Question	**Hypothesis**	**Sampling plan (e.g. power analysis)**	**Analysis Plan**	**Interpretation given to different outcomes**
1.) Do ER strategies reduce emotional arousal? (Manipulation check)	1a) Subjective arousal (arousal rating) is lower after using an emotion regulation strategy (distraction, distancing, suppression) compared to active viewing.	F tests - ANOVA: Repeated measures, within factors Analysis: A priori: Compute required sample size <u>Input:</u> Effect size f = 0.50 (η_p^2 = 0.20) (Scheffel et al., 2021) α err prob = 0.05 Power (1-β err prob) = 0.95 Number of groups = 1 Number of measurements = 4 Corr among rep measures = 0.5 Nonsphericity correction ε = 1 <u>Output</u>: Noncentrality parameter λ = 20.0 Critical F = 2.9603513 Numerator df = 3.0 Denominator df = 27.0 Total sample size = 10 Actual power = 0.95210128	Repeated measures ANOVA with four linear contrasts, comparing the subjective arousal ratings of four blocks (active viewing, distraction, distancing, suppression). ANOVA is calculated using aov_ez() function of the afex-package, estimated maginal means are calculated using emmeans() function from the emmeans-package, pairwise contrasts are calculated using pairs(). Bayes factors are computed for the ANOVA and each contrast using the BayesFactor-package.	ANOVA yields $p < .05$ is interpreted as arousal ratings changing significantly with blocks. Values of arousal ratings are interpreted as equal between blocks if $p > .05$. Each contrast yielding $p < .05$ is interpreted as arousal ratings being different between those two blocks, magnitude and direction are inferred from the respective estimate. Values of arousal ratings are interpreted as equal between blocks if $p > .05$. The Bayes factor *BF10* is reported alongside every *p*-value to assess the strength of evidence.
	1b) Physiological arousal (*corrugator* muscle activity) is lower after using an emotion regulation strategy (distraction, distancing, suppression) compared to active viewing.	F tests - ANOVA: Repeated measures, within factors Analysis: A priori: Compute required sample size <u>Input:</u> Effect size f = 0.1605 (Zaehringer et al., 2020) α err prob = 0.05 Power (1-β err prob) = 0.95 Number of groups = 1 Number of measurements = 4	Repeated measures ANOVA with four linear contrasts, comparing the *corrugator* muscle activity of four blocks (active viewing, distraction, distancing, suppression). ANOVA is calculated using aov_ez() function of the afex-package, estimated maginal means are calculated using	ANOVA yields $p < .05$ is interpreted as *corrugator* muscle activity changing significantly with blocks. Values of *corrugator* muscle activity are interpreted as equal between blocks if $p > .05$. Each contrast yielding $p < .05$ is interpreted as *corrugator* muscle activity being different between those two blocks, magnitude and direction are

		Corr among rep measures = 0.5 Nonsphericity correction ε = 1 <u>Output</u>: Noncentrality parameter λ = 17.5169700 Critical F = 2.6404222 Numerator df = 3.0 Denominator df = 252 Total sample size = 85 Actual power = 0.9509128	emmeans() function from the emmeans-package, pairwise contrasts are calculated using pairs(). Bayes factors are computed for the ANOVA and each contrast using the BayesFactor-package.	inferred from the respective estimate. Values of *corrugator* muscle activity are interpreted as equal between blocks if $p > .05$. The Bayes factor *BF10* is reported alongside every *p*-value to assess the strength of evidence.
	1c) Physiological arousal (*levator* muscle activity) is lower after using an emotion regulation strategy (distraction, distancing, suppression) compared to active viewing.	F tests - ANOVA: Repeated measures, within factors Analysis: A priori: Compute required sample size <u>Input:</u> Effect size f = 0.1605 (Zaehringer et al., 2020) α err prob = 0.05 Power (1-β err prob) = 0.95 Number of groups = 1 Number of measurements = 4 Corr among rep measures = 0.5 Nonsphericity correction ε = 1 <u>Output</u>: Noncentrality parameter λ = 17.5169700 Critical F = 2.6404222 Numerator df = 3.0 Denominator df = 252 Total sample size = 85 Actual power = 0.9509128	Repeated measures ANOVA with four linear contrasts, comparing the *levator* muscle activity of four blocks (active viewing, distraction, distancing, suppression). ANOVA is calculated using aov_ez() function of the afex-package, estimated maginal means are calculated using emmeans() function from the emmeans-package, pairwise contrasts are calculated using pairs(). Bayes factors are computed for the ANOVA and each contrast using the BayesFactor-package.	ANOVA yields $p < .05$ is interpreted as *levator* muscle activity changing significantly with blocks. Values of *levator* muscle activity are interpreted as equal between blocks if $p > .05$. Each contrast yielding $p < .05$ is interpreted as *levator* muscle activity being different between those two blocks, magnitude and direction are inferred from the respective estimate. Values of *levator* muscle activity are interpreted as equal between blocks if $p > .05$. The Bayes factor *BF10* is reported alongside every *p*-value to assess the strength of evidence.

