

Identification of Affective- and Social-Reinforcement Functions of Driven Exercise: Evidence From Three Samples



David R. Kolar¹, Ann F. Haynos², Shirley B. Wang³,
Theresa Lask⁴, Stuart B. Murray⁵, Ulrich Voderholzer^{6,7,8},
and Sasha Gorrell⁹

¹Department of Psychology, University of Regensburg; ²Department of Psychology, Virginia Commonwealth University; ³Department of Psychology, Yale University; ⁴Department of Child and Adolescent Psychiatry, University Medicine Mainz; ⁵Department of Psychiatry and the Behavioral Sciences, University of Southern California, Los Angeles; ⁶Department of Psychiatry and Psychotherapy, University Hospital, Ludwig Maximilian University of Munich; ⁷Schoen Clinic Roseneck, Prien am Chiemsee, Germany; ⁸Department of Psychiatry and Psychotherapy, University Hospital of Freiburg, University of Freiburg; and ⁹Department of Psychiatry and Behavioral Sciences, UCSF Weill Institute for Neurosciences, University of California, San Francisco

Clinical Psychological Science
1–16

© The Author(s) 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/21677026241290455

www.psychologicalscience.org/CPS



Abstract

Driven exercise is a transdiagnostic maladaptive behavior, especially common in eating disorders (EDs); however, its maintenance mechanisms remain poorly understood. To address this gap, we examined the applicability of previously empirically derived and validated reinforcement-function models in explaining driven-exercise maintenance in studies varying across developmental stages and clinical presentations (Study 1: $N = 279$ adolescents/adults oversampled for EDs; Study 2: $N = 118$ adolescent/adult inpatients with severe EDs; Study 3: $N = 52$ adults oversampled for athletes and/or EDs). Results supported the utility of a four-function model (automatic positive reinforcement [APR]/automatic negative reinforcement [ANR; increase positive/decrease negative affect], social positive reinforcement [SPR]/social negative reinforcement [SNR; engage in/avoid interpersonal situations]) in explaining driven exercise. APR was most frequently endorsed, followed by ANR, SNR, and SPR in all studies. APR correlated with last-month driven-exercise episodes across studies; associations between other functions and ED psychopathology varied between studies. Furthermore, results suggested a separate control function could be considered in samples with more severe EDs.

Keywords

eating disorders, compulsive exercise, exercise, athletes, social function

Received 3/8/23; Revision accepted 8/8/24

Driven exercise (i.e., feeling compelled or urged to engage in physical activity or exercise) is a common transdiagnostic behavior, especially in eating disorders (EDs), but also frequently observed in the general population. For instance, driven-exercise behaviors are often reported in samples of adolescents and young adults (Ganson et al., 2022; Goodwin et al., 2012, 2014), gym attendees (Lejoyeux et al., 2008), people that regularly participate in high-intensity functional-movement training (Lichtenstein & Jensen, 2016), fitness instructors (Gjestvang et al., 2021), and college students (Edlund et al., 2022). Prevalence rates of driven exercise

have ranged from 5% to 42% in these studies; increased rates have occurred in samples diagnosed with EDs, including 66% of individuals with bulimia nervosa (Stiles-Shields et al., 2015) and 80% of individuals with anorexia nervosa (Dalle Grave et al., 2008). Typically, driven exercise is aerobic activity that is excessive in intensity, duration, or frequency (Solenberger, 2001) and continued even when adverse consequences, such as

Corresponding Author:

David R. Kolar, Department of Psychology, University of Regensburg

Email: david.kolar@psychologie.uni-regensburg.de

pain or injuries, occur (Lichtenstein & Jensen, 2016; Noetel et al., 2017). Besides being associated with ED behaviors, driven exercise is also associated with significant emotion-regulation difficulties (Goodwin et al., 2012, 2014), general mental-health problems (Ganson et al., 2022; Gjestvang et al., 2021), and depression (Edlund et al., 2022; Gjestvang et al., 2021).

Driven exercise has been historically conceptualized as either “exercise dependence” (Hausenblas & Downs, 2002) or “compulsive exercise” (Meyer et al., 2011). In brief, exercise dependence highlights that engagement in driven exercise often leads to a need to increase the amount of exercise (tolerance) and is associated with withdrawal symptoms (e.g., negative affect) and loss of control over when and for how much time a person exercises (Hausenblas & Downs, 2002). The compulsive-exercise model, however, emphasizes affect regulation together with compulsive (e.g., perceived negative consequences such as feeling guilty when stopping exercise routines), perfectionistic, and rigid (e.g., rule-driven exercising) personality traits as main components of the behavior (Meyer et al., 2011). There is some agreement that the compulsive-exercise model is a better fit to describe driven exercise, especially in EDs (Scharmer et al., 2020); however, some model components did not fit well across different samples (A. Harris et al., 2020; Schlegl et al., 2022). Although these models are derived from distinct lines of research, both emphasize short-term affective consequences of driven exercise as key components in the maintenance of the behavior (for a comparison, see Kolar & Gorrell, 2020a, p. 281). There is empirical support for affective change after physical activity and exercise in individuals with EDs from ecological-momentary-assessment (EMA) studies that have shown short-term increases in positive affect and decreases in negative affect following driven exercise (Engel et al., 2013; Kolar et al., 2020; Lekkas et al., 2022; Reichert et al., 2020). Recent research also indicated that affective trajectories after exercising may differ depending on whether the exercise itself was subjectively classified as adaptive or maladaptive (the latter category including driven exercise), at least for positive affect and in individuals with binge-spectrum EDs (Lampe et al., 2023). Other maintaining factors of driven exercise have been proposed beyond affective reinforcement. For instance, recent research indicates that interpersonal processes might also be related to driven exercise, at least in individuals with anorexia nervosa, such as sensitivity toward (perceived) social threats (Kolar, Kaurin, et al., 2022).

Although theoretical models have been proposed to explain the maintenance of driven exercise, little research has investigated the subjectively deemed intended outcomes of driven exercise reported by

individuals engaging in this behavior. Using such a functional perspective might elucidate the purpose of driven-exercise behaviors and which processes ultimately maintain them. A useful framework to identify potential functional-maintenance mechanisms of driven exercise is the four-function model by Nock and Prinstein (2004). This model posits that maladaptive behaviors are reinforced by automatic or intrapersonal and social or interpersonal functions. Automatic negative reinforcement suggests that the behavior is intended to alleviate negative affective states, whereas behaviors enacted for automatic positive reinforcement function to increase positive affective states. Social negative reinforcement occurs when a behavior serves to avoid unpleasant social experiences, whereas social positive reinforcement occurs when the behavior seeks to increase desired interactions with others, such as being comforted. There is evidence supporting the four-function model in nonsuicidal self-injury (NSSI; Hepp et al., 2020; Nock & Prinstein, 2004), binge eating and purging (Wedig & Nock, 2010), and restrictive eating (Wang et al., 2021). In these studies, NSSI, binge eating, and purging were found to predominantly serve automatic-negative-reinforcement functions, whereas Wang et al. (2021) found that automatic-positive-reinforcement functions were most highly endorsed for restrictive eating across three studies, with only little variability across type and severity of EDs and developmental stage. Control has been discussed as an additional independent maintenance mechanism for ED behaviors (Wang et al., 2021; Wedig & Nock, 2010); however, in prior studies, the original four-function model has provided a better fit to the data than a model incorporating control as a fifth factor. However, the importance of control functions has been emphasized in previous, more broad assessments of maintenance functions in EDs (Mikschl et al., 2016; Slade, 1982); therefore, engaging in exercise to feel in control may serve as a reinforcer for driven exercise for individuals with elevated ED symptoms, extending the four-function model to a five-function model.

Despite the demonstrated relevance of functional models to other ED behaviors, a functional analysis that applies the four-function framework to driven exercise is missing, especially across different populations with varying ED severity and levels of physical activity. Because driven exercise occurs frequently in individuals with anorexia nervosa, high levels of automatic positive and negative reinforcement may be expected. However, given prior emphasis on control as an important maintenance function of EDs in general (Mikschl et al., 2016; Slade, 1982), we also exploratorily examined if the addition of a fifth factor, control, improved functional models of driven exercise. To investigate

potential maintenance functions in driven exercise, we conducted three studies across different cultural contexts (Germany and United States), age groups (older adolescents and adults), ED presentations and severity (including an ED inpatient sample), and physical-activity levels (i.e., oversampling for athletes).

Study 1 assessed functions of driven exercise in a German community sample that oversampled for individuals with EDs. We hypothesized that (a) a four-factor structure (automatic positive/negative-reinforcement and social positive/negative-reinforcement functions) or five-factor structure (with an additional control function) similar to restrictive eating (Wang et al., 2021) would emerge; (b) functions would be endorsed to varying degrees; (c) automatic functions of driven exercise would demonstrate positive correlations with emotion-regulation difficulties, depression, anxiety and stress, body dissatisfaction, compulsive exercise severity, and ED severity; and (d) social functions of driven exercise would demonstrate negative correlations with perceived social support and positive correlations with ED severity. In Study 2, we investigated whether the same hypotheses as in Study 1 would replicate in a sample of German adolescent and adult inpatients with severe EDs. In Study 3, we aimed to replicate the functions of driven exercise in a U.S. adult sample that consisted of a substantial portion of athletes; recruitment for this sample also oversampled for individuals with self-reported ED pathology. We further investigated correlations of the functions with compulsive exercise and ED severity. Study 3 was not preregistered but followed similar procedures to preregistered Studies 1 and 2.

Transparency and Openness

Preregistration

The data-collection and analytic procedures of Studies 1 and 2 were preregistered at <https://osf.io/gjfc2>.¹

Data, materials, code, and online resources

Data and code to reproduce main findings of Study 1 and Study 2 are available on OSF at <https://osf.io/bfwej>. Ethics approval was not obtained to share data of Study 3. Analytical code for Study 3 is identical to Study 1 and Study 2. In the Supplemental Material available online, Table S1 reports how we handled missing data for all studies and results of the bias assessment, Table S2 provides a list of corresponding items for each factor, and Figures S1 through S3 report the correlation

patterns of the five-function models of Study 1, Study 2, and Study 3, respectively.

Reporting

Studies 1 and 2. We report all data exclusions, all manipulations, and all measures used in the study. Sample size was determined as reported in the preregistration. Additional measures were collected for the master thesis of T. Lask but not analyzed for this article.

Study 3. We report all data exclusions, all manipulations, and all measures used in this study. The sample is a partial sample of an ongoing project that was conducted until data collection in Study 2 was completed.

Ethical approval

For Study 1, the study design was approved by the ethics committee of the Institute of Psychology of the Johannes Gutenberg-University, Mainz, Germany. For Study 2, the study design was approved by the ethics committee of the medical faculty of the Ludwig-Maximilians-University Munich, Germany (No. 21-0043). For Study 3, the study was approved by the Internal Review Board of the University of Southern California (No. 20-00467).

Study 1

Method

Participants and procedures. Participants were recruited online using REDCap (P. A. Harris et al., 2009) hosted at University Medicine Mainz, Germany, between November 2020 and January 2021. Recruitment occurred via mailing lists of university students in Germany and via SurveyCircle (<https://www.surveycircle.com>; Mannheim, Germany). Psychology students were compensated with course credits. To increase the proportion of participants with self-reported EDs, we also recruited via online communities and social media profiles focusing on EDs and mental health. Data collection was anonymous, except for participants who requested course credits, in which case, identifiable information was stored separately from study data.

Interested users ($N = 1,176$) were able to give informed consent online, and 863 participants who fulfilled inclusion criteria (ages 16–50 and provided informed consent) completed the questionnaire battery. We then excluded participants who (a) failed both attention checks and (b) did not report at least 1 day on which they exercised in a driven manner (see driven-exercise screener [DES] below).

This resulted in a final sample of 279 participants (age: $M = 25.29$ years, $SD = 6.90$). Most participants were assigned female sex at birth ($n = 237$; 85%) and identified as female gender ($n = 233$; 84%). Eight participants (3%) identified as either transgender, intersexual, or nonbinary. Most participants ($n = 147$; 53%) were currently single, followed by participants who were married or in permanent relationships ($n = 127$; 46%). Ethnicity and sexual orientation were not recorded in this study because of German legislation limiting the collection of these data. All participants that had any amount of partial missing data in the Functional Assessment of Maladaptive Behaviors (FAMB; $N = 15$; 5%) were excluded from confirmatory factor analysis (CFA).²

Measures.

Body mass index. Body mass index (BMI; kg/m^2) was calculated using self-reported height and weight. Research indicates that self-reported height and weight are sufficiently accurate for our study purposes (Lin et al., 2012).

DES. The DES (Kolar & Gorrell, 2020b) is a two-item measure created by members of the study team to assess driven exercise within the last month. It provides participants with a definition of “driven” and three examples of potential driven-exercise behaviors (exercising excessively more than others in similar situations, exercising even when injured, and exercising when neglecting other important life domains). Participants then report whether they exercised in a driven manner within the last month and if yes, how often (range = 1–31 days). Participants were excluded (and did not fill in the functional-assessment questionnaire) if they reported no driven-exercise days within the last month.

FAMB. We translated the FAMB (Wedig & Nock, 2010) version as adapted by Wang et al. (2021) to German to assess functions of driven exercise. Translation was conducted by D. R. Kolar and T. Lask, tested in a focus group of patients with EDs, adjusted, and then back-translated to English by a native speaker. The back-translation was discussed with A. F. Haynos, S. B. Wang, and S. Gorrell to ensure semantic overlap. The FAMB was used to measure how often participants engage in driven exercise for intended automatic-positive-, automatic-negative-, social-positive-, and social-negative-reinforcement functions; options ranged from 0 (*never*) to 3 (*often*). The functions of driven exercise were not mutually exclusive, and individuals could endorse multiple different functions of this behavior. An adapted version was previously created to better capture potential functions of restrictive eating; we employed this adapted version in this investigation of driven exercise. In brief, two items regarding automatic positive reinforcement, one item regarding negative

reinforcement, and four items assessing an additional control function were added (for details, see Wang et al., 2021).

Eating Disorder Examination–Questionnaire. We used the German version (Hilbert et al., 2007) of the Eating Disorder Examination–Questionnaire (EDE-Q; Fairburn & Beglin, 1994) to assess disordered-eating symptoms within the last 4 weeks. Items were rated from 0 (*no day*) to 6 (*every day*). The EDE-Q is a reliable instrument with strong validity and test–retest reliability (Berg et al., 2012). We used the EDE-Q global score in this study as a measure of ED severity.

Compulsive Exercise Test. The Compulsive Exercise Test (CET; Taranis & Meyer, 2011) is a self-report measure assessing psychological components of driven exercise. It consists of 24 items rated from 0 (*never true*) to 5 (*always true*). It is widely used in ED research; however, the originally proposed factor structure with five subscale factors does not replicate well (for an overview, see A. Harris et al., 2020). We therefore used only the CET total score obtained with the German version by Schlegel et al. (2022) as a measure of driven-exercise severity.

Difficulties in Emotion Regulation Scale–18. The Difficulties in Emotion Regulation Scale–18 (DERS-18; Victor & Klonsky, 2016) is the short-form of the Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004). It comprises 18 items addressing different types of emotion-regulation problems, for example, nonacceptance of emotional reactions, impulsivity, or deficits in emotional clarity. Participants reported the items from 1 (*almost never*) to 5 (*almost every time*). Because of the lack of an independent German DERS-18 version, we used the corresponding 18 items of the German DERS (Gutzweiler & In-Albon, 2018) to compute the total DERS-18 score as a measure of emotion-regulation problems.

Depression, Anxiety Stress Scales. The Depression Anxiety Stress Scales (DASS-21; Antony et al., 1998) assesses depression, anxiety, and stress symptoms as a measure of general psychopathology. The German version (Nilges & Essau, 2015) showed good psychometric properties in a sample of chronic-pain patients.

Multidimensional Scale of Perceived Social Support. The Multidimensional Scale of Perceived Social Support (MSPSS; Zimet et al., 1988) assesses perceived social support from family, friends, and significant others, rated on a 7-point scale from 1 (*disagree strongly*) to 7 (*agree strongly*). The German-version MSPSS showed good psychometric properties in a sample of Austrian students (Greimel et al., 2016). We used the total MSPSS score as a measure of general perceived social support in our study.

Data analysis. Data were analyzed in R (R Core Team, 2022) and RStudio (Rstudio Team, 2022) using the *tidyverse* (Grolemund et al., 2019), *psych* (Revelle, 2022), *ggplot2* (Wickham, 2016), and *lavaan* (Rosseel, 2012) packages. We used CFA to test two models, including the four-function model as found in Wang et al. (2021) for restrictive eating and a five-function model with control-function items restricted to loading on a discrete control function instead of automatic positive reinforcement (for a list of specific items for each factor, see Table S2 in the Supplemental Material). We evaluated goodness of fit using standard fit indices following Hu and Bentler (1999), including comparative fit index (CFI; cutoff $\geq .95$), Tucker-Lewis Index (TLI; cutoff $\geq .95$), root mean square error of approximation (RMSEA; cutoff $\leq .06$), and standardized root mean square residual (SRMR cutoff $\leq .08$). Model fit of four-function and five-function models was compared using a scaled χ^2 difference test (Satorra, 2000). Differences in endorsement of functions were tested using linear mixed models because functions were nested within participants. The endorsement of reinforcement functions was defined by inspecting the estimated marginal means of these models. A reinforcement function was considered endorsed when it showed a mean score greater than zero. Tukey's honestly significant difference test was used to assess pairwise significant differences. Pearson correlations were calculated to evaluate associations with ED and general psychopathology, emotion-regulation difficulties, and perceived social support. Differences in correlations of control and automatic positive reinforcement with clinical features were explored with *cocor* (Diedenhofen & Musch, 2015) using Hotelling's *t* test (Hotelling, 1940).

Results

Many participants reported a current ($n = 86$; 31%) or past ED ($n = 31$; 11%), mostly anorexia nervosa spectrum (including full-threshold and atypical anorexia nervosa; $n = 78$; 67%), followed by bulimia nervosa spectrum (including full-threshold and atypical bulimia nervosa; $n = 32$; 27%), binge-eating disorder ($n = 3$; 3%), and avoidant-restrictive food intake disorder ($n = 2$; 2%). Average BMI was 21.86 kg/m^2 ($SD = 4.22 \text{ kg/m}^2$; range = 13.53 kg/m^2 – 50.50 kg/m^2). Participants endorsed driