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 Input: Effect size $f = 0.50 \ (\eta_p^2 = 0.20)$ (Scheffel et al., 2021) α err prob = 0.05 Power $(1-\beta \text{ err prob}) = 0.95$ Number of groups = 1 Number of measurements = 4 Corr among rep measures = 0.5 Nonsphericity correction $\epsilon = 1$ Output: Noncentrality parameter $\lambda = 20.0$ 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 Input: Effect size $f = 0.1605$ (Zaehringer et al., 2020) α err prob = 0.05 Power $(1-\beta$ err prob) = 0.95 Number of groups = 1 Number of measurements = 4	Repeated measures ANOVA with four linear contrasts, comparing the <i>corrugator</i> muscle activity of four blocks (active viewing, distraction, distancing, suppression). ANOVA is calculated using aov_ez() function of the afexpackage, 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 ran maggurag = 0.5	emmeans() function from the	informed from the respective estimate
	Corr among rep measures = 0.5	· · · · · · · · · · · · · · · · · · ·	inferred from the respective estimate.
	Nonsphericity correction $\varepsilon = 1$	emmeans-package, pairwise	Values of <i>corrugator</i> muscle activity
		contrasts are calculated using	are interpreted as equal between blocks
	Output:	pairs().	if $p > .05$.
	Noncentrality parameter $\lambda =$		
	17.5169700	Bayes factors are computed for	The Bayes factor <i>BF10</i> is reported
	Critical $F = 2.6404222$	the ANOVA and each contrast	alongside every <i>p</i> -value to assess the
	Numerator $df = 3.0$	using the BayesFactor-package.	strength of evidence.
	Denominator $df = 252$		
	Total sample size = 85		
	Actual power = 0.9509128		
1c) Physiological	F tests - ANOVA: Repeated	Repeated measures ANOVA	ANOVA yields $p < .05$ is interpreted as
arousal (levator	measures, within factors	with four linear contrasts,	levator muscle activity changing
muscle activity) is	Analysis: A priori: Compute	comparing the <i>levator</i> muscle	significantly with blocks. Values of
lower after using an	required sample size	activity of four blocks (active	levator muscle activity are interpreted
emotion regulation	Input:	viewing, distraction, distancing,	as equal between blocks if $p > .05$.
strategy (distraction,	$\overline{\text{Effect}}$ size f = 0.1605 (Zaehringer	suppression).	as equal between blocks if $p > .05$.
distancing,	et al., 2020)	11 /	Each contrast yielding $p < .05$ is
suppression) compared	$\alpha \text{ err prob} = 0.05$	ANOVA is calculated using	interpreted as <i>levator</i> muscle activity
to active viewing.	Power $(1-\beta \text{ err prob}) = 0.95$	aov_ez() function of the afex-	being different between those two
to deal to the triang.	Number of groups = 1	package, estimated maginal	
	Number of groups = 1 Number of measurements = 4	means are calculated using	blocks, magnitude and direction are
	Corr among rep measures = 0.5	emmeans() function from the	inferred from the respective estimate.
	Nonsphericity correction $\varepsilon = 1$	emmeans-package, pairwise	Values of <i>levator</i> muscle activity are
	Tronspherietty correction $e - 1$	contrasts are calculated using	interpreted as equal between blocks if <i>p</i>
	Output:	pairs().	> .05.
	Noncentrality parameter $\lambda =$	pans().	
		David factors on commutal for	The Bayes factor <i>BF10</i> is reported
	17.5169700	Bayes factors are computed for the ANOVA and each contrast	alongside every <i>p</i> -value to assess the
	Critical $F = 2.6404222$		strength of evidence.
	Numerator $df = 3.0$	using the BayesFactor-package.	
	Denominator $df = 252$		
	Total sample size = 85		
	Actual power = 0.9509128		

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 Input: Effect size $f = 0.2041241$ ($\eta_p^2 = 0.04$) (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 $\epsilon = 1$ Output: Noncentrality parameter $\lambda = 17.6666588$ 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 Input:	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 $p > .05$.

3b) Subjective arousal ratings negatively predict subjective values of ER strategies. 3c) Corrugator muscle activity negatively predict subjective values of ER strategies. 3d) Levator muscle activity negatively predict subjective values of ER strategies.	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 Output: 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. 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. 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. Fixed effects yield $p < .05$ are interpreted as subjective values are related to $levator$ activity. 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, <i>corrugator</i> and <i>levator</i> muscle activity.	t tests - Linear multiple regression: Fixed model, single regression coefficient Analysis: A priori: Compute required sample size Input: 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 $(1-\beta$ err prob) = 0.95 Number of predictors = 4 Output: Noncentrality parameter $\delta = 3.4$ 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) 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 Input: Tail(s) = One 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 $(1-\beta$ err prob) = 0.95 Number of predictors = 2 Output:	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 slope as predictors and FlexER score as criterion.	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. The Bayes factor $BF10$ is reported alongside every p -value to assess the strength of evidence.

	Noncentrality parameter $\delta = 3.316662$ Critical $t = 1.69665997$ Df = 71 Total sample size = 74 Actual power = 0.95101851		
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.	