2.) Do ER strategies require cognitive effort? (Manipulation check)	2a) Subjective effort (effort ratings) is greater after using an emotion regulation strategy (distraction, distancing, suppression) compared to active viewing.	F tests - ANOVA: Repeated measures, within factors Analysis: A priori: Compute required sample size <u>Input:</u> Effect size f = 0.2041241 (η_p^2 = 0.04) (Scheffel et al., 2021) α err prob = 0.05 Power (1-β err prob) = 0.95 Number of groups = 1 Number of measurements = 4 Corr among rep measures = 0.5 Nonsphericity correction ε = 1 <u>Output</u>: Noncentrality parameter λ = 17.6666588 Critical F = 2.6625685 Numerator df = 3.0 Denominator df = 156.0 Total sample size = 53 Actual power = 0.95206921	Repeated measures ANOVA with four linear contrasts, comparing the subjective effort ratings of four blocks (active viewing, distraction, distancing, suppression). ANOVA is calculated using aov_ez() function of the afex-package, estimated maginal means are calculated using emmeans() function from the emmeans-package, pairwise contrasts are calculated using pairs(). Bayes factors are computed for the ANOVA and each contrast using the BayesFactor-package.	ANOVA yields $p < .05$ is interpreted as effort ratings changing significantly with blocks. Values of effort ratings are interpreted as equal between blocks if $p > .05$. Each contrast yielding $p < .05$ is interpreted as effort ratings being different between those two blocks, magnitude and direction are inferred from the respective estimate. Values of effort ratings are interpreted as equal between blocks if $p > .05$. The Bayes factor *BF10* is reported alongside every *p*-value to assess the strength of evidence.
	2b) Majority of participants reuse the strategy that was least effortful for them.	-	Subjects are asked about the reasons for their choice in the follow-up survey. These answers are classified into categories and counted.	The percentage choice of strategies is described descriptively.
3.) Which variables can predict individual subjective values of ER strategies?	3a) Subjective effort ratings negatively predict subjective values of ER strategies.	t tests - Linear multiple regression: Fixed model, single regression coefficient Analysis: A priori: Compute required sample size <u>Input</u>:	Multilevel model of SVs with level-1-predictors subjective effort, subjective arousal, *corrugator*, and *levator* muscle activity using subject specific	Fixed effects yield $p < .05$ are interpreted as subjective values are related to subjective effort. Subjective values are interpreted as not being related to subjective effort if $p > .05$.

		Tail(s) = One Effect size f^2 = 0.34 (Since there are no findings in this respect yet, we have inferred from the effect size in the closest-similar model: Westbrook et al., 2013) α err prob = 0.05 Power (1-β err prob) = 0.95 Number of predictors = 4 <u>Output</u>: Noncentrality parameter δ = 3.4 Critical t = 1.6991270 Df = 29 Total sample size = 34 Actual power = 0.9529571	intercepts and allowing random slopes for ER strategies. The null model and the random slopes model are calculated using lmer() of the lmerTest-package. Bayes factors are computed for the MLM using the BayesFactor-package.	The Bayes factor *BF10* is reported alongside every *p*-value to assess the strength of evidence.
	3b) Subjective arousal ratings negatively predict subjective values of ER strategies.			Fixed effects yield $p < .05$ are interpreted as subjective values are related to subjective arousal. Subjective values are interpreted as not being related to subjective arousal if $p > .05$. The Bayes factor *BF10* is reported alongside every *p*-value to assess the strength of evidence.
	3c) *Corrugator* muscle activity negatively predict subjective values of ER strategies.			Fixed effects yield $p < .05$ are interpreted as subjective values are related to *corrugator* activity. Subjective values are interpreted as not being related to *corrugator* activity if $p > .05$. The Bayes factor *BF10* is reported alongside every *p*-value to assess the strength of evidence.
	3d) *Levator* muscle activity negatively predict subjective values of ER strategies.			Fixed effects yield $p < .05$ are interpreted as subjective values are related to *levator* activity. Subjective values are interpreted as not being related to *levator* activity if $p > .05$. The Bayes factor *BF10* is reported alongside every *p*-value to assess the strength of evidence.

4.) Is the effort required for an ER strategy the best predictor for subjective values of ER strategies?	4a) Subjective values decline with increasing effort, even after controlling for task performance measured by subjective arousal ratings, *corrugator* and *levator* muscle activity.	t tests - Linear multiple regression: Fixed model, single regression coefficient Analysis: A priori: Compute required sample size <u>Input</u>: Tail(s) = One Effect size $f^2 = 0.34$ (Since there are no findings in this respect yet, we have inferred from the effect size in the closest-similar model: Westbrook et al., 2013) α err prob = 0.05 Power (1-β err prob) = 0.95 Number of predictors = 4 <u>Output</u>: Noncentrality parameter δ = 3.4 Critical t = 1.6991270 Df = 29 Total sample size = 34 Actual power = 0.9529571		Fixed effects yield $p < .05$ are interpreted as subjective values changing significantly with ER strategy. Subjective values are interpreted as equal between ER strategies if $p > .05$. The Bayes factor *BF10* is reported alongside every *p*-value to assess the strength of evidence.
5.) Are subjective values related to flexible emotion regulation?	5a) The higher the subjective value, the more likely the respective strategy is chosen.			
	5b) Subjective values are lower and decline stronger when ER flexibility is lower.	t tests – Linear multiple regression: Fixed model, single regression coefficient Analysis: A priori: compute required sample size <u>Input:</u> Tail(s) = One	SVs will be ordered by magnitude. Values will be fitted in a GLM to estimate the individual intercept and slope. A linear regression will be computed with intercept and	β yield $p < .05$ are interpreted as significant association between predictor (intercept, slope) and ER flexibility. The direction of effect is interpreted according to sign (negative or positive). *p* – values > .05 are interpreted as no association between predictor and ER flexibility.

		Effect size f^2 = 015 (as there is no evidence in the literature, we assume a medium sized effect) α err prob = 0.05 Power (1-β err prob) = 0.95 Number of predictors = 2 <u>Output:</u> Noncentrality parameter δ = 3.316662 Critical t = 1.69665997 Df = 71 Total sample size = 74 Actual power = 0.95101851	slope as predictors and FlexER score as criterion.	The Bayes factor *BF10* is reported alongside every *p*-value to assess the strength of evidence.
Exploratory: Are individual subjective values of ER strategies related to personality traits?			Multilevel model of SVs with level-1-predictors subjective effort, subjective arousal, *corrugator*, and *levator* muscle activity and level-2-predictors NFC and self-control using subject specific intercepts and allowing random slopes for ER strategies. The null model and the random slopes model are calculated using lmer() of the lmerTest-package. Bayes factors are computed for the MLM using the BayesFactor-package.	Fixed effects yield $p < .05$ are interpreted as subjective values are related to NFC and self-control. Subjective values are interpreted as not being related to subjective effort if $p > .05$. The Bayes factor *BF10* is reported alongside every *p*-value to assess the strength of evidence.