 (After VVG) compared to T1 (Baseline) ($p = .005$) and T2 (During Cognitive Task) ($p = .009$). Although no significant difference was observed from T2 to T3 ($p = .335$), the overall increase in RMSSD (see Table 2 and Figure 4) suggested heightened physiological relaxation from Baseline to after VVG. There was no significant between-subjects ($F(1, 48) = 0.46$, $p = .499$, $\eta^2_p = .10$) or interaction effect ($F(3, 144) = 1.35$, $p = .262$, $\eta^2_p = .03$).

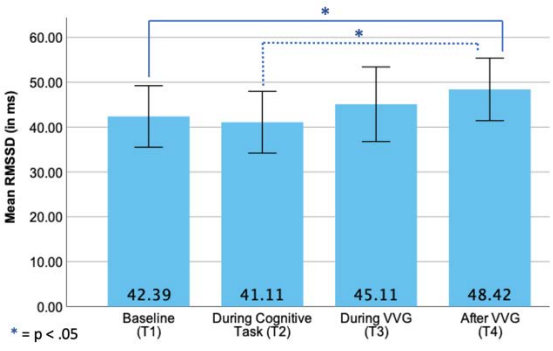
Table 2

Descriptive Statistics for RMSSD in ms

		Baseline (T1)		During Cognitive Task (T2)		During VVG (T3)		After VVG (T4)	
	N	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Experimental	27	44.17	25.04	44.08	26.34	48.77	33.02	48.63	24.17
Control	23	40.30	23.44	37.61	21.56	40.81	24.10	48.18	25.52
Total	50	42.40	24.15	41.11	24.24	45.11	29.25	48.42	24.54

Figure 4

Mean RMSSD in ms across timepoints



Note. Error Bars: 95% CI

Mean high Frequency Power (in ms2). A Mixed ANOVA was calculated, with measurement occasion (HF-Power at T1 vs. HF-Power at T2 vs. HF-Power at T3 vs. HF-Power at T4) as repeated measure and condition as between-subjects factor (experimental vs. control condition).

Mauchly's test of sphericity indicated a violation of the assumption ($\chi^2(5) = .36$, $p < .001$). Therefore, the degrees of freedom were adjusted using the Greenhouse-Geisser correction ($\varepsilon = .65$).

The results revealed that there was only a significant within-subjects effect, $F(1.84, 86.54) = 4.42$, $p = .017$, $\eta^2_p = .07$, while there was no significant between-subjects ($F(1, 47) = 0.81$, $p = .373$, $\eta^2_p = .02$) and no significant interaction effect ($F(1.84, 86.54) = 1.66$, $p = .198$, $\eta^2_p = .03$).

Pairwise comparisons revealed a significant increase in mean HF-Power at T4 (After VVG) compared to T2 (During Cognitive Task) ($p = .010$). Additionally, mean HF-Power was significantly higher fat T3 (During VVG) compared to T2 (During Cognitive Task) ($p = .022$). This suggests heightened physiological relaxation during and after VVG, irrespective of condition (see Table 3 and Figure 5).

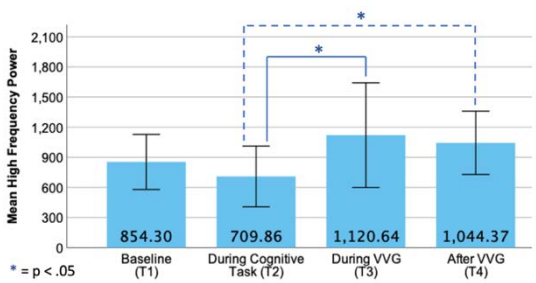
Table 3

Means and Standard Deviations for High Frequency Power in ms²

		Baseline (T1)		During Cognitive Task (T2)		During VVG (T3)		After VVG (T4)	
	N	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Experimental	27	947.30	1057.13	830.39	1234.08	1399.06	2226.71	1094.05	1119.52
Control	23	740.17	821.39	516.93	785.37	778.95	1085.38	983.40	1093.92
Total	50	854.30	954.64	709.86	1054.98	1120.64	1816.10	1044.37	1097.96

Figure 5

Means HF-Power in ms² across timepoints



Note. Error Bars: 95% CI

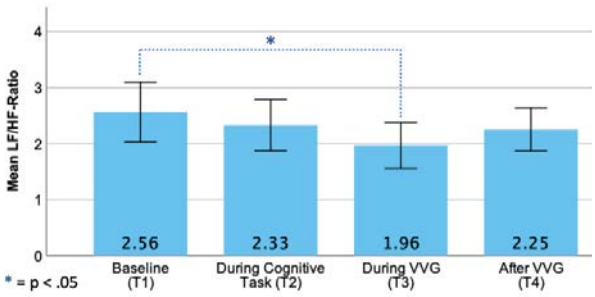
Mean LF/HF-Ratio. *The LF/HF-Ratio reflects the interplay between sympathetic and parasympathetic activity. A lower LF/HF-Ratio indicates heightened parasympathetic activity and reduced sympathetic activity, indication a relaxation response. A Mixed ANOVA was*

calculated, with measurement occasion (LF/HF-Ratio at T1 vs. LF/HF-Ratio at T2 vs. LF/HF-Ratio at T3 vs. LF/HF-Ratio at T4) as repeated measure and condition as between-subjects factor (experimental vs. control condition).
Mauchly's Test of Sphericity indicated that the assumption of sphericity had not been violated, $\chi^2(5) = .89$, $p = 0.36$. The results revealed that there was only a significant within-subjects effect, $F(3, 141) = 3.23$, $p = .024$, $\eta^2_p = .06$, while there was no significant between-subjects ($F(1, 47) = 0.003$, $p = .958$, $\eta^2_p < .01$) and no significant interaction effect ($F(3, 141) = 1.19$, $p = .32$, $\eta^2_p = .03$).
Pairwise comparisons revealed a significant decrease in the mean LF/HF-Ratio from T1 (Baseline) to T3 (During VVG) ($p = .037$), see Table 4 and Figure 6. This suggests greater physiological relaxation during VVG than at baseline, irrespective of condition.

Table 4
Means and Standard Deviations for LF/HF-Ratio

		Baseline (T1)		During Cognitive Task (T2)		During VVG (T3)		After VVG (T4)	
	N	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Experimental	27	2.72	2.22	2.30	1.66	1.81	1.38	2.23	1.29
Control	23	2.37	1.32	2.36	1.54	2.15	1.49	2.27	1.43
Total	50	2.56	1.86	2.33	1.59	1.97	1.42	2.25	1.34

Figure 6
Mean LF/HF-Ratio across timepoints



Note. Error Bars: 95% CI

3.4. Perception of Game and Cognitive Task

Familiarity with the VVG. In terms of familiarity with the VVG, no significant difference emerged between the experimental ($M = 1.87$; $SD = 1.38$) and control condition ($M = 1.64$; $SD = 1.19$), $F(1, 53) = 0.42$, $p = .552$, $\eta^2 < .01$. Participants in both conditions indicated the same level of (un)familiarity with *Battlefield 1*.

Game Characteristics. Regarding game perception, there were no differences between both conditions for any of the items (e.g., perceiving the game as frustrating / stressful / easy / fun / brutal / meaningful / difficult, perceived difficulty of the game's controls and satisfaction with gaming performance), $F \leq 1.84$, $p \geq .181$.

Immersiveness and Empowerment. There were no differences between both conditions regarding immersion in the game, identification with the character they played or the perception of feeling powerful in the game, $F \leq 0.68$, $p \geq .425$.

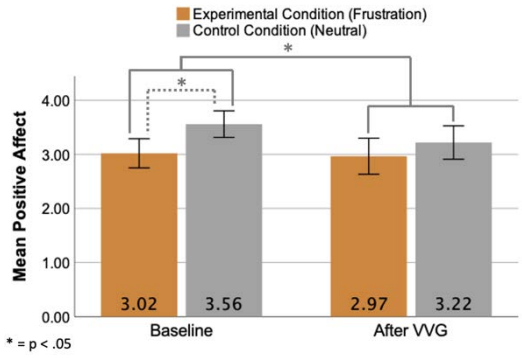
Cognitive Task-Related Frustration. A one-way ANOVA revealed a significant difference in the perception of frustration related to performance in the cognitive task between the experimental ($M = 3.43$; $SD = 1.41$) and control conditions ($M = 2.04$; $SD = 1.10$), $F(1, 53) = 16.25$, $p < .001$, $\eta^2 < .24$ As anticipated, participants in the experimental condition reported a higher level of frustration about their performance in the cognitive task than the control group. Additionally, there was a significant difference in frustration levels concerning the perceived difficulty of the cognitive task, between the experimental ($M = 2.90$; $SD = 1.24$) and control conditions ($M = 1.68$; $SD = .95$), Welch's $F(1, 52.61) = 17.09$, $p < .001$, $\eta^2 = .24$.[4]

3.5. Positive and Negative Affect

[4] Levene's test of homogeneity of variances was significant ($p = .024$) and therefore, Welch's test was used.

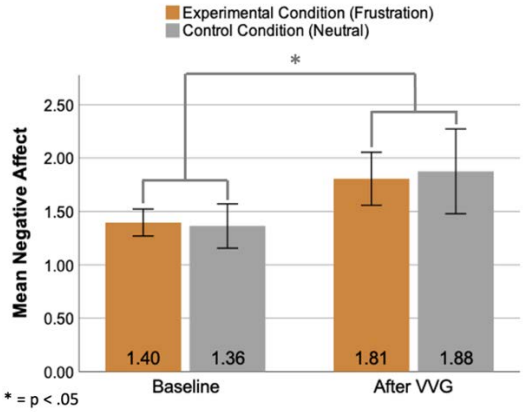
POSITIVE AFFECT. Firstly, a one-way ANOVA revealed that there was a significant difference between the experimental and control condition regarding positive affect at baseline, $F(1,53) = 8.90$, $p = .004$, $\eta^2 = .14$. The experimental group reported significantly lower levels of positive affect ($M = 3.02$, $SD = .72$) than the control group ($M = 3.56$, $SD = .59$). After the VVG, there was no significant difference between both conditions regarding positive affect $F(1,53) = 1.27$, $p = .265$, $\eta^2 = .23$.
A Mixed ANOVA was calculated with measurement occasion as repeated measure (PA at Baseline vs. PA after the VVG) and condition as between-subjects factor (experimental vs. control condition). There was a significant within-subjects effect, showing a significant decrease in PA from Baseline ($M = 3.27$, $SD = .72$) to after the VVG ($M = 3.08$, $SD = .83$), $F(1,53) = 4.55$, $p = .038$, $\eta^2_p = .08$, *see Figure 7.* There was also a significant between-subjects effect ($F(1,53) = 4.74$, $p = .034$, $\eta^2_p = .08$), suggesting that there is a notable difference in the average PA between conditions. There was no significant interaction effect ($F(1,53) = 2.42$, $p = .126$, $\eta^2_p = .44$).

Figure 7
Mean Positive Affect at Baseline and After VVG



NEGATIVE AFFECT. Firstly, a one-way ANOVA revealed that there were no significant differences in negative affect at baseline between the experimental group ($M = 1.40$, $SD = .34$) and the control group ($M = 1.36$, $SD = .50$), $F(1,53) = 0.08$, $p = .775$, $\eta^2 < .01$. After the VVG, there were also no significant differences in negative affect between the experimental group ($M = 1.81$, $SD = .67$) and the control group ($M = 1.88$, $SD = .96$), $F(1,53) = 0.10$, $p = .754$, $\eta^2 < .01$.
A Mixed ANOVA with measurement occasion (NA at Baseline vs. NA after the VVG) as repeated measure and condition as between-subjects factor (experimental vs. control condition) was calculated. Results revealed that there was a significant within-subjects effect, showing a significant increase of overall NA from Baseline ($M = 1.38$, $SD = .42$) to after the VVG ($M = 1.83$, $SD = .81$), $F(1,53) = 29.78$, $p < .001$, $\eta^2_p = .36$, *see Figure 8.* There was no significant between-subjects effect ($F(1,53) = 0.01$, $p = .906$, $\eta^2_p < .01$) and no significant interaction effect ($F(1,53) = 0.36$, $p = .549$, $\eta^2_p = .01$), suggesting that the pattern of change in NA is similar for participants in both conditions.

Figure 8
Mean Negative Affect at Baseline and After VVG



3.6. Pain Sensitivity

A one-way ANOVA showed no significant difference in mean pain sensitivity between both the experimental ($M = 4.09$, $SD = 1.41$) and the control condition ($M = 4.32$, $SD = .1.90$), $F(1,53) = 0.27$, $p = .605$, $\eta^2 < .01$. The absence of differences suggests that there was consistent pain sensitivity to the cold across conditions.

Discussion

This experimental study aimed to investigate the cathartic potential of playing VVGs by exploring their effects on frustration, aggressive behavior, and physiological relaxation.

Subjective Frustration. We hypothesized that playing a VVG after a frustrating task is associated with a significant *decrease in subjective frustration levels.* Our data supported this hypothesis, showing *a significant decrease of subjective frustration in the experimental group from after the cognitive task to after playing the VVG.* However, the mean negative affect rose from pre-game to post-game, while the mean positive affect decreased. This indicates that the observed mood repair effect through the VVG might only apply to particular negative emotions, such as frustration. Therefore, our data only partially supports Mood *Management Theory, which suggests that media is an effective strategy for mood improvement. Nevertheless, the findings are consistent with the results of Tyack et al.'s work (2020), which indicated that playing a moderately violent video game after a frustration task led to mood improvement.*
It is also important to note that the control group, which initially participated in a neutral cognitive task and had much lower frustration levels at T2, exhibited an unexpected increase in subjective frustration after playing the VVG. This suggests that the emotional state preceding the exposure to virtual violence might exert a notable influence on post-game emotional experiences. The unexpected increase of frustration in the control group cannot be explained by differences between the two conditions regarding the perception of the game as frustrating, familiarity of the game, or gaming experience as this was controlled for.
However, an alternative explanation emerges. Engaging in the neutral task left participants in a non-frustrated, and therefore more positive mood than the experimental group. The contrast between this state and the emotionally impactful content of the VVG might have contributed to the observed rise in frustration levels among the participants, even if the game itself might not have been perceived as particularly frustrating. For the experimental group, the shift to the emotionally charged content of the VVG might not have been as intense due to the heightened frustration and arousal following the frustrating cognitive task. Instead, the VVG served as an outlet for pent-up frustration.
It is also important to note that unexpected challenges can cause an increase of frustration (Wurhofer et al., 2015). Participants in the control group may have approached the study with expectations that diverged from the actual study procedures. For instance, the VVG could have been more unexpected in the control group than in the experimental group, resulting in heightened frustration. Likewise, it is possible that frustration increased in the control group because participants harbored a fear of failure when entering the VVG (Britt & Janus, 1940), especially after experiencing the positive outcome of successfully completing the cognitive task. The negative emotional impact stemming from a fear of failure is often not offset by positive overall outcomes (Wong, 1979). Consequently, the control group might have experienced increased frustration, even if the overall satisfaction with gaming performance was equivalent in both conditions.
Lastly, specific personality traits may interact with the exposure to virtual violence, impacting frustration levels. For instance, Wagener et. al (2024) demonstrated that higher scores on the Machiavellianism scale were associated with increased physiological relaxation after engaging in a VVG. Integrating measures of personality traits and examining their interplay with the VVG content could offer valuable insights into the responses observed in this study.

Aggressive behavior. We hypothesized that engaging in a VVG following a frustrating task would lead to a substantial decrease in levels of aggressive behavior.
According to the frustration-aggression hypothesis, frustration is likely to lead to aggressive behavior. Therefore, one would expect that the experimental group, which experienced higher frustration levels before playing the VVG, exhibits higher levels of aggressive behavior compared to the control group. However, the study did not find a significant differ-

ence in aggressive behavior between the two groups after playing the VVG.
One possible explanation of these findings is that despite the higher frustration levels in the experimental group before the VVG, the act of playing the game might have served as a release or outlet for their frustration, resulting in a comparable post-game frustration level between the experimental and control groups. This suggests that engaging in virtual violence, rather than exacerbating frustration into aggression, might have indeed had a cathartic effect on frustration and subsequent aggression. Therefore, the study's data seems to support the frustration-aggression hypothesis, as well as Mood Management Theory.
However, it is important to note that aggressive behavior was only assessed at one timepoint (after the VVG). For future research, aggressive behavior should be assessed at multiple time points (before and after playing the VVG) without inducing testing effects. By doing so, it becomes feasible to examine and compare aggression levels both within and between both groups. This would enable a more comprehensive analysis of changes in aggression and the intricate interplay it has with frustration.
Additionally, the immediate post-gameplay assessment may not fully encapsulate the complexity of how frustration and aggression after the VVG unfold over time. Differences in aggressive behavior might require a more extended period or different conditions to manifest.

Physiological relaxation. We hypothesized that playing a VVG leads to physiological relaxation. The analyses indicated that engaging in a VVG leads to increased parasympathetic activity, suggesting greater physiological relaxation (Nakao, 2019), particularly during and after playing the VVG. Therefore, our data further supports the results of Wagener et. al's (2024) work that discovered that VVGs can have a calming effect.
In the experimental group, the observed increase physiological relaxation during and after the VVG could be attributed to a in frustration and aggressive arousal. However, as the control group experienced an increase in frustration, while also experiencing physiological relaxation. Therefore, it is possible that engaging in virtual violence counteracts the manifestation of the physiological effects of frustration.

One notable limitation of the study is that the intended sample size was not fully achieved due to the limited timeframe of the EMPRA, limiting the power of the study. Additionally, a more extensive and diverse selection of VVGs and frustrating tasks could provide a better understanding of the full range of potential effects. While the data shows that participants in the experimental group perceived the frustrating task as frustrating, there is a question regarding its potency. There might have been stronger physiological effects with different or more intense cognitive tasks. Furthermore, this study focuses on short-term effects, leaving the longer-term consequences of engaging in VVGs unexplored. A more comprehensive investigation into the persistence and evolution of these effects over time is necessary for a thorough understanding of the phenomena under study.

While previous research has often focused on the potential negative effects of violent content, the present study suggests that, in the short term, engaging in virtual violence may alleviate frustration and increase physiological relaxation. Mood Management Theory was supported by a notable decrease in subjective frustration in the experimental group after engaging in the VVG. However, an unexpected increase in subjective frustration was observed in the control group after playing the VVG, suggesting that the emotional state preceding exposure to virtual violence significantly influences post-game emotional experiences. Consequently, the cathartic potential of VVGs may be influenced by various factors and nuances. The lack of a significant difference in aggressive behavior between conditions after the VVG might be due to the absence of a significant difference in frustration, supporting the proposed link in frustration-aggression hypothesis.
The hypothesis that playing a VVG leads to physiological relaxation was supported by the data. Both groups exhibited greater physiolog-

ical relaxation during and after engaging in the VVG, challenging common assumptions about the stress-inducing nature of VVGs.

Overall, the findings suggest that an integration of VVGs into contexts where frustration reduction is the desired outcome might be beneficial. However, the pre-game emotional context significantly influences post-game experiences, suggesting that interventions could be more effective if they consider the initial mood of the participants.

References

Aitken, R. C. (1969). Measurement of feelings using visual analogue scales. *Proceedings of the Royal Society of Medicine*, *62*(10), 989–993.

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., Rothstein, H. R., & Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in Eastern and Western countries: A meta-analytic review. *Psychological Bulletin*, *136*(2), 151–173. https://doi.org/10.1037/a0018251

Bond, A., & Lader, M. (1974). The use of analogue scales in rating subjective feelings. *British Journal of Medical Psychology*, *47*(3), 211–218. https://doi.org/10.1111/j.2044-8341.1974.tb02285.x

Breyer, B., & Bluemke, M. (2016). Deutsche Version der Positive and Negative Affect Schedule PANAS (GESIS Panel). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)*. https://doi.org/10.6102/ZIS242

Britt, S. H., & Janus, S. Q. (1940). Criteria of frustration. *Psychological Review*, *47*(5), 451–470. https://doi.org/10.1037/h0061381

Calvert, S. L., Appelbaum, M., Dodge, K. A., Graham, S., Nagayama Hall, G. C., Hamby, S., Fasig-Caldwell, L. G., Citkowicz, M., Galloway, D. P., & Hedges, L. V. (2017). The American Psychological Association Task Force assessment of violent video games: Science in the service of public interest. *American Psychologist*, *72*(2), 126–143. https://doi.org/10.1037/a0040413

Dollard, J., Doob, L. W., Miller, N. E. (Neal E., Mowrer, O. H., Sears, R. R. (Robert R., & Yale university. Institute of human relations. (1939). *Frustration and aggression*. New Haven, Pub. for the Institute of human relations by Yale University Press; London, H. Milford, Oxford University Press. http://archive.org/details/frustrationaggre00doll

European Commission, Directorate-General for Communications Networks, Content and Technology. (2023). *Understanding the value of a European video games society: Final report*. Publications Office of the European Union. https://data.europa.eu/doi/10.2759/332575

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

Ferguson, C. J. (2007). Evidence for publication bias in video game violence effects literature: A meta-analytic review. *Aggression and Violent Behavior*, *12*(4), 470–482. https://doi.org/10.1016/j.avb.2007.01.001

Ferguson, C. J. (2010). Blazing angels or resident evil? Can violent video games be a force for good? *Review of General Psychology*, *14*(2), 68–81. https://doi.org/10.1037/a0018941

Ferguson, C. J. (2020). Aggressive video games research emerges from its replication crisis (Sort of). *Current Opinion in Psychology*, *36*, 1–6. https://doi.org/10.1016/j.copsyc.2020.01.002

Geen, R. G., & Quanty, M. B. (1977). The catharsis of aggression: An evaluation of a hypothesis. In *Advances in Experimental Social Psychology* (Vol. 10, pp. 1–37).

Elsevier. https://doi.org/10.1016/S0065-2601(08)60353-6

Gündoğdu, S., Çolak, Ö. H., Doğan, E. A., Gülbetekin, E., & Polat, Ö. (2021). Assessment of mental fatigue and stress on electronic sport players with data fusion. *Medical & Biological Engineering & Computing*, *59*(9), 1691–1707. https://doi.org/10.1007/s11517-021-02389-9

Lee, E.-J., Kim, H. S., & Choi, S. (2021). Violent video games and aggression: stimulation or catharsis or both? *Cyberpsychology, Behavior, and Social Networking*, *24*(1), 41–47. https://doi.org/10.1089/cyber.2020.0033

Liu, S., Kaufmann, C., Labadie, C., Ströhle, A., Kuschpel, M. S., Garbusow, M., Hummel, R., Schad, D. J., Rapp, M. A., Heinz, A., & Heinzel, S. (2019). Short-term effects of video gaming on brain response during working memory performance. *PLOS ONE*, *14*(10), e0223666. https://doi.org/10.1371/journal.pone.0223666

Luria, R. E. (1975). The validity and reliability of the visual analogue mood scale. *Journal of Psychiatric Research*, *12*(1), 51–57. https://doi.org/10.1016/0022-3956(75)90020-5

Nakao, M. (2019). Heart rate variability and perceived stress as measurements of relaxation response. *Journal of Clinical Medicine*, *8*(10), 1704. https://doi.org/10.3390/jcm8101704

Prescott, A. T., Sargent, J. D., & Hull, J. G. (2018). Metaanalysis of the relationship between violent video game play and physical aggression over time. *Proceedings of the National Academy of Sciences*, *115*(40), 9882–9888. https://doi.org/10.1073/pnas.1611617114

Reinecke, L. (2016). Mood Management Theory. In P. Rössler, C. A. Hoffner, & L. Zoonen (Eds.), *The International Encyclopedia of Media Effects* (1st ed., pp. 1–13). Wiley. https://doi.org/10.1002/9781118783764.wbieme0085

Ruscheweyh, R., Marziniak, M., Stumpenhorst, F., Reinholz, J., & Knecht, S. (2009). Pain sensitivity can be assessed by self-rating: Development and validation of the Pain Sensitivity Questionnaire. *Pain*, *146*(1), 65–74. https://doi.org/10.1016/j.pain.2009.06.020

Schwabe, L., & Schächinger, H. (2018). Ten years of research with the socially evaluated Cold Pressor Test: Data from the past and guidelines for the future. *Psychoneuroendocrinology*, *92*, 155–161. https://doi.org/10.1016/j.psyneuen.2018.03.010

Tyack, A., Wyeth, P., & Johnson, D. (2020). Restorative play: Videogames improve player wellbeing after a need-frustrating event. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3313831.3376332

Wagener, G. L., Schulz, A., & Melzer, A. (2024). Games, hormones, and "dark" personalities: Dark tetrad and the effects of violent gaming on aggression, cortisol, and testosterone. *Physiology & Behavior*, *274*, 114421. https://doi.org/10.1016/j.physbeh.2023.114421

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Wechsler, D. (2012). *Wechsler Adult Intelligence Scale—Fourth Edition* [dataset]. https://doi.org/10.1037/t15169-000

Wong, P. T. (1979). Frustration, exploration, and learning. *Canadian Psychological Review / Psychologie Canadienne*, *20*(3), 133–144. https://doi.org/10.1037/h0081509

Wurhofer, D., Krischkowsky, A., Obrist, M., Karapanos, E., Niforatos, E., & Tscheligi, M. (2015). Everyday commuting: Prediction, actual experience and recall

of anger and frustration in the car. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 233–240. https://doi.org/10.1145/2799250.2799251

Acknowledgments

We express our sincere gratitude to the University of Luxembourg for providing the funding that made this study possible. Additionally, we extend our appreciation to our dedicated tutor, Gary Lee Wagener, whose guidance and support were invaluable throughout the research journey.

Body Vigilance, Anxiety, and Heart Rate Variability: Exploring the Links

Leila Agovic, Selma Dautbasic, Rachel Dos Santos Pinto,Katrin Oé, Linda Weisenstein

Supervisor : Doctoral Researcher Dominika Repcikova

Background. Inconsistent results have been found concerning the link between anxiety and low heart rate variability. Moreover, Body Vigilance is assumed to be associated with anxiety. In this paper an overview is provided of the link between anxiety, Body Vigilance and Heart rate variability.
Methods. Data collection involved administering the State & Trait anxiety Inventory for assessing anxiety levels and the Multidimensional Assessment of Interoceptive Awareness, Version 2, for evaluating body vigilance. Heart rate variability was measured using the Polar RS800CX device.
Conclusion. This study reveals a mediating link between anxiety symptoms and body vigilance. We have found that state and Trait anxiety are both negatively associated with attention regulation and self regulation. Further, higher scores on State anxiety implied more worry and Trait anxiety associates negatively with non-distracting behavior. The male population of our study showed anxiety symptoms that were significantly associated with reduced HRV as a function of high frequency activity.
Keywords : HRV, Anxiety, Attention regulation, Self regulation

1 Introduction

Anxiety disorders stand out as widespread mental health issues, imposing substantial financial and health-related challenges and posing a significant problem in society (Chalmers et al., 2014).
As per the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR), bodily sensations play a pivotal role in the manifestation of anxiety and serve as significant symptoms across various anxiety disorders. Physical manifestations of anxiety include sensations like lightheadedness, abdominal discomfort, accelerated heartbeat, muscle tightness, chest pressure, and other related symptoms (American Psychiatric Association , 2000). These symptoms are especially important when it comes to diagnosing panic disorder (PD): Research on cardiac perception in PD reveals that individuals experiencing panic attacks demonstrate heightened awareness of their heart rate. Patients with PD outperform controls in heart rate perception tests. Individuals who are more attentive to their elevated heart rate are likely to express greater somatic concerns and sensitivity to anxiety compared to those with lower perception. This underscores the importance of heightened internal cue awareness, commonly referred to as body vigilance, in the emergence of panic-related issues (Olatunji et al., 2007).
Biologically, anxiety results in an inhibition breakdown, leading to a reduced ability to regulate cognitive, emotional, behavioral, and physiological responses. This inefficiency results in a decrease in cardiac vagal outflow and a reduction in heart rate variability (HRV). As a result, anxiety disorders are linked to altered cardiovascular activity (Chalmers et al., 2014; Held et al., 2021).
Current research reveals that anxiety disorders elevate the risk of future cardiovascular disease (CVD) and premature mortality, even after controlling for factors like smoking, lifestyle, and socioeconomic status, regardless of any pre-existing medical conditions (Olatunji et al., 2007). In short, one potential mechanism connecting anxiety disorders to CVD is reduced heart rate variability (HRV) and low emotional regulation, which indicates impaired vagal function.
HRV stands for heart rate variability, a metric that assesses fluctuations in the time intervals

between successive heartbeats, also known as interbeat intervals. It serves as an indicator of the effectiveness of regulating autonomic balance, blood pressure, gas exchange, gut function, and the activity of the heart and vascular tone. Vascular tone pertains to the contractile activity of blood vessels, influencing blood pressure and potentially affecting facial muscles.

A faster heart rate results in less time between successive beats and therefore less IBI variance. In contrast, a slower heart rate allows for more variability in IBIs, resulting in a higher HRV.

Our study measures three different frequency domains: very low frequency (VLF), low frequency (LF) and high frequency (HF). In our study we only use a short recording period of 5 to 10 minutes.

Several research studies have shown a strong connection between very low frequency (VLF) power and the overall risk of mortality. However, there is a weak correlation between mortality and the power of low frequency (LF) or high frequency (HF).

Additionally, VLF has been linked to arrhythmic death, post-traumatic stress disorder (PTSD), elevated inflammation, and reduced levels of testosterone. The nervous system plays a crucial role in the VLF rhythm. The sympathetic nervous system (SNS) for instance influences the extent and alternance of its oscillations. Physical activity and various physiological factors, such as thermoregulation, the renin-angiotensin system, and endothelial influences, may affect VLF power in the heart. According to researchers, the peripheral nervous system (PNS) contributes to VLF. Subsequently, parasympathetic blockade almost eliminates VLF. Opposed to this phenomenon, sympathetic blockade does not have any influence in VLF. The VLF band varies between 0.0033 and 0.04 Hz, and within these 5 min, there should be approximately 0–12 complete periods of oscillation.

The LF is responsible for steady blood pressure. Also, it is not only guided by the SNS, moreover by the PNS and unspecified factors. Consequently, interactions between the SNS and the PNS are non-linear and non-reciprocal. The LF band varies between 0.04 to 0.15 Hz and should be measured during at least 2 minutes.

As for the HF, it is associated with parasympathetic activity and the respiratory cycle. Concerning the analysis, one should consider that HR accelerates during inhalation, because the vagal outflow is blocked. Concerning expiration, HR slows down because vagal outflow is reestablished, and Adrenocorticotropic hormone (ACTH) is released. Research has shown that lower HF power correlates with conditions such as stress, panic, anxiety or worry. HF is measured from 0.15 to 0.40 Hz and should be recorded for at least 1 minute (Shaffer & Ginsberg, 2017).

The third variable in our study is body vigilance and it is assessed by the body vigilance scale (BVS). Body vigilance describes the capability of the conscious awareness of their own body‘s signals is assumed to be a group of symptoms associated with anxiety disorders such as panic disorder (PD). Researchers assume that there is a correlation between individuals suffering from anxiety disorders and cardiac awareness (Olatunji et al., 2007).

Previous studies have reported inconsistent findings concerning the link between anxiety disorders and HRV, underscoring the necessity for further research. Our study seeks to explore the relationship among heart rate variability (HRV), anxiety, and the body vigilance scale (BVS). The primary focus is on understanding the correlation between HRV and mental health, specifically comparing individuals with high anxiety levels to those with low anxiety levels. We further hypothesize that body vigilance may operate as a mediator between altered HRV and anxiety symptoms. We anticipate that individuals experiencing anxiety will exhibit lower HRV. Therefore, we hypothesize that:

H1: Anxiety symptoms are linked to altered Body Vigilance.

H2: Anxiety symptoms are linked to altered HRV.

H3: Body Vigilance mediates the relationship between anxiety and symptoms.

2 Methods

2.1 Search criteria

44 Participants were recruited among students and research staff of the University of Luxembourg (25% male / 75% female, mean age = 25.7). We used flyers, social media, and student platforms to recruit participants. Additionally, participants were offered a lottery prize of a gift voucher. The exclusion criteria for our study were chronic diseases or current psychiatric treatment. All participants were required to sign an informed consent form and data protection statement before proceeding with the study.

2.2 Methods-Procedure

First, each participant received a survey by email to be filled out at home. Participants were required to read an information page before proceeding with the survey. Participants could choose between German or English versions of the survey based on their preference. The online survey included the short version items from the State and Trait Anxiety Inventory (Grimm & Jürgen, 2009)– and the Multidimensional Assessment of Interoceptive Awareness (Mehling WE et al., 2018), Version 2 (MAIA-2).

The questionnaires were distributed using LimeSurvey. The items from the State Trait Anxiety Inventory (STAI-S) include different items such as feelings associated with anxiety. Some examples are “I feel calm”, “ I feel secure”, or “I am tense”. For the STAI-S questionnaire,(Spielberger et al., 1983), we used a 7-point Likert scale (0 = almost never; 7 = almost always).

The Multidimensional Assessment of Interoceptive Awareness, Version 2 (MAIA-2) measures multiple dimensions of introspection connected to physiological symptom awareness. For the purposes of our study, we used four subscales, namely Not distracting, Not worrying, Attention regulation and Self regulation. The subscale Non distracting contains six items and specifies the propensity to deny or divert attention from painful or uncomfortable experiences. Not worrying contains five items indicating emotional distress or worrying about pain or discomfort. Attention regulation contains seven items which measure the ability to maintain and control attention to bodily sensations. Self regulation contains four items which show the capacity to control mental trouble by consideration to body sensations.

For the heart interbeat interval (IBI) measurement, participants first completed a resting phase, which consisted of sitting still for 10 minutes. This was to control for confounding which could be introduced by physical activity shortly before the measurement. Subsequently, we introduced the recording phase, where resting-state IBIs were recorded for 10 minutes using an electrode belt placed around the chest. The participants were allowed to read or look at their cellphone during the resting and recording phase, except for contents that could accelerate the heartbeat. All IBIs were recorded with the Polar RS800CX and Polar Pro Trainer 5.

To protect participants’ privacy, we used pseudonymization for all demographic and health data.

3 Statistics

After collecting all data, we first removed erratic beats from the IBIs. All post-hoc analyses were performed using IBM SPSS Statistics (Version 28) and Artiifact (Kaufmann et al., 2011).

As for the survey data, we first explored the bivariate correlations and linear regression among the following items: Trait anxiety (TraitAnx), State anxiety (StateAnx), Attention regulation (AttReg), Self regulation (SelfReg), Non distracting (NonDis) and Not worrying (NotWor).
Also, we analyzed the correlation among survey and heart data, such as LF, HF and VLF.
Our survey sample showed a highly symmetric normal distribution. High and very low frequency values in the male sample demonstrate a normal distribution, but the low frequency

and LF/HF ratio values illustrate right-skewed distribution.

4 Results

4.1 Bivariate correlation (including only survey data)

TraitAnx and SelfReg showed a high negative correlation with Pearson's *r* = -.527 (in our analysis we used solely Pearson's *r*), *sig*. = .000. TraitAnx showed a moderate negative correlation with AttReg (*r* = -.351, *sig*. = .020) and NonDis (*r* = -.355, *sig*. = .018).
StateAnx correlates highly negatively with both SelfReg (*r* = -.489, *sig*. = .001) and AttReg (*r* = -.411, *sig*. = .006).
These results suggest that there is a significant link between body vigilance and anxiety.

4.2 Linear & Multiple Regression including Sobel-Test (including only survey data)

Our correlation analysis of the survey data revealed two positive linear regressions, TraitAnx with SelfReg and AttReg with R^2 = 0.401 (Fig.1) and StateAnx with SelfReg and AttReg with R^2 = 0.408 (Fig.2). Direct effects between TraitAnx and StateAnx as well as SelfReg and AttReg were confirmed by Sobel test (*p* = .038). Trait and StateAnx (*p* = .002) are significant mediators for AttReg and SelfReg, thus proving that the positive correlation between AttReg and SelfReg is caused by TraitAnx and StateAnx.
In addition, a negative relationship between StateAnx and NotWorr with R^2 = 0.102 (Fig.3) and TraitAnx and NonDis with R^2 = 0.126 (Fig.4) was found.

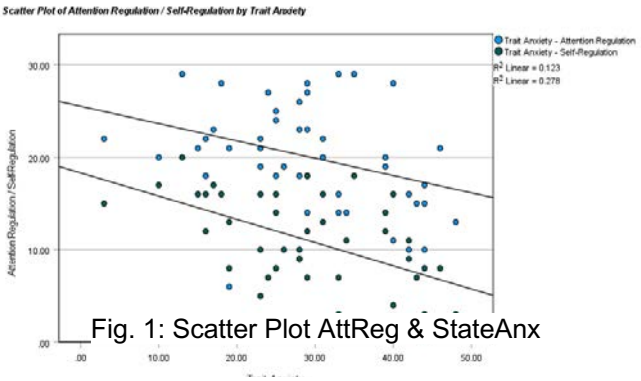


Fig. 1: Scatter Plot AttReg & StateAnx

Note. The higher AttReg and SelfReg, the lower TraitAnx.

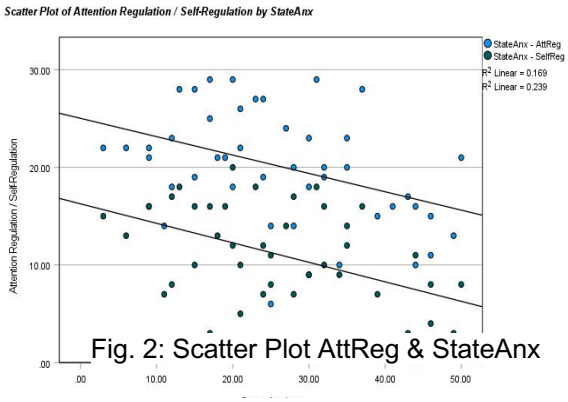


Fig. 2: Scatter Plot AttReg & StateAnx

Note. The higher the StateAnxiety the less control over Selfreg and Attreg.

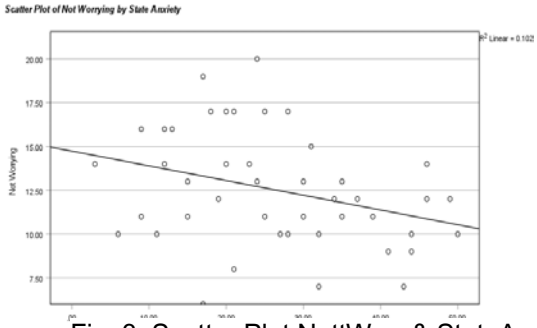


Fig. 3: Scatter Plot NottWorr & StateAnx

Note. The higher the StateAnx, the lower NotWorr.

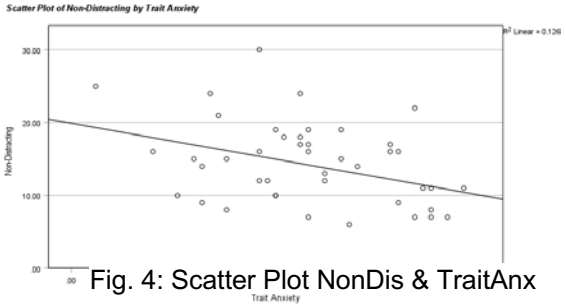


Fig. 4: Scatter Plot NonDis & TraitAnx

Note. The higher the TraitAnx, the lower Non-Dis.

4.3 Bivariate correlation (including survey & heart rate data)

There is a negative correlation between HF and StateAnx (p = .02; r = -.678) Another negative correlation has been found between HF and TraitAnx (p = .01; r = -.761).

4.4 Linear regression / correlation (including survey & heart rate data)

The direct effects between HF and StateAnx with R^2 = .459 (Fig.5) and HF and TraitAnx with R^2 = .579 (Fig.6) were analyzed by linear regression. This analysis excludes the data of women as there were only significant results for men. The results suggest a negative relation, so that lower HF is associated with higher TraitAnx and StateAnx.

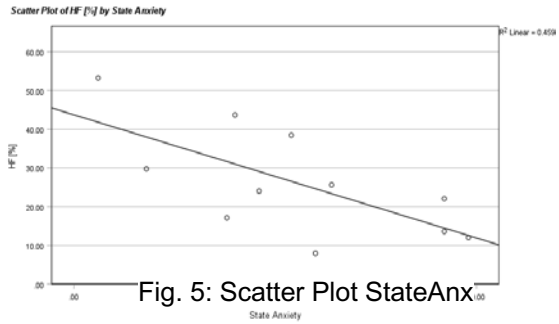


Fig. 5: Scatter Plot StateAnx

Note. The higher the State anxiety, the lower the HF.

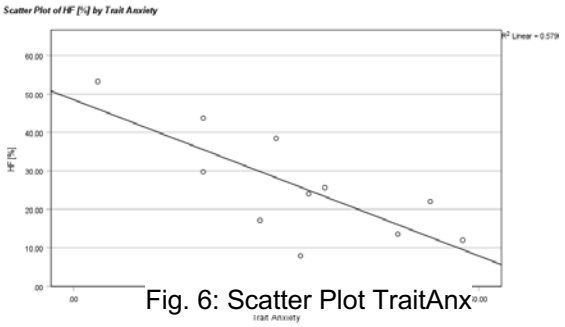


Fig. 6: Scatter Plot TraitAnx

Note. The higher the Trait anxiety, the lower the HF.

5 Discussion

The results of our study suggest that anxiety disorders are associated with significant reductions in heart rate variability (HRV), especially among men, with effects of a moderate effect size. The gender-specific findings indicate different impacts of anxiety on physiological symptoms among women and men.
Moreover, we found that BVS is associated with anxiety; we could neither confirm, nor accept that BVS mediates the relationship between anxiety and symptoms.

Thus, the purpose of our study was to test three hypotheses:

1. Hypothesis 1: Anxiety Symptoms are linked to altered Body Vigilance.

This hypothesis could be confirmed, as both Trait and State anxiety could be found as mediators of Attention regulation and Self regulation. These cognitive processes have been identified as being part of body vigilance. This implies that the correlation between the capacity for Self regulation and Attention regulation is depending upon the presence of Trait and State anxiety. St
ate and Trait anxiety serve as a mediating factor in bringing to life the association between the independent and dependent variables, providing an explanation for the occurrence of the relationship between Attention regulation and Self regulation. As all correlations are ne-

gative, we conclude that if one item changes the other correlated item changes in the other direction. Anxiety and Attreg and SelfReg have a negative relation, this means the greater the anxiety the less effective is the Self regulation and Attention regulation.

2. ***Hypothesis 2: Anxiety Symptoms are linked to altered HRV.***

Our findings only partially confirmed this hypothesis. Significant associations were observed solely for male participants. Both Trait and State anxiety correlated with high-frequency (HF) HRV in men, while no significant correlations were observed for women. A low HRV leads to a deteriorated ability to adapt to rapidly changing situations in the environment. Studies have indicated that reduced high-frequency (HF) power is associated with states of stress, panic, anxiety, or worry (Shaffer & Ginsberg, 2017). Thus, the negative correlation indicates that the lower the HF, the higher anxiety. These results highlight especially a gender-specific impact of anxiety on HR, implicating to consider gender differences in the physiological interpretation of anxiety symptoms.

3. ***Hypothesis 3: Body Vigilance mediates the relationship between anxiety and symptoms.***

We found no mediation effects between body vigilance, anxiety and HRV in our sample.
The limitations of our study include a small and specific sample (psychology students, potentially professional athletes unknown), methodological constraints (short heart rate measurement, limited frequency testing), and time pressure, which may have resulted in sampling errors. Our power analysis with G-Power revealed satisfactory but suboptimal statistical power (actual power: 0.71 with a total sample size of 44).

Chalmers et al. (2014) explores the correlation between anxiety disorder and heart rate variability (HRV), yielding similar results. Their research also delves into various anxiety disorders such as panic disorder or PTSD, whereas our study is limited to anxiety states and traits.
Olatunji et al. (2007) also used, for their data, factor analysis to investigate the structure of the Body Vigilance Scale. They also measured a gender difference on body vigilance scores whereas we only found a difference in the heart rate values. Their second study identifies a significant association between body vigilance items and anxiety. Like our study, where a significant link between some body vigilance items and anxiety was observed. In contrast to our study which is limited to exploring the link between anxiety, body vigilance and HRV, their study also examines if body vigilance could lead to health and illness concerns. In contrast to our observations, the study by Held et al. (2021) found no significant main effect of gender on HR at baseline.
In summary, our research has shown significant associations between anxiety disorders and reduced heart rate variability (HRV). We have established a critical link between anxiety symptoms, and both altered body vigilance and HRV. Notably, our findings do not support the notion that body vigilance acts as a mediator in the relationship between anxiety and its symptoms. While these insights contribute to a more comprehensive understanding of the physiological aspects of anxiety, it is crucial to acknowledge the limitations within our study, including a relatively small and specific sample, potential methodological constraints, and time pressures that may have introduced unintended errors. Our power analysis, while indicating satisfactory statistical power, also highlights room for improvement. These insights contribute to a more comprehensive understanding of the physiological groundwork of anxiety, providing a foundation for future research and simpler interventions in the field of mental health.

6 References

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.). *Washington, DC: Author.*

Chalmers, Quintana, Abbott, and Kemp, (2014). *Anxiety disorders are associated with reduced heart rate variability: a meta-analysis.* Frontiers in psychiatry, 5, 80.

Grimm, Jürgen (Hg.) (2009).: State-Trait-Anxiety Inventory nach Spielberger. Deutsche Lang- und Kurzversion.- Methodenforum der Universität Wien. MF-Working Paper 2009/02.

Held, J., Vîslă, A., Wolfer, C., Messerli-Bürgy, N., & Flückiger, C. (2021). Heart rate variability change during a stressful cognitive task in individuals with anxiety and control participants. *BMC psychology*, *9*(1), 1-8.

IBM Corp. Veröffentlicht 2021. IBM SPSS Statistics für Windows, Version 28.0. Armonk, NY: IBM Corp.

Kaufmann, T., Sütterlin, S., Schulz, S., & Vögele, C. (2011). ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. *Behavior Research Methods*, 43(4).

Mehling WE, Acree M, Stewart A, Silas J, Jones A. (2018). The Multidimensional Assessment of Interoceptive Awareness, Version 2 (MAIA-2).

Olatunji, B. O., Deacon, B. J., Abramowitz, J. S., & Valentiner, D. P. (2007). Body vigilance in nonclinical and anxiety disorder samples: structure, correlates, and prediction of health concerns. *Behavior Therapy*, *38*(4), 392-401.

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in public health*, 258.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

7 Appendix

Principal Factor Analysis (PCA)

Based on the results of the factor analysis, the primary factor, SelfReg, explained 46.5% of the variability, while the second factor, TraitAnx, contributed an additional 19.5%. The remaining factors made smaller contributions to the structure of STAI-T, STAI-S, and MAIA. Therefore, we propose that the concepts of anxiety are best measured by the two factors, SelfReg and TraitAnx.
Considering eigenvalues greater than 1 and examining the scree plot, principal component analysis with varimax rotation suggests that the four extractable factors are associated with the dimensions of StateAnx, NonDis, NotWorr, and AttReg.

Analyzing absolute loadings greater than 0.4 reveals that items from the Self regulation and Attention regulation dimensions strongly load onto Factor 1, while only items from the Not worrying dimension load strongly onto Factor 2.
Furthermore, we theorize that Self regulation and Attention regulation, as well as Trait anxiety and State anxiety, form clusters within Factor 1, as each pair of dimensions shows similar correlations with Factor 1. No such clusters were observed for Factor 2.
However, caution is advised when interpreting unrotated solutions, as they depict loadings where the first factor explains the maximum variance (notably, most high loadings are concentrated in the first factor).

Luxemburger Experimentalpraktikum Journal (LEx), 18, 2024, 24-38

How do young people see the role of Instagram on body image?

Chiara Guinet-Venzi, Lis Krack, Katherine Marsh, Stacy Mauer, Sara Valerio de Sousa

Supervision: M.A. Kriti Kelkar, Asst. Prof. Dr. André Melzer

Research demonstrates that social media use is consistently and positively associated with negative body image (Fardouly & Vartanian, 2016). Instagram's influence on body image specifically however is still a new area of research. The purpose of the study was to explore Instagram's role in influencing body image of young people. As less is known about this topic, a qualitative approach was employed. Two face-to-face focus groups, one with all female participants and one with all male participants ($N = 9$, $n_f = 5$, $n_m = 4$), aged 19-25 years, were conducted. The data were transcribed and thematically analysed, separately for each group. Five main themes emerged for the male group: Instagram use, concept of body image, Instagram and body image, consequences of Instagram on body image and Instagram policy. The female group produced four main themes: use of Instagram, body image, role of social media and (gender) stereotypes. We discuss the themes with the help of psychological theories such as peer influence, self-objectification, cognitive dissonance, selective self-presentation, gender stereotypes or media literacy.
Key words: Instagram, body image, focus groups, social media, young people, qualitative analysis, protective strategies, positive body image

Introduction

Social media and body image are related. Body image and social media are linked because people tend to compare themselves to others, and because Instagram, for example, is a highly visual social media application, a negative and unrealistic body image can be displayed and seen by many people. There is also a concern because body image is how we perceive and feel about our bodies, and when this is out of balance it affects our well-being. By promoting unrealistic body ideals, people begin to perceive their bodies in a more negative way, which can lead to unhealthy behaviors such as the development of eating disorders (Sharifi et al., 2016). As a result of this concern, researchers investigated this phenomenon to find out which aspects had an impact on men and women regarding a negative body image (Fardouly & Vartanian, 2016). It has been shown that there are more images of thin women in media or magazines than they exist of men (Strahan et al., 2006). In addition, it has been proven that digital image editing leads to unrealistic body sizes and that diets are being criticized. At the same time, however, women are encouraged to look a certain way when being exposed to media. American girls and women for example gained weight in the last 30-40 years but became thinner in media (Grogan, 2021).

Moreover, the cultural norms for men's bodies are more flexible compared to women. Nevertheless, the pressure of society and adverts regarding male bodies lead to an increase of dissatisfaction and low self-esteem in men. There is even a cultural ideal for male beauty which includes men to be lean, muscular, have a masculine appearance and should look young and healthy. Additionally, it was stated that men can and should achieve this ideal (Grogan, 2021).

Depending on the content people follow on Instagram, the insecurities they face might differ, but in most cases, they are linked to their body image (Cohen et al., 2017). Since there is less known about male responses to body image and social media, we wanted to include an all-male focus group to get more insight on the less explored gender. Currently,

there are over 2.35 billion monthly active Instagram users (Ruby, 2023). As of January 2023, 13.9% of users were women aged 18-24 and 16.9% were men (Statista, 2023). Instagram has gained in range over the last couple of years. Therefore, we decided to further investigate the idea of Instagram use and body image. It was important to include men and women in the study to determine whether there were gender differences.

Theoretical Background

An increase of interest in the topic of Body image has been seen in research over the last decades. This kind of research has shown that mainstream media like social media has a negative impact on body image (Baker et al., 2019). For instance, the use of Instagram has shown that the platform increases factors like body dissatisfaction and also causes more social comparison and self-objectification (Baker et al., 2019).

With our study, we want to find answers on how young people see the role of Instagram on body image. We want to find out what impact Instagram has on the body image of young people. The data was collected with the help of two focus group discussions as mentioned in the methodology part. A focus group discussion is qualitative research to gather a range of opinions and perspectives on a specific topic. In our study, we used it for Exploratory, Explanatory, and Evaluation Research (Hennink & Leavy, 2014). The first group comprised four male participants and the second group comprised five female participants.

First, we need to define some key terms that are used throughout our study.

Body Image: Body image is a multidimensional construct consisting of perceptual, cognitive, affective, and behavioral components relating to the image that one forms of their body and appearance (Rusticus, 2014).

Influencers: Social media influencers are individuals who create and share content on social media platforms while they have built important relationships and connections with organizational stakeholders (Enke & Borchers, 2021).

Stereotype: A cognitive structure that is our knowledge, beliefs, and expectations about a social group of people (Jonas et al., 2007, p. 607).

Self-Objectification: Self-objectification is a procedure in which individuals, especially women view themselves as objects to be judged and looked at according to their appearance (Grippo & Hill, 2008).

The Tripartite Influence model (Thompson et al., 1999) proposes three sources of influence, namely parents, peers, and media. These influences contribute to the development of body image and disordered eating. Understanding the socio-cultural perspective from the Tripartite Influence Model is crucial. In focus group discussions, we can investigate participants' experiences with appearance-related pressures from classmates, family, and the media, and how these external factors alter their judgments of body image, contributing to thin-ideal internalization and comparisons (Keery et al., 2004).

In the qualitative context, Social Comparison Theory (Festinger, 1954) relates to participants' tendencies to compare themselves with others on Instagram especially when they perceive similarities. (Jonas et al., 2007). This can provide insights into how Instagram fosters social comparisons, influencing body image perceptions.

Objectification Theory (Fredrickson& Roberts, 1997) is relevant to understanding how constant exposure to Westernized societal norms on Instagram may impact women's body image (Grogan, 2021). In focus group discussions, we explore participants' experi-

ences of being sexualized or the emphasis on physical appearance and the internalization of beauty ideals influenced by Instagram.

In a qualitative exploration, the Social Cognitive Theory of Mass Communication (Bandura, 2001) informs the analysis of how participants model their behaviors and values based on Instagram content. This could explain how the symbolic environment of Instagram influences participants' perceptions of body shapes and sizes through the extensive modeling present in mass media.

The Employ Uses and Gratification Theory (Lazarsfeld and Stanton, 1944) could help to understand the motivations driving participants' engagement with Instagram. The theory focuses on how the gratifications that are sought and obtained, are addressing both psychological and social needs. Through the focus group discussions, we wanted to find out the specific ways individuals use Instagram to fulfill these needs, shedding light on the platform's role in shaping body image perceptions.

Methods

Participants

Participants were young people, respectively students from different cultural backgrounds currently undertaking an undergraduate or master's degree qualification at the University of Luxembourg. In total, our sample consisted of five women and four men (N = 9, n_f = 5, n_m = 4) aged between 19 to 25 years (M = 22, SD = 1.63). Furthermore, participants were all English speakers and Instagram users, who actively and regularly used the platform and expressed enthusiasm about partaking in a group discussion regarding the topic of Instagram's role on body image. A further inclusion criterion was written informed consent, which all recruited participants were able to give.

Recruitment

Convenience sampling was utilized to recruit participants. Flyers were distributed throughout the main university building, the Maison du savoir (MSA). Moreover, flyers were also displayed in the Maison des arts et des étudiants (MAE) and the student lounge. Each flyer contained information outlining the study title, the eligibility criteria, and an advertisement of the study incentive (i.e., a 15€ GoGift voucher and study credits for BAP students) along with the contact details, the email address discussbodyimage@gmail.com, which had been specifically set up for the purpose of this study. Furthermore, a Moodle call was sent out to all students at the University via email, containing more elaborate information about the study. Both the flyer and the Moodle call considered that the topic of discussion may be triggering to some, and persons should thus apply at their own discretion.

Students who expressed an interest in the study were subsequently emailed the exact schedule (date, time, and room number) and were asked to confirm their presence on that date to be sure of the number of participants.

For eligible students, the first females and males to confirm their presence were recruited. If a person wanted to participate on a specific date, but the slots had already been filled, they were offered the choice of either being assigned another spot on a different date or being kept as a reserve in case a slot reopened. The focus group discussions were carried out on two predetermined dates. Hence after that, recruitment ceased.

Data collection

Focus group discussions were deemed the most appropriate qualitative methodology to use for this study. The reason being that they can be used for academic purposes, specifically in this case for exploratory research. Since

there is still little knowledge on the influence of Instagram on body image, especially concerning men, this approach was deemed the most fruitful for uncovering more about the topic.
Additionally, we opted for focus group discussions for our research design, due to us intending to capture a range of views and experiences, wanting to encourage group interaction and debate, and considering it ideal for gaining community-level information (Hennink, 2014).

Each focus group was directed by a self-constructed topic guide (see Table 1), including a seven-phase process: (1) defining the term body image; (2) relating body image to Instagram; (3) how body image is displayed on Instagram; (4) the influence this had on participants; (5) how it changed their usage of Instagram; (6) preventing Instagram from affecting body image; (7) space for additional thoughts and feedback on the discussion. The topic guide and the question route were developed collaboratively. Project supervisor Kriti Kelkar first developed several drafts of the topic guide and the question route, which formed the basis of the finalized version. Questions evolved, and new questions emerged over the course of several months. Consulting project supervisor Dr. André Melzer reviewed the questions. As an example, a question which started out as “What are the ideals for bodies presented on the platform- Instagram?” was modified to “What kind of bodies are displayed on Instagram?”. The EMPRA students were tasked with brainstorming questions for each topic independently. Some questions which arose, such as “Do you think there’s a difference between the promoted body image for men and women on Insta?”, were then also integrated into the final topic guide.
The question route was designed to be direct, clear, and simple, promoting an informal environment. The questions were open-ended, free of technical language and neutral, to reduce interview bias (Kreuger, 1998).

Consequently, ethics approval for the study was sought by filling out the “Application for Ethics Approval to Student’s Research”, which encompassed personal details of all researchers, detailed information on the project, possible ethical implications, information on consent, anonymity, and storage of data. Ethical implications merely included the potential risk of minimal emotional and/or psychological toll due to the possibility that some participants may find the discussion triggering or relating to some personal experience. To keep this complication at bay, as mentioned before, a caution was placed on both the flyer and the Moodle call. Additionally, the option of withdrawal was given in case participants felt sudden and unexpected distress during the discussion. Regarding consent and anonymity, it was made clear in the ethics form that participants had to consent to the collecting and processing of personally identifiable information (audio- and video material) and that their names would be coded into pseudonyms for organisation of the data in the transcripts. As for storage of data the ethics form included that the data would be stored on password-protected computers, with only the researchers having access to them and that it would be kept for five years before deletion.

The topic guide and the general proceedings were reviewed and refined following a pilot focus group (N = 5). It was determined that additional examples needed to be included as probes or follow-ups to questions participants of the pilot study found difficult to answer (see Table 1). The examples were carefully chosen to avoid influencing participants’ responses in a specific direction. Furthermore, modifications were implemented in the seating plan based on the conclusion that the setting should foster a sense of intimacy, closeness, and safety. Thus, participants in both focus groups were arranged around a table, with only the project supervisor Kriti Kelkar as the moderator seated among them, while the other facilitators remained in the background. Data from

the pilot study focus group was not included in the final analysis of results, it was only used to ensure the ease of understanding of the questions and as a transcribing exercise to prepare the EMPRA students for data analysis.

Thus, focus group discussions took place on the 7th and 8th of November 2023 in seminar rooms on the fourth floor of the MSA and lasted for up to 90 min, with the men's group concluding sooner. The moderator led both focus groups, presenting the participants with questions from the topic guide. Sara Valerio de Sousa (for the male group) and Lis Krack (for the female group) were tasked with reading the introductory slides to the participants to ensure their knowledge of the duration, the audio- and video recording, the compensation, and their rights as a participant to withdraw at any time without any negative repercussions, and to receive their data if requested. In addition, within the introduction, participants were asked to remain respectful and were guaranteed a safe space where they could freely express their thoughts, feelings, and opinions. Other facilitators, namely Katherine Marsh (for the male group) and Lis Krack (for the female group), took detailed notes of discussion points, giving participants the opportunity to review what was said by them after the focus group concluded and to get a first idea of the data at hand. The notes were later also used to facilitate transcription in case of doubt. Stacy Mauer and Chiara Guinet-Venzi were assigned the task of constructing a plan of the setting, specifically to map out where the participants were sitting with their respective fictitious names to be sure of the assigned names in the transcripts being correct.

Before commencement of the focus groups, participants had no prior acquaintance with the moderator (i.e., Kriti Kelkar). The same applies to the other facilitators, however there was one exception. One male participant, with the fictitious name of Joseph, studies the BAP in the third semester and thus belongs to the same cohort as Stacy Mauer, Katherine Marsh and Sara Valerio de Sousa, who were present during the focus group which he participated in. Upon arrival, written informed consent and essential demographic information were obtained before the beginning of the study.

Table 1: Topic guide and the question route

Topic 1: Body image and its understanding/idea • When you hear the words Body Image, what comes to mind? [What shaped this concept of BI?] • When did this concept start to shape for you/develop for you? • If you must list the things that can affect BI, what all would it include
Topic 2: Instagram use and understanding of the application • How do you see Instagram use by young people in relation to their BI today?
Topic 3: Body Image <-> Instagram (features of Insta, content, posting behaviors) • How would you describe your activities on Instagram? – What all kinds of content do you post? Or what do you keep in mind while posting your own photos? (maybe when what kinds of photos do you prefer to post- what do you look out for?) (maybe give examples- full-body or mostly selfies) • What Instagram features do you make use of while posting on Instagram? • Personally, what do you think about the importance of the comments/reactions/replies that you get on your posts? (Content seen- body related) • What all kinds of content do you come across on Instagram? • What kind of bodies are displayed on Instagram? • What words (adjectives) do you

use to describe the bodies that are shown on Instagram?
- Do you think there is a difference between the promoted body image of men and women on Insta? If so, how?
- Would you like to share any thoughts about the captions or comments you see on these posts or if they differ based on gender?

(Shift towards personal experience)
- What do you think about Instagram's role in affecting your very own body image?

Topic 4: Signs for anything different - Signifiers/Indicators (emotions, cognition, behaviors, physical sensations)
- When you are exposed to these "-" images, what are your thoughts? (describe in detail)
- How do you feel when you come across these images?
- How does your own behavior regarding your BI change due to Insta?
- Have you experienced physical sensations due to viewing or posting a photo on Instagram? If so, please describe them. (for example, anxiousness, change in breathing or heart rate, or maybe you feel heavy or down, maybe arousal…?) – Examples given if the need is felt.

Topic 5: Attention - Difference on and off Instagram
- Is there a difference in how you see your BI on and off Insta? If so, in what way?
- Is there a difference between how you show (present) your own body on Insta and how you see it yourself?

Topic 6: Prevention/Management - Regulation (Policy)
- I've been seeing body positive narratives and "reality" posts- what do you think about the role of these posts on Insta?
- What strategies do you find yourself using to protect your body image?
- In your view, what can be done to prevent body images of young people from being affected by Instagram Use?

Instagram has some policies… like (advertisement, no nudity, no hate speech, give credit to the content owner)- specify for body related. (Are you familiar with policies when people's bodies are shown?)

What are your thoughts on how Insta's own policy making could promote a healthy body image?

Topic 7: Summarize
- So, we've come to conclude this session and I would like to mention that these are the points that we've have discussed… (maybe written on a board- key words, section wise)
- Would you like to discuss or mention any other point from your experience related to the topic…?
- We're going to be holding this discussion in the coming weeks with other groups, so if you have any feedback for us, it will be appreciated!

- How did you find this discussion, what did you like about this, what did you not like, what could be improved?

After summarizing the focus group discussions, each participant received a 15-euro GoGift voucher as a token of gratitude for their time. Field notes, detailing participants' responses, and seating plans were taken during the focus groups to aid in transcribing the data. No repeated or duplicated focus group sessions were conducted. The

researchers did not organize multiple sessions with the same group of participants to explore the discussed topic of how young people view the role of Instagram on body image further. This decision was due to the limited time frame of one semester to complete the study.

Data analysis

Focus groups were audio- and video-recorded. The first step of the analysis involved manually transcribing the audio tapes, with the aid of the video tapes for additional observations of the participants' mimics, gestures, overall behavior, and attitude. Study facilitators Katherine Marsh, Stacy Mauer and Sara Valerio de Sousa transcribed equal sections of the male focus group recordings, while Chiara Guinet-Venzi and Lis Krack divided and transcribed the recordings of the female focus group.

In a second step, each transcript was thematically analyzed, utilizing Braun and Clarke's six step framework (2006). Initially, the transcripts were thoroughly read by each researcher to establish familiarity with the data. Subsequently, each facilitator independently identified initial codes within the data sections they had transcribed using an inductive approach. The generated initial codes, along with the transcripts were documented in Microsoft Word. During a subsequent collective review involving all researchers, comparisons of codes based on similarities were undertaken and analytical memos were added. Codes were then grouped based on similarities, and if needed were additionally refined through shortening, editing, or merging. As an example, the code "perception of Instagram imposing body image and sexualized content onto users" and the code "content is made for profit (advertisements), presented with it regardless of want" were found to be similar and thus grouped. The collated codes were then sorted into subthemes. Codes that were deemed irrelevant in relation to the research question, such as the code "engagement with memes", were discarded in this process. The subthemes were then combined, and main overarching themes were formed. Following this process, five latent themes for the male group and four for the female group, encompassing several subthemes were devised. Previously established and newly arisen psychological concepts, such as peer influence, self-objectification, cognitive dissonance, selective self-presentation, gender stereotypes or media literacy, were added to each subtheme within a theme. This was done to relate the subthemes to relevant psychological theories.

To assess the reliability of the raters in measuring the same codes and thus ensuring the trustworthiness of this research, a third step involved calculating inter-rater reliability using Fleiss Kappa, given the presence of more than two raters for each transcript.
Since in the case of Fleiss Kappa the variable to be measured is categorical, the codes were compiled into a Microsoft Excel table, and each rater indicated their agreement with codes using the number 1.00 and disagreement with .00. The data was then transferred and statistically calculated using the SPSS software. The Fleiss Kappa showed a high agreement among the raters for both the male and the female group, with $\kappa_m = .872$, $p < .001$, 95% CI [0.765, 0.979] and $\kappa_f = .794$, $p < .001$, 95% CI [0.687, 0.900].

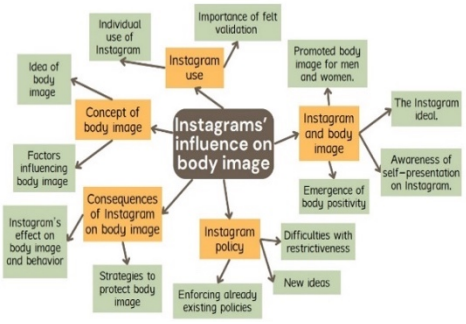
All researchers worked reflexively throughout the process of data collection and analysis, ensuring that their personal preconceptions and usage of Instagram did not affect the study results or introduce bias.

Results

Disclaimer: To ensure the confidentiality and anonymity of our participants, the names used in this report are entirely fictitious. Participants were given the option to choose a name of their preference to preserve anonymity during the discussion.

Male group

To the best of our knowledge, there is a shortage of in-depth research investigating the impact of Instagram on men's body image, which is worth highlighting. Most of the knowledge that is currently available focuses on how it affects women's body image. This gap in research motivated our study to explore the specific experiences and perceptions of men in relation to body image influenced by Instagram.



1. Concept of Body Image

1.1. Idea of body image

There has always been a concept of the "ideal body", but over time, society's definition of the ideal has evolved. Nowadays, most people's conception of an ideal physique tends to be more athletic in shape.

Tono: "For me personally when I heard the word body image, it's how beautiful, how fat, how thin that person is, and yes."

Joseph: "Well, I'd say there always was kind of a body image (…) only the standard shifted."

1.2. Factors Influencing Body Image

The study identified public opinion and influencers as significant factors influencing body image, with participants emphasizing the impact of influencers on personal body perception.

Jos: "Influencers can affect a great part of your personal body perception."

2. Instagram Use

2.1. Individual Use of Instagram and its Features

Participants revealed diverse ways they engage with Instagram, emphasizing the platform's role in documenting personally significant moments and reflecting on past experiences through posts. They highlighted the use of Instagram tools, such as filters, music, and captions. Moreover, the majority expressed using Instagram for social interaction, entertainment, and staying connected with personal interests, public figures, and influencers.

Jos: "Often even without noticing it, I just sort of end up doomscrolling memes for memes. Um other than that, I have some interesting persons from public life, influencers I know."

2.2. Importance of Felt Validation through Instagram

The significance of receiving recognition and attention on Instagram emerged as a prominent subtheme. Participants expressed that positive feedback on their posts contributes to feelings of happiness, with some actively strategizing the timing of their posts to maximize attention.

Jos: "I do feel happy, yes, when I get comments or reactions from my friends when I post on Instagram (…)"

Joseph: "(…) most of us do it for the sole purpose of getting reactions or getting commented by some people (…)."

3. Instagram and Body Image

3.1. Promoted Body Image for Men and Women

Men acknowledged that idealized body standards on Instagram affect individuals of both genders, but participants pointed out distinctions in their reactions. Men in the group mentioned observing a higher occurrence of inappropriate comments on women's posts, indicating societal pressures related to appearance.

Jos: "(…) always body related, influencers, lightly dressed people (…)."

Jos: "(…) those bodies which are very good looking, fitness influencers or gym dudes or girls who always train and showing off their progress, or very oversexualized content (…)."

3.2. The Instagram Ideal

Participants noted the sexualization of the female body on Instagram, highlighting the prevalence of perfect but unrealistic bodies and self-objectification.

Tono: "Well-shaped bodies and it's like everybody wants to take that kind of shape, but nobody has it in the real world."

3.3. Awareness of Self-Presentation on Instagram

Participants acknowledged the importance of humorous posts and consideration of others' perspectives in their Instagram content. Additionally, they emphasized the value of privacy and the importance of respecting everyone's right to privacy by getting consent before sharing pictures with other people.

Tono: "For me, the perspective of the people that are following me."

Joseph: "Of course, I want to look presentable; I don't want to look like garbage."

3.4. Emergence of Body Positivity

While participants acknowledged body positivity as a recent trend and expressed its importance in challenging typical ideals, they also voiced concerns about the misrepresentation of body positivity by some influencers. Specifically, they pointed out that certain influencers are promoting obesity under the guise of body positivity.

Jos: "Compared to the masses of oversexualized content and very slim waists, it's nothing still but that can, of course, change."

Jos: "Most of the American influencers who are patronizing talking very positively about, how do you say it, obesity, for example, and glorifying that as body positivity."

4. Consequences of Instagram on Body Image

4.1. Instagram's Effect on Body image and Behavior

Positive effects on motivation for fitness and health were noted by participants, indicating that Instagram content can inspire positive behavior and reduce insecurities.

Tom: "It also gives a positive impact for me because when I see something or some content that recommends me to go for exercise for ten minutes today."

4.2. Strategies to Protect Body Image

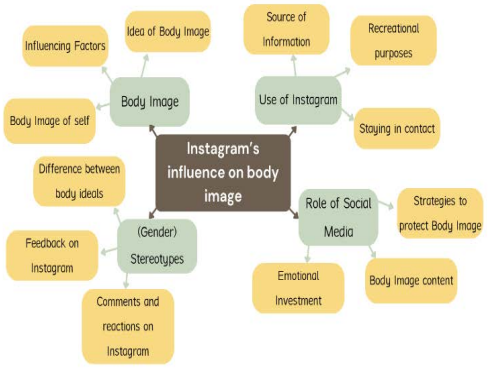
Participants suggested various strategies to protect their body image, including following body-positive content, avoiding social comparison, and learning to cope with hateful comments.

Jos: “Yeah, that's important, and sometimes when I’m scrolling my feed and I see too much of these over-the-top influencers, I just tend to close the phone and put it in my pocket.”

5. Guidelines

Participants highlighted the need for stricter guidelines on Instagram, including giving credit to content creators, reporting negative comments, and using AI to detect policy violations. However, they acknowledged the challenges of implementing and enforcing these guidelines effectively.

Female group



1. Use of Instagram

1.1. source of information

Instagram is used by our participants as an easy and quick access to a broad range of information available online. It updates them on what happens all around the world at any given time.

1.2. recreational purposes

Instagram allows individuals to express themselves freely and view the content they like to see. This includes looking at content that is related to hobbies and individual lifestyle. Some of our participants also mentioned that they use the app to pursue content creation in the form of photography and share their hobbies with others.

Vavi: “I have 200 something (posts) but it´s because I like photography, so I take pictures, but I also have pictures of me doing crazy stuff (...).”

1.3. staying in contact

Many of our participants reported that they primarily use Instagram to stay in contact with their friends and reconnect with people they’ve met. Some also explained that they have a public account and a private account for sharing more in-depth experiences of their lives with close friends.

Sophia: “I also have to say I have a second Instagram account, like my private account and there I post like everything, I post when I cry, I post when I´m happy, I post when something funny happens in my life, I post when something cringe happens in my life, I, but it´s only my close friends.”

2. Body Image

2.1. Idea of Body Image

Body awareness and body shaming begin in childhood and early adolescence, giving a certain image of our own body and of others.

2.2. Influencing factors

There are several influencing factors that make one aware of one's own body. Some may be related to the standards of beauty and gender stereotypes one is exposed to. Others may come from our immediate environment, such as family, friends, and peers. In addition, the me-

dia around us and the culture we have grown up in can also have an impact. Furthermore, in the clothing industry, the sizes on offer often make us feel uncomfortable in our own bodies. Finally, the importance of public opinion also plays a role in influencing factors that some may face.

2.3. Body Image of self

When one's own body image is explored, the focus lies on how someone looks, and when posing in pictures, the importance lies in making sure the person looks as flattering as possible. This behavior can sometimes trigger mental illnesses such as eating disorders. This phenomenon has already been explored in which it was said that a predictor of body image and eating disorders in adolescent girls came from social comparison and critical body image processing (Botta, 2003: 389). In addition, the way one sees themselves can change as one gets older, making one less confident if the image that is presented of one self online does not match how someone may look. This can lead to cognitive dissonance because there is no control over how people might perceive the one sharing content online.

3. Role of social media

3.1. Body Image content

The promotion of body ideals, body trends and diet culture can cause eating disorders when people are exposed to negative and/or edited content.

3.2. Emotional Investment

When being exposed to body related content, either negative or positive feelings can arise from it. Existing surveys found out that there is a higher body dissatisfaction when using apps with visual content, as for example Instagram (Grogan, 2021). When seeing negative body content feelings of distress can be dominant. On the other hand, positive feelings are felt when seeing body positive content on Instagram. Nevertheless, it is noticeable users may feel disappointment when influencers that are often seen as role models, capitalize and lie about their body.

3.3. Strategies to protect body image

Sometimes it can be a challenge to protect people's body image, but there are some things that can help. Being aware of comparisons may make it easier to avoid them or to be more aware of them, which may lead to a less damaging outcome. Also, limiting the filtering of content viewed and promoting body positive and inclusive content may be helpful. In addition, media literacy and regulation could be further adapted to ensure more protective strategies.

4. (Gender) stereotypes

4.1. Difference between body-ideals

Stereotypes, especially gender stereotypes were a big part of our discussion and every participant had experiences to share ranging from insecurities on gender stigmas and gender stereotypes, double standards on how they should look and act but also the difference the perceived in feedback compared to the one, men get. It was mentioned that especially for the whole fitness and gym scene, women are shamed for looking to muscular when men are praised for it. Women should look thin and athletic but also not too much as it can look too "manly". The talked how there was more shaming towards women depending on body size and lifestyle choices. It was also discussed that the bodies of women are much more sexualized compared to the bodies of men while women seem to look beyond appearance.

Sophia: "(...) I also feel like the body of a woman is much more sexualized than

the body of a man (...), me personally I don't sexualize the male body (...)."

4.2. Feedback on Instagram

There was a difference noted on the feedback received by men and women where men did not seem to care as much about their own but also other's bodies, particularly men's bodies. Men also seem to be more supportive towards other men while the comments under a woman's post tend to shame the woman for normal bodily features such as body hair or a different size of body.

The female body is also being used to capitalize and women face the problem of self-objectification where many influencers profit off their bodies as a main selling point of their page. It has also been noted that participants feel like nudity is viewed differently based on gender.

Georgia: "(...) sort of like gym influencer on this and his comments are all men being like oh sick bro, that's so cool bro, GG, good job(...)."

Jane: "(...) I've really seen is when girls have like body hair and stuff, they're, all the comments are just men getting mad at what nothing."

4.3. Comments and reactions on Instagram

To go into more detail on how comments come across differently based on gender, participants claimed, that comments from other women seemed much more genuine than the comments received by men. Those on the other hand were seen as unwanted and men apparently made use of specific features can mean different things including to sexually harass other's by sending them the fire emoji for example. In short, the difference in attention based on gender was noted.

Sophia: "(...) I love it more when girls are saying like "you look gorgeous" than when boys are saying like "you look gorgeous". (...) I feel like, they are not you know as genuine as girls."

Lena: "You know, like you can get the emojis, like the predefined emojis and eh there was this thing between boys (...), to approach a girl, they would just send the fire emoji on like a picture in a story (...)."

Discussion

In our study, we aimed to analyze how young people see the role that Instagram has on body image. For this, we conducted two focus group discussions that were fruitful and resulted in a variety of different answers. The discussions covered key terms like the concept of body image and how it is influenced, gender stereotypes, self-representations including the issue of self-objectification for women, the body positive movement but also eating disorders. We discovered that it is a tripartite model of influencing factors for the body image: family, friends and media. They all play a key role in the creation and moreover the development of one's body image.

For the female group, they started off by talking about how they use Instagram and what for. Then it was quickly established that for every participant, Instagram has played a key part in the development of their body image. This can be led back to Festinger's theory of social comparison that explains that individuals tend to compare themselves to people they see on Instagram in this case (Festinger, 1954). It was mentioned that it also may have impacted the forming of eating disorders during younger years highlighting the severity of the impact the platform can have. This was particular for the female groups because the

male group didn't mention eating disorders. Additionally, women talked about the topics of gender stereotypes in combination of self-objectification and the double standards persisting on the app. The objectification theory explains this in further detail, but in short, especially women are exposed to a constant representation of the westernized norms with the emphasis on physical appearance (Fredrickson and Roberts, 1997). This constant comparison of your own body to unrealistic beauty standards and the emphasis by society on physical appearance can lead to a cognitive dissonance in individuals. Cognitive dissonance describes the fact that two things are not consistent with each other. This is a state of dissonance, it is not comfortable for the individual and they try to reduce it. This can be achieved by changing one of the two opinions. If we take the example of your body image and beauty ideals, one might try to fit these set standards. If this is not attained, dissonance is created (Festinger, 1957).

The male group had very similar topics. They mentioned the importance of felt validation, strategies to protect your own body image and policies Instagram could introduce the negative effect it has on people's view of themselves. It was discussed how these policies bring a lot of restrictions and difficulties but also about media literacy and the importance of education in this context.

Many theories back up our findings like the ones that were mentioned earlier in the discussion. Festinger's theory of social comparison explains why people are so prone to be influenced by what they see on social media. It demonstrates how individuals perceive others, especially if they notice similarities (Festinger, 1954). The same goes for the objectification theory that provides an insight on why especially women tend to distance themselves from their body due to the constant exposure to westernized norms and the heavy focus on physical appearance. This also gives us an explanation why some influencers make money off their body and build their page all around it. Bandura explained in his theory of social cognitive mass communication why influencers are even able to do so. People tend to model their behavior and their values of what they see, in this case the content posted on social media (Bandura, 2001).

In recent years however, a certain type of countermovement has taken over social media. More and more influencers talk about the importance of being real on the app and show their bodies in a normal light. Additionally, a lot more bodies are being represented to normalize real bodies and not give so much importance to these unrealistic beauty standards.

Limitations and further research

Limitations include that the interrogators were all female. Due to that, men seemed to regularly change opinions and factors like social desirability played a role. They were also most likely influenced by societal norms about how men must behave and think in certain situations. The female team could also pose an issue in the analysis part of the study. As another result, women seem to get exposed to different reactions and different experiences in everyday life than men. This "priming effect" could have influenced the interrogators' reactions and way of thinking. It was probably also a lot easier for the participants to confide in interrogators of the same gender, expecting them to have had similar experiences based on gender.

Future studies can analyze beyond our findings and include aspects like culture differences and other social media applications. In some cultures, African or Asian especially, commenting on other people's bodies is much more common than in the western cultures and this

could also impact how people react to the exposure to other people's bodies but also the feedback received on their own body on social media. Maybe they are more sensible to it because of their prerequisites or maybe they are more used to it, and it doesn't affect them as much anymore. As for social media platforms, one could expand one's studies by including multiple others like Youtube or Tiktok, including also auditorial information or simply just focus on another form of social media completely. Future studies could also further analyze the background of the clear difference in the perception and experience of men and women on social media. We have already discussed the effect of societal norms on this but there could also be other important factors, maybe even biological ones? Is it possible that certain biomarkers give way for certain reactions and could count as a predisposition for the effects that Instagram has on the body image? Discussing the gender difference in general could be interesting for future studies. Why do our participants claim that men are quicker to comment on posts and criticize others? Was it maybe just our group of participants that noticed such effects?

References

Baker, N., Ferszt, G. G., & Breines, J. G. (2019). A Qualitative Study Exploring Female College Students' Instagram Use and Body Image. Cyberpsychology, Behavior, and Social Networking, 22(4), 277–282. https://doi.org/10.1089/cyber.2018.0420

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology,* *3*(2),77-101. https://doi.org/10.1191/1478088706qp063oa

Enke, N., & Borchers, N. S. (2021). Social Media Influencers in Strategic Communication: A Conceptual Framework for Strategic Social Media Influencer communication. In Routledge eBooks (pp. 7–23). https://doi.org/10.4324/9781003181286-2 :

Fardouly, J., & Vartanian, L. R. (2016). Social Media and Body Image Concerns: Current Research and Future Directions. *Current Opinion in Psychology, 9,* 1-5. https://doi.org/10.1016/j.copsyc.2015.09.005

Festinger, L. (1957). A theory of cognitive dissonance. In Stanford University Press eBooks. https://doi.org/10.1515/9781503620766

Grippo, K. P., & Hill, M. S. (2008). Self-objectification, habitual body monitoring, and body dissatisfaction in older European American women: Exploring age and feminism as moderators. *Body Image*, *5*(2), 173–182. https://doi.org/10.1016/j.bodyim.2007.11.003

Grogan, S. (2021). Body image: Understanding body dissatisfaction in men, women and children. Routledge. https://doi.org/10.4324/9781003100041

Hennink, M. M. (2014). Introducing Focus Group Discussions. In *Understanding Focus Group Discussions, Understanding Statistics*. New York: Oxford Academic. https://doi.org/10.1093/acprof:osobl/9780199856169.003.0001

Jonas, K., Stroebe, W., & Hewstone, M. (2007). Sozialpsychologie. In Springer-Lehrbuch. https://doi.org/10.1007/978-3-540-71633-4

Krueger, R. A. (1998). *Developing Questions for Focus Groups*. SAGE Publications, Inc. Focus Group Kit. https://doi.org/10.4135/9781483328126

Keery, H., Van Den Berg, P., & Thompson, J. K. (2004). An evaluation of the Tripartite Influence Model of body dissatisfaction and eating disturbance

with adolescent girls. *Body Image*, *1*(3), 237–251. https://doi.org/10.1016/j.bodyim.2004.03.001

Rusticus, S. A. (2014). Body image. In *Springer eBooks* (pp. 420–423). https://doi.org/10.1007/978-94-007-0753-5_224

Sharifi, S. M., Omidi, A., & Marzban, B. (2016). The Impact of Instagram Use on Body Image Concerns among Iranian University Female Students: A Phenomenological Approach. *International Journal of Academic Research in Psychology, 3* (1), 26-36. http://dx.doi.org/10.6007/IJARP/v3-i1/2280

Encouraging young children's peer collaboration for math problem-solving

Drockenmuller Morgane, Hellenbrand Samantha, Meier Melissa, Schmit Ben, Signore Laura

Tutor: (PhD candidate) Dipl.-Psych. Hoyeon Lee

This research focuses on early childhood numeracy and social collaboration between young children. The study wants to examine the possible role of peer collaboration on mathematical problem-solving in a preschool setting. Precisely, the main goal of this study is to explore the impact of social cooperation and the effects of group. Children (N=100) from seven preschools in Luxembourg, aged between 5 and 6 years, were tested. The research was divided into 2 parts, both of which were performed one after the other. The first part was an individual session with measures of visual perception and mathematical skills. The second part included mathematical problem-solving tasks, where the children worked individually or in a paired setting from 2 children. This is examined by a mirror task to assess the visuospatial skills and a cookie task for counting and size comparison. The analysis was conducted with parametric and non-parametric tests to explore the impact of collaboration on children's problem-solving skills. In addition, correlation analysis was made to study further relationships between the different kinds of tasks. The results indicate significant differences between the tasks done individually versus in collaboration. The children, who worked in pairs, scored significantly higher. Furthermore, positive correlations were observed between different variables, such as visual perception and size comparison tasks. The results suggest that peer collaboration positively influences children's mathematical problem-solving skills. However, the potential impact of item difficulty on the discriminative power of the task is acknowledged in the paper. It should also be noted that the study does not record the specific interactions of the children. Future research could similarly delve deeper into the more precise interactions between children

Keywords: Early childhood numeracy, peer collaboration, preschool children, mathematical problem-solving

Introduction

Background and context

Early numeracy, also called early number literacy, is a fascinating discipline that requires a focus on basic mathematical skills such as understanding numbers, (verbal) counting, solving (e.g. simple problems or basic operations), measuring (e.g. example height, length, capacity) or estimating (e.g. quantity estimation, measurement estimation, comparative estimation) that help preschool children develop a fundamental understanding of numbers, quantities, and relationships especially.(Reid, 2016)

Children need the ability to solve problems and to manage different everyday tasks like recognising patterns, finding solutions, using data or financial management (French, 2013). The relevance of this area of research must therefore be emphasized.

In the realm of innate number sense, studies such as those conducted by Izard et al. (focusing on non-symbolic comparison) in 2009, Starr et al. (emphasis on predicting power) in 2013, and Wynn (simple arithmetic operations) in 1992 have indicated that early numeracy surpasses counting. It includes both number sense and symbolic knowledge, involving the acquisition of information (with early interventions in environments like the home and kindergarten playing a vital role). These foundations are closely linked to children's later

academic success, as Duncan et al. (2007) found in their comprehensive study.

A notable finding was also the malleability of early childhood numeracy, as highlighted in studies by Cornu et al. (2017). It suggests that kids' math skills in early childhood are not set in stone but can be influenced and improved. It highlights that the environment they are in, along with educational efforts, contribute to shaping their math abilities. In simpler terms, the study implies that with the right support, like good educational programs and a nurturing environment, young children's math skills can see improvement and development.

Furthermore, research, like the one conducted by Mix & Cheng (2012), underscores the significance of visual-spatial skills and their connection to children's mathematical proficiency. In fact, visual-spatial skills are crucial for understanding and manipulating visual information in space. The relationship between visual-spatial skills and math is multifaceted and involves several aspects such as spatial reasoning, problem solving, geometry and visualization, number sensed, memory, graphics, etc. These findings emphasize that nurturing these skills is vital for establishing a solid foundation in mathematics.

Early Numeracy and cognitive development

The primary abilities of pre-schoolers' mathematical development, early numeracy, consist of several important components that are the subject of our investigation. According to Wynn (1992) and Starr et al. (2013), fundamental components such as counting, quantity discrimination, and basic mathematical operations form the basis for mathematical comprehension (Wynn, 1992; Starr et al., 2013). Wynn's (1992) study highlights that 9-month-old infants surprisingly can discern the physical laws of numbers (Wynn, 1992).

One of the earliest mathematical abilities, as observed in Wynn's study (1992), is the ability to discriminate between quantities that we identified as number sense. This skill serves as the foundation for comprehending numbers and quantities which develop during kindergarten.

As highlighted by Mix & Cheng (2012), a crucial aspect in investigating early childhood numeracy is the emphasis on the importance of visual and spatial abilities. Early on in development, visual abilities—which include things like pattern identification and visual-spatial understanding—are crucial for promoting mathematics knowledge. Simultaneously, spatial abilities, including the capacity for spatial visualization and orientation, provide a valuable foundation for the development of numerical skills in childhood (Mix & Cheng, 2012).

These abilities underline the value of developing knowledge of mathematics in early childhood education since they are closely linked to it.

According to the study by Mix & Cheng (2012), kids who have higher visuospatial abilities typically do better in math (Mix & Cheng, 2012). Given the connection between proficiency in mathematics and visuospatial abilities, it is important that these skills continue to grow in early childhood education.

In the present research, we not only explore children's numerical abilities, but also investigate how these abilities are related to their visuospatial abilities. This aspect of our study will provide us with important insights into children's solving skills and the development of mathematical understanding.

Social Collaboration and cognitive development

A significant dimension in early childhood development is social interaction. This influence has been extensively studied and explained by theorists such as Lev Vygotsky (1978) and Albert Bandura (1977).
Vygotsky (1978) introduced the concept of the zone of proximal development, which shows

how children can expand their cognitive abilities through social interactions. This concept refers to the area between a child's current level of development and their potential level of development that they can achieve through support. Social interactions with more competent individuals enable children to expand their cognitive abilities in this zone.
Social learning theory, particularly according to Albert Bandura (1977), emphasizes the role of observation, imitation, and modelling in cognitive development. Children learn by observing others and imitating behaviours, which can lead to enhanced cognitive development.
Our research will build on these findings to investigate how peer collaboration influences children's mathematical abilities. Specifically, we will examine factors that might influence the effectiveness of social interaction, such as children's social skills, cognitive ability, and specific task type.
Investigating these influencing factors will help us understand how peer collaboration affects the children's problem-solving ability.

These findings build the foundation of our research, which seeks to investigate the influence of peer collaboration on children's problem solving.

Aim and research question

Our study aim is to gain a deeper and better understanding of the development of early childhood (5–6-year-old children numeracy and the impact of social collaboration on mathematical skills. We plan to conduct a mathematical test in a play school with a laboratory setting that provides a place where we can control the conditions and where we have an accurate measurement for studying children's mathematical skills in collaboration or alone.

Thus, our hypothesis is the following:

Hypothesis 1: When children collaborate during math problem solving tasks, they have higher results than children who do the tasks alone.

Hypothesis 2: If Children obtain high scores in tasks related to visual perceptual skills, they will also obtain high scores in tasks related to symmetry.

The multi-faceted aspects of early childhood numeracy and social collaboration could be explained by the help of these research questions and to gain new insights into how children's mathematical skills can be developed.

Methods

Participants

The study involved a sample of children (*N*=100) recruited from seven kindergartens in five regions of Luxembourg. More precisely, the sample comprised more male (*N*=54) than female (*N*=46) participants. The age range was between 5 to 6 years (*M*=68,19 months; *SD*=4,563 months). Only the age was relevant, there was no consideration of the ethnicity of the children if they were in a kindergarten in Luxembourg and understood the basics of the Luxembourgish language. The decision to limit the sample to this specific age group assumed that children of this age are capable of effective communication and comprehension of the task demands. In addition, the main reason was also the language. The focus was on the fact that early childhood development plays an important role in the rest of development. Identifying the effects of cooperation on problem solving tasks at such an early age can be helpful for further research into child development. The data were collected between November and December 2023.

Procedures

The criteria for the selection were limited to children between 5 and 6 years and to preschool.
In the first step, we reached out to preschools situated in Luxembourg and established contact with the kindergarten teacher. Further-

more, the study required the consent from the mayor of the municipality and from the school director.
The teachers were then briefed about the objectives of our research. Additionally, consent forms were disseminated amongst the pupils. The child's parents were given the chance to permit or refuse their child's involvement in our experiment. Consequently, the number of children taking part in our study was at times quite restricted in some classes. It was important due to ethical reasons that both the teachers and the parents of the children were thoroughly informed about our approach and the purpose of our study.

The study took place in a typical laboratory setting. Even though it was conducted in schools, in an area separate from the classroom, we did not directly work in the classroom with the cooperation of the teacher as in a normal school day. The provision of a quiet environment enabled us to exclude as many confounding variables as possible. To provide the same condition for all the subjects, the assessment took place in the morning for all the participants. This enabled us to exclude possible fatigue effects among the peers.

The test is always administered by two test supervisors: One experimenter evaluates the children's performance through the completion of task-specific sheets while the other one focuses on the interaction with the child and gives the needed instructions and explanations.

Throughout the various tasks, neither of the two experimenters took an active role, whether the children worked individually or in pairs. Both the mathematical and problem-solving tasks took approximately 20 minutes to assess, and the order in which the testing sections were administered was consistent across all participants. Clear instructions were given to all testers in advance to standardize the interaction and the test procedure. The interaction guidelines were detailed in written format and familiar to all test administrators to avoid discrepancies between testers and ensure optimal inter-rater validity. Furthermore, test scoring was standardized to ensure a high level of objectivity in our research.

The exams were conducted in either Luxembourgish or French, to ensure that the language barrier did not affect the children's test scores.

Measures

The investigation employed two distinct tasks: a fundamental math aptitude task and a problem-solving task. Both tasks lasted together around 40-60 minutes. The math aptitude task section comprises activities that test abilities in verbal counting, non-symbolic comparison, visual perception (FEW-II figure ground) and mental transformation (CMTT).

Fundamental Math Aptitude Task

Specifically, the FEW-II activity requires children to recognise shapes in a picture. There were 18 items in total with an increasing difficulty. For the next task the child was presented with two parts of a shape and had to determine the full shape resulting from their combination. They then select the correct answer from a set of four options (CMTT). The CMTT had 16 items. After that, they were asked to count aloud as far as they could. In the first trial they did it independently and afterwards got help with one of 3 different possibilities: A) Can you start counting? (if the child has difficulties with starting to count), B) Wait, what comes after number X or C) Wait, what comes after number X + 1?. The trial stopped when A) the children did 3 mistakes directly afterwards, B) did 4 mistakes on one single decimal number or C) did not count any further even with one of the assistants. After the second trial, the last correctly stated number was recorded.
Subsequently, the child was instructed to begin counting from a specific number and only awarded a point when all the 4 following numbers were accurately stated. There were 5 items and stopped when there were 3 non-solved items in a row. This was followed by a task where the child was required to count

backwards from a specific number. Here we have the same requirements as for the forwards counting.
The subsequent non-symbolic task, a motor speed test, entailed identifying the box (out of two) which included the greatest number of points by marking it off. They were allotted two minutes to mark off as many as possible and the score was calculated based on the ones that were correctly marked off. For the concluding task of the mathematical segment, they had to mark off as many black shapes as possible within 20 seconds.
The mathematical component enables the assessment of each child's math ability, and all participants worked individually for this section.

Problem-solving Task

The second component, the problem-solving task, assesses number skills through a mirror and a cookie task. The children got introduced with "Petzi", the teddy bear, who will help them through the tasks. Petzi should support and motivate the children during the process. For this part, the experimental group was paired with a partner of the same gender from their own class. This grouping strategy ensured that the participants were comfortable working together, which is based on the social learning theories, which supports the ideas that the presence of others promotes acquisition of knowledge, improves the attention, and enhances the thinking. For the mirror task, children must replicate the construction in a symmetric manner, imagining the blocks in the mirror. The task score was calculated based on the placement, perception, symmetry, overall shape, and finishing of the blocks. First, the children got training to get familiar with the concept of the mirror and the task. There were 5 items in total. For the cookie task, children were tasked with identifying the largest cookie from a selection of cookies made up of identical small blocks arranged in various shapes and sizes. The question for each item was, which cookie is the biggest and they should feed it to Petzi. Here again, they got training to get to know the small blocks and to discuss how you can approach the task. Various strategies are available to solve this task, such as breaking the objects into smaller pieces, comparing their weight, size or overlaying them. The different approaches were scored with different points.

The points were distributed like the following:

- Picked the right figure (+1 or no point)
- Disassemble (+0.5): count the blocks (+1), compare the amount (+0.5), make a line (+0.5), play (0)
- Assembled (0): Overlay (+0.5), compare side by side (+0.5), compare the weights (+0.5), play (0)
- Explored all the given blocks? (+0.25 or no point)

The number and size of the cookies increased over the items. Here again we had 5 items in total. In evaluating, we do not only consider if the child/children pick*s the right cookie, but we also consider the specific method employed by the children.

The problem-solving task was carried out with two variables: one group working in pairs and the other completing the task individually.

At the end of the testing session, the children got the opportunity to evaluate the overall situation and their current emotion after the session. They also got a certificate with a sticker.

Data

To verify our first hypothesis, we compared problem-solving scores of the experimental group (children who worked in pairs) with those of the control group (children who worked alone). Consequently, we conducted a t-test to compare the means of both groups. This approach allows us to answer our hypothesis if children, who worked in pairs, will present higher scores than those who worked individually. Furthermore, the mathematical assessment is used to evaluate the mathematical ability of each child, thus enabling us to calculate the correlation between mathematical ability and problem-solving scores.

Results

As the aim of our current study was to identify whether collaboration between peers leads to better results in mathematical-problem-solving tasks, we conducted an independent sample t-test for the cookie task (table 2). This analysis enables us to verify if there is a significant difference between the mean of the two group conditions. In fact, for this task we observed that the mean of the group condition (*M*=8.42 *SD*=2.95) is significantly higher than the mean of the individual condition (*M*=5.66 *SD*=3.30) *t*(98)=4.25, *p*<.001.

Regarding the cookie task, we can observe a mean of *M*=8.42 and a standard deviation of *SD*=2.95 for the group condition, whereas the control group disposes of a *M*= 5.66 and a standard deviation of *SD*=3.30.
Furthermore, for this task the results of the participants follow a normal distribution, which indicates that the task was neither too easy nor too difficult and was perfectly adapted for the children's age.
Considering the mirror task, we conducted a normality test, which was significant. Meaning that the data doesn't follow a normal distribution. Therefore, due to violation of normality, we conducted a nonparametric t-test (Mann-Whitney U) and obtained the following results: *W*=.824, *p*<. 001.Thus we can conclude that collaboration lead to significantly higher scores when performing the mirror task than when the children worked individually. (Table 3)In the following table (Table 1) the descriptive statistics of our results are being displayed. We observe a median of *M*=36.7 and a standard deviation of *SD*=3.11 for the mirror task in the group condition, whereas the children who worked individually obtain a median of *M*=33.7 and a standard deviation of *SD*=4.92 for this task. The highest possible score for the mirror task is equal to 38.5.

Table 1

Descriptives of Individuals and Groups

	Condition	Sex	M Total	C Total	F+C	Counting with IHC	Counting with FHC
N	G	66	66	66	66	66	66
	I	34	34	34	34	34	34
Missing	G	0	0	0	0	0	0
	I	0	0	0	0	0	0
Mean	G		36.7	8.42	18,6	10,5	12,0
	I		33.7	5.66	17,5	8,94	10,1
Median	G		38.3	9.00	18,0	10,0	11,5
	I		34.6	4.50	17,5	8,0	9,00
Standard deviation	G		3.11	2.95	4,68	5,91	6,41
	I		4.92	3.30	4,39	6,42	6,84
Minimum	G		24.0	3.00	10	1	2
	I		20.0	1.00	11	1	1
Maximum	G		38.5	13.8	28	30	30
	I		38.5	13.8	29	29	29

Table 2

Indepedent Sample T-Test (cookie task)

		Statistic	df	p
C total	Student's t	4,25	98,0	<.001

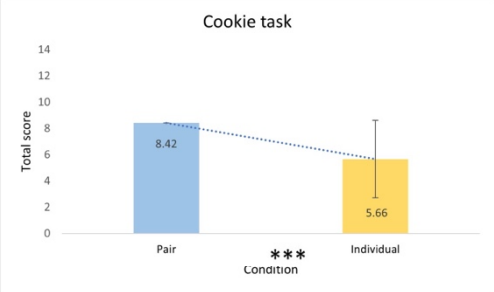
Table 3

Independent Sample T-Test (mirror task)

		Statistic	p
M total	Mann-Whitney U	661	<.001

Having a closer look at the data, we also observe numerous meaningful correlations among the different variables. We observe a positive correlation of r=.43, p=.006 between visual perception skills (figure ground) and the mirror task (symmetry). For this correlation, we analysed all. individual performances for the children (N=100) of the visual perceptual skills and the results of the two experimental groups Individual and Group for the mirror task.

Figure 1



Pair/individual cookie task

Figure 2

Pair/individual mirror task

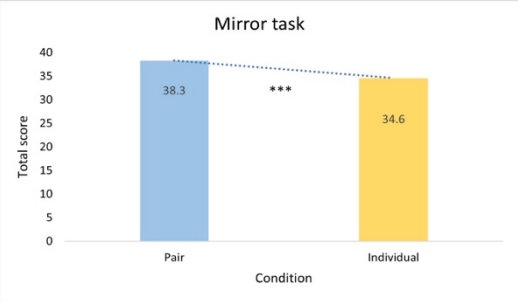
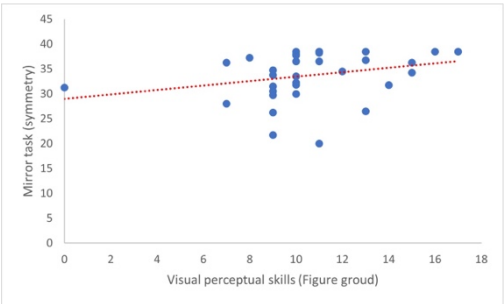


Figure 3

Distribution of the Mirror Task and Figure Ground



Discussion

Using various statistical tests and descriptive statistics, the study examined the relation between children's problem-solving skills and collaboration. This study provides evidence of a positive peer collaboration effect among young children on their performance in mathematical problem-solving tasks. Children performed better in pairs in aspects of the mirror task, which implies the visuospatial skills, and the cookie task for counting and size comparison. However, there are still enough further research possibilities to deepen the previous results.

The mirror task showed a right-skewed distribution. Despite the violation of normality, we proceeded with the analysis because it allowed us to compare results between working in pairs versus individually and observed significant effects. It would be beneficial if future studies used more difficult items for the mirror task or used a different alternative to measure visuospatial skills to avoid the ceiling effect.

This limitation restricts the comprehensiveness of the data set, especially in areas where visual insights could give useful background information. However, an easy resolution of this issue would be finding alternative visualization techniques. These changes would improve the clarity of the study.

A further research idea would be to analyse deeper the different strategies of the children and gain valuable insights into the reason for these varying results. For instance, exploring the nature and quality of interactions in groups or looking at individual characteristics that may contribute to better performance. Further research can also focus on the link between the motor speed tests and its impact on the results. Since the children do paper and pencil comparison tasks which can have an impact on the motricity development and the results.

In conclusion, although this study contributes evidence on the positive impacts of collaboration on children's problem-solving skills, future studies could complement these findings. Addressing the identified limitations would contribute to a more perceptive interpretation of the results.

Future research efforts should carefully consider these challenges to improve our understanding of the dynamics between collaboration and the early childhood numeracy.

Conclusion

The research investigated further on the early childhood numeracy and peer collaboration. The results of the study that were identified therefore suggest that working with peers in kindergartens on mathematical problem solving has a significant positive effect. In general, positive correlations can also be found particularly in areas related to visual perception, visual-motor skills and counting tasks. The results provide to the existing knowledge of the importance of social interactions for children and the early childhood numeracy by providing the importance of peer collaboration and its possible influence on mathematical problem solving.

In summary, the study contributes valuable information and underlines the need to consider the limitations for a more reliable and complete research on the complex relationship between collaboration and early childhood numeracy.

References

Berch, D. B. E. (n.d.-b). *Why is math so hard for some children? The nature and origins of Mathematical learning Difficulties and Disabilities.*

Cornu, S. L., et al. (2017). The malleability of early numeracy: A longitudinal study. *Journal of Educational Psychology, 109*(3), 428-437.

Duncan, R. S., et al. (2007). *Long-term effects of early numeracy on academic achievement. Educational Psychology, 27*(4), 503-518.

French, G. (2013). Early literacy and numeracy matters.

Izard, A. B. (2009). *Early numeracy in young children*. *Journal of Early Numeracy, 4*(2), 123-137.

Mix, K. S., & Cheng, Y. L. (2012). Visuospatial skills and mathematics: What we know and what we need to know. In D. Berch & M. M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 73-89).

Reid, K. (2016). *Counting on it: Early numeracy development and the preschool child.* Australian Council for Educational Research (ACER). https://research.acer.edu.au/learning_processes/19

Starr, C. L. (2013). *Quantitative development in early childhood. Child Development Research*, *2013*, 1-12.

Wang, L., & Cook, S. A. (2017). Staff development programs on teaching skills and curriculum integration of academic and information literacy at the University of Auckland. In *Elsevier eBooks* (pp. 107–118).

Wynn, J. (1992). *The foundations of early numerical understanding. Trends in Cognitive Sciences*, *4*(12), 534-541.

Examining Math Performance and Math Anxiety In German and French Instructional Sessions

Arnotte Anne-Sophie, Becker Sara, de Nerée Vincent

Supervisor: PhD candidate Styliani Politi

Luxembourg's linguistic diversity, with Luxembourgish, German and French as its three official languages, provides a unique context for exploring the interplay between language and mathematical competence. This study delves into the relationship between math performance and math anxiety in a multilingual setting, focusing on sessions conducted in both German and French. Building on existing research, particularly Barroso et al.'s (2021) meta-analysis, which highlights a negative correlation between math anxiety and achievement, and Greisen et al.'s (2021) study on language proficiency's impact on mathematical achievement, our investigation aims to deepen understanding in this area. Our study, involving 71 multilingual participants examines the relationship between math anxiety and performance across instructional sessions in German and French. Despite language preferences, we found a consistent negative correlation, with mathematic ability mediating this relationship. These findings underscore the pervasive influence of math anxiety on performance, irrespective of the language of instruction. Future research should address limitations and explore language preferences and proficiency to enhance educational interventions, particularly in multilingual contexts like Luxembourg.

Introduction

With a large number of languages, including the three official languages, Luxembourgish, German and French, and a strong linguistic diversity, Luxembourg is an extremely polyglottic country (*Welche Sprachen spricht man in Luxemburg?* o. J.)This cultural diversity is reflected not only in the everyday life of its inhabitants, but also in the educational system, where students are taught in Luxembourgish, French and German from an early age. This multilingual approach prepares students for a multicultural society and reflects the nation's commitment to embracing its diverse heritage. This linguistic diversity also provides a unique opportunity to study the impact of different languages on mathematical performance. Of particular interest is the fact that non-native speakers often experience increased anxiety when faced with examinations in languages with which they are unfamiliar (Education nationale, Enfance et Jeunesse - Luxembourg, 2024). Understanding these relationships is crucial for developing educational strategies that can better support multilingual students and improve their math performance.

Research framework

Many studies have produced interesting findings on math performance in multilingual educational settings. The meta-analysis by Barroso et al. (2021) provides an in-depth examination of the complex dynamics between math anxiety and performance. Spanning from 1992 to 2018, this comprehensive study synthesized findings from 223 individual studies, comprising 747 correlation coefficients from 332 different samples and approximately 385,441 participants (Barroso et al., 2021).

The results of this meta-analysis show a consistent trend: a slightly to moderately negative correlation ($r = -0.28$) between math anxiety and achievement across different educational contexts. These findings demonstrate the pervasive impact of math anxiety on academic

achievement, from early education through to adulthood. Of particular note is the in-depth exploration of moderating factors, which highlights the complexity of this relationship (Barroso et al., 2021). The meta-analysis suggests that grade level, level of mathematic ability, type of anxiety scale, mathematic topic of the anxiety scale, mathematic assessments and publication status are crucial factors influencing the relationship between mathematic anxiety and mathematic achievement. The observed variability in effect sizes, as indicated by the Q and I2 statistics, highlights the diversity of contexts in which the relationship between math anxiety and achievement operates (Barroso et al., 2021). For example, in our study focusing on math achievement and anxiety among university students, understanding these contextual variations can shed light on why some students with high math anxiety perform differently across different types of math assessments or academic years. This variability not only complicates the interpretation of results, but also highlights the need to implement tailored interventions that take into account specific educational contexts and individual differences in math anxiety levels.

On the other hand, research by Greisen et al. (2021) examines the impact of language proficiency on mathematical achievement. By examining the performance of third graders in a highly multilingual educational environment, the study finds significant relationships between language proficiency, particularly in German reading comprehension, and mathematic achievement. The findings highlight the importance of considering language proficiency, especially in the context of the language of instruction, when assessing mathematic achievement in multilingual societies such as Luxembourg. The mediation analysis highlights the role of German reading comprehension in mediating the relationship between mother tongue proficiency and mathematic achievement. Overall, this study provides valuable insights into the factors influencing mathematic achievement in multilingual settings, with implications for educational practices and interventions aimed at supporting students from diverse linguistic backgrounds.

Additionally, da Silva Santos and Finger (2020) explore how bilingualism affects numerical cognition, highlighting that language familiarity impacts mathematical problem-solving. Specifically, their study aimed to determine if Portuguese-English bilinguals exhibit a language preference when reading numbers during addition problems presented in digit, Portuguese, or English formats. Results showed higher accuracy in answering addition problems when presented in Portuguese, followed by the digit format and then the English format. The participants' language background did not significantly affect their performance, with the main factor influencing response time and accuracy being the format in which numbers were presented. Bilinguals performed better in their preferred language for arithmetic, especially when calculations were presented in digit format (da Silva Santos & Finger, 2020).

Marchand et al. (2019) examined whether the language of the test impacted estimation performance in bilingual children aged 5 to 7 years. Results showed a significant three-way interaction of Age, Numerosity, and Language of Test, indicating differences in estimation behavior between French and English, especially among younger children. Even when considering only numbers within the child's highest common estimate, the interaction between Numerosity and Language of Test persisted, showing differences in accuracy between French and English estimates. This suggests that bilingual children have different mappings between approximate number system values and familiar number words in their two languages (Marchand et al., 2019).

Finally, Zaripova et al. (2019) examine the effects of the intensity of bilingual education on math achievement. Their findings on the potential influence of bilingualism on the development of abstract thinking are directly applicable. This study can be used to discuss the potential impact of varying language instruction intensities on math performance in Lux-

embourg's multilingual educational context (Zaripova et al., 2019).

All these studies help make the argument that even though there are reasons to believe that a multilingual educational context could be important in the study and understanding of math anxiety, up to date no studies have investigated this issue in multilingual and multicultural samples. This is particularly crucial for our project, as Luxembourgish students often switch the language of math instruction from German to French. This shift could significantly impact math anxiety and performance, given the established importance of language in math achievement.

This study

In our study, we aim to explore the relationship between math performance and math anxiety in multilingual individuals, building on these existing findings.

Specifically, we aim to investigate how these factors interact in both German (GE) and French (FR) testing contexts.

Hypotheses

Four hypotheses were developed on the basis of the results presented in the theoretical background. The aim of the study was to replicate the results of previous studies and meta-analyses and to investigate new correlations. The hypotheses can be differentiated according to their specificity. Two main hypotheses (H1.a and H1.b) and two sub-hypotheses (H2.a and H2.b) were formulated. The hypotheses we developed are listed below.

H1.a: Math anxiety (MA) is negatively associated with math performance (MP) in GE and FR sessions.
H1.b: The MA-MP relationship is the same in GE and FR sessions but with different effect sizes.
H2.a: The MA-MP relationship is mediated by math fluency (MF).
H2.b: The mediating role of MF is the same in GE and FR sessions.

Methods

Participants

In our study, participants had to meet certain inclusion criteria. They had to be proficient in German, French and English, be between 18 and 40 years old, have no learning difficulties and have attended a Luxembourgish school. These criteria ensured that participants had the necessary language skills and educational background to engage meaningfully with the study tasks.

Our sample consisted of 71 participants, aged between 18 and 34, with an average age of 22.35 (SD= 3.622). Among them, 25 preferred German, while 46 preferred French as their language of choice for math exams. Participants volunteered to take part and were recruited through posters and personal referrals as well as through the university online platform for recruiting participants. They were rewarded with €20 and credit hours for psychology students.

Material and procedure

The research procedure involved administering a diverse range of questionnaires as part of a larger research project. Below, we provide a detailed description of the procedure; however, it is important to note that this report focuses on analyzing only a specific portion of the data.
Given the multilingual context of Luxembourg, the choice of language in our study was crucial. Participants completed tasks in both German and French, reflecting the linguistic diversity of Luxembourg.
Our study consisted of two separate sessions:

First session: During the first session, participants completed a series of tasks designed to assess distinct aspects of their cognitive and mathematical abilities. First, they completed State-Trait Anxiety Questionnaires (STAI) to assess their anxiety levels, distinguishing between transient (state) and persistent (trait) anxiety (Spielberger et al., 1983). These questionnaires provided insight into participants' levels of general anxiety and whether they experienced anxiety as a temporary state or a

more persistent trait. Following the anxiety assessments, participants completed the Timed Arithmetic Test (TTR) to measure math fluency (TTR; De Vos, 1992). Here, participants are presented with 5 columns of 40 math operations each, totaling 200 problems per page. Each page focuses on a different type of mathematical operation: addition, subtraction, multiplication, division, and a mix of all four. This setup is designed to comprehensively assess the participants' math fluency across various basic arithmetic skills. This test was designed to measure participants' mathematical ability and performance under a time constraint, providing insight into their ability to solve mathematical problems quickly and accurately.

Following the TTR test, participants completed computerized tasks including a general knowledge quiz (GKQ) and a mathematical performance test (MPT). These tasks were designed to in these tasks, participants were instructed to solve as accurately and as quickly as possible, general knowledge questions in a multiple-choice format. Each trial began with a question shown on a light grey screen with a white background and black letters. The answer choices were listed below the question. The question stayed on the screen until the participant clicked on their chosen answer using the mouse or touchpad.

Participants then underwent an assessment specifically designed to assess their level of anxiety during cognitive tasks, known as Cognitive Test Anxiety (CTAS-2; Thomas et al., 2016). The aim of this component was to determine the extent to which participants experienced anxiety when undertaking a cognitive task.

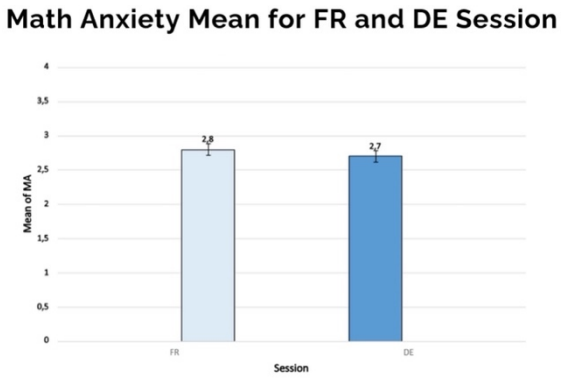
Finally, participants completed assessments to assess their anxiety in relation to mathematics, known as the Math Anxiety (MA) test. The Math Anxiety Questionnaire comprises 20 items that assess math anxiety across four dimensions: general, computational, geometrical, and linguistic (see Marinova et al., 2023). This questionnaire was newly developed, incorporating novel items and drawing from previously validated tools such as the Abbreviated Math Anxiety Scale (Hopko et al.,2003) and the Children's Anxiety in Math (Jameson, 2013) Scale. Participants indicated how much each statement related to them on a Likert scale ranging from 1 (not true at all) to 5 (very true). The questionnaire was administered in both German and French, with the item order randomized in each session. This test focused on assessing participants' anxiety specifically in mathematical contexts.

Second session:
Participants completed a subset of tasks compared to the first session. Specifically, they completed computerized tasks, the Cognitive Test Anxiety (CTA) and the Math Anxiety (MA) tests. This streamlined approach allowed us to focus on key aspects of participants' cognitive and anxiety-related responses in the alternate language. In the second session, we deferred tasks such as the general questionnaire and the TTR test, given that participant age and general information are unlikely to undergo significant change. This was done in order to concentrate on the collection of essential data on anxiety levels and mathematical performance in an efficient manner. This session provided further insight into participants' responses in the same problems but in a different linguistic context, thereby enhancing our understanding of the role of language in mathematical performance and anxiety.

Results

The mean scores for math anxiety (MA) did not significantly differ depending on the language of the session (German or French) ($t(70) = -0.932$, $p = 0.355$, $d = -0.111$). Similarly, the mean scores of math performance did not significantly differ between the two language sessions ($t(70) = 1.166$, $p = 0.248$, $d =$

0.138).

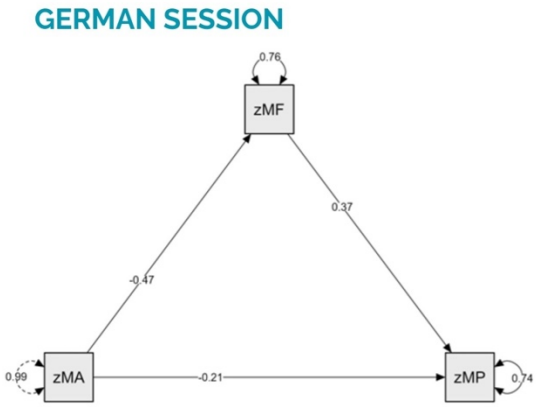


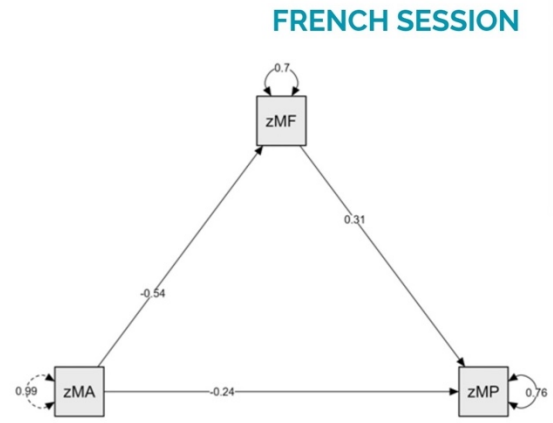
Findings on H1

To address our initial hypothesis, we conducted a Pearson's correlation analysis between MA and MP, revealing a negative relationship in German ($r = -0.405$, $p < 0.001$) and in French ($r = -0.383$, $p < 0.001$). To investigate whether effect sizes significantly differed between the German and French session we used Fisher's r-to-z transformation. However, the analysis showed no significant differences in effect sizes ($p = 0.924$). The relationship between MA and MP remained consistent for both French and German sessions, with similar effect sizes.

Findings on H2

Given the established relationship in the literature between math fluency (MF), MA, and MP, we performed a mediation analysis to explore the extent of MF's mediating role. Our independent variable (IV) was MA, the dependent variable (DV) was MP, and MF acted as the mediator. The analysis revealed that MA significantly influenced MF, and MF significantly influenced MP, with a marginal direct effect between MA and MP. The analysis revealed that MA significantly influenced MF (path a), and MF significantly influenced MP (path b), with a marginal direct effect between MA and MP (path c').





Mediation Analysis Results

Direct Effect (path c'): The direct effect of math anxiety on math performance was marginally significant (Estimate = -0.239, SE = 0.123, p = 0.052, 95% CI = [-0.481, 0.002]).

Indirect Effect: Math fluency significantly mediated the relationship between math anxiety and math performance (Estimate = -0.166, SE = 0.073, p = 0.024, 95% CI = [-0.309, -0.022]), suggesting that the primary influence of math anxiety on performance is indirect.

Total Effect: The total effect showed that math anxiety has a significant negative impact on math performance when considering indirect pathways
(Estimate = -0.405, SE = 0.109, p < 0.001, 95% CI = [-0.618, -0.193]).

These findings confirm that the relationship between math anxiety and math performance is mediated by math fluency in both German

and French sessions, supporting our hypothesis about the mediating role of MF.

Limitations-Surmounting Research Restraints

Future research should conduct studies with larger thus more diverse samples so as to improve the generalizability of the findings and utilize longitudinal designs to investigate the temporal dynamics and causal relationships between math anxiety, instructional language, and math performance with the aim to overcome the limitations of our study. For the purpose of having a more complete picture of math anxiety and performance, future research should concentrate on the use of a combination of self-reported measures, behavioral assessments, and teacher reports. Furthermore, in the hope of fully comprehending the impact that language competency plays in math anxiety and performance, the study design and analysis should take contextual aspects like teaching quality, the classroom setting, and socioeconomic status into account. Additionally, comprehensive assessments of language proficiency should be included. And lastly, consider other psychological factors, such as general anxiety and test anxiety, to provide a more holistic understanding of the influences on math performance.

Discussion

Talking about the connection between math anxiety and math performance, with an emphasis on our most recent findings and possible paths for future study and intervention:

Low Math Performance and High Math Anxiety Levels

First off, poorer math performance has been linked to high levels of math anxiety, as has been extensively studied. Our study, which is in line with earlier research by Barroso et al., (2021) discovered that arithmetic anxiety had a detrimental effect on academic results, particularly math performance. For instance, Ashcraft and Krause (2007) demonstrated that individuals with high math anxiety tend to perform worse on math tasks due to increased cognitive load. Additionally, Hembree (1990) showed that math anxiety significantly hampers problem-solving skills and overall math achievement. Our findings confirmed the negative correlation between math anxiety and math performance, indicating that math anxiety substantially impairs the ability to perform well in arithmetic. Moreover, our study replicated and extended these findings in multilingual participants.

Math Fluency as a key mediator

We then looked at arithmetic fluency as a crucial mediator in this relationship. According to our research, math anxiety can be lessened by increasing one's fluency in the subject. This suggests that programs designed to improve arithmetic fluency may lessen the negative effects of math anxiety on academic performance. This finding aligns with the work of Ramirez et al. (2013) who demonstrated that students with higher math fluency experienced less anxiety and performed better on math assessments. Our study replicated and extended these findings by highlighting the importance of arithmetic fluency in multilingual participants.

Instructional Language's Function and Role

We examined the role of the language used in instruction on the relationship between math anxiety and math performance. Our initial hypothesis was that math anxiety would differ between the German and French sessions due to participants' linguistic backgrounds, particularly since their first language for learning math was German before switching to French. The lack of a significant difference in math anxiety between the German and French sessions could be due to several factors. For instance, participants might have developed coping mechanisms that mitigate the anxiety's impact when switching languages, or their proficiency in both languages could have reached a level where the language of instruction no longer makes a significant difference. Additionally, the consistency in the negative correlation suggests that interventions to reduce math anxiety could be effective across multiple

languages, potentially benefiting multilingual learners in diverse educational settings.

Potential Advancements for the Future

Looking ahead, several potential advancements could enhance our understanding of the impact of language on math anxiety and performance. Future research could investigate how different levels of proficiency in a second language impact the relationship between math anxiety and math performance. This would involve evaluating language competency levels and their influence on math-related anxiety and outcomes. Furthermore, determining the preferred language for math instruction among multilingual students could reveal whether instruction in a student's dominant language reduces math anxiety more effectively than in a second language. What's more, offering more in-depth insights into how language preference and proficiency affect math performance and anxiety could lead to more targeted interventions. Understanding these dynamics would allow for the development of instructional approaches and treatments that better support students' learning and well-being in arithmetic. By addressing these areas, future research can build upon our findings to create more effective interventions and educational strategies, ultimately enhancing the support provided to multilingual learners in diverse educational settings. By including determining the preferred language for math instruction, evaluating language competency, and offering more in-depth understanding of how language preference and proficiency affect math performance and anxiety, could advance our understanding of this problem.

All things considered; by addressing these limitations, future research can build upon our findings to develop more effective interventions as well as instructional, educational approaches and treatments that are more effective in supporting students' learning and wellbeing in arithmetic.

References

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review*, *14*(2), 243–248. https://doi.org/10.3758/BF03194059B

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). *A meta-analysis of the relation between math anxiety and math achievement.* Psychological Bulletin, 147(2), 134–168.

da Silva Santos, C., & Finger, I. (2020). Bilingualism and numerical cognition. *Veredas - Revista de Estudos Linguísticos*, *24*(1), 340–360. https://doi.org/10.34019/1982-2243.2020.v24.30632

De Vos, T. (1992). *Arithmetic Number Fact Test*.

Education nationale, Enfance et Jeunesse - Luxembourg. (2024, Juni 12). Public.lu. https://men.public.lu/fr.html

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), 103456. https://doi.org/10.1016/j.actpsy.2021.103456

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for research in mathematics education*, *21*(1), 33. https://doi.org/10.2307/749455

Hopko, D. R., Mahadevan, R., Bare, R. L., & Hunt, M. K. (2003). The Abbreviated Math Anxiety Scale (AMAS): Construction, validity, and reliability. *Assessment*, *10*(2), 178–182. https://doi.org/10.1177/1073191103010002008

Jameson, M. M. (2013). The development and validation of the children's anxiety in Math Scale. *Journal of Psychoeducational Assessment*, *31*(4), 391–395. https://doi.org/10.1177/0734282912470131

Marchand, E., Wade, S., Sullivan, J., & Barner, D. (2019). Language-specific numerical estimation in bilingual children. In *PsyArXiv*. https://doi.org/10.31234/osf.io/afdg8

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development: Official Journal of the Cognitive Development Society*, *14*(2), 187–202. https://doi.org/10.1080/15248372.2012.664593

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press.

Thomas, C. L., Cassady, J. C., & Finch, W. H. (2018). Identifying severity standards on the Cognitive Test Anxiety Scale: Cut score determination using latent class and cluster analysis. *Journal of Psychoeducational Assessment*, *36*(5), 492–508. https://doi.org/10.1177/0734282916686004

Vos, H., Marinova, M., De Léon, S. C., Sasanguie, D., & Reynvoet, B. (2023). Gender differences in young adults' mathematical performance: Examining the contribution of working memory, math anxiety and gender-related stereotypes. *Learning and Individual Differences*, *102*(102255), 102255. https://doi.org/10.1016/j.lindif.2022.102255

Welche Sprachen spricht man in Luxemburg? (o. J.). Public.lu. Abgerufen 14. Juni 2024, 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

Impressum

Luxemburger Experimentalpraktikum Journal
Band 18, Jahrgang 2024

Herausgeber

Dr. Andreia Costa
Koordinatorin des Practical Training in Empiricial Research
Université du Luxembourg
Maison Sciences Humaines
11, Porte des Sciences
L-4366 Esch-sur-Alzette

Redaktion

Clémentine Offner, Studentin des BAP

Mitherausgeber

Gary Lee Wagener
Dominika Repcikova
Kriti Kelkar
Dr. André Melzer
Hoyeon Lee
Styliani Politi

Manuskriptrichtlinien

Die Beiträge richten sich nach dem „Publication Manual of the American Psychological Association" (2010) bzw. den „Zur Gestaltung von Haus- und Abschlussarbeiten" (2004) der Deutschen Gesellschaft für Psychologie.

Erscheinungsweise

Ein- bis zweimal jährlich

Bezug

Ass.-Prof. André Melzer, Université du Luxembourg, Maison Sciences Humaines 11, Porte des Sciences, L-4366 Esch-sur-Alzette

Online unter

https://bap.uni.lu

ISSN 3093-1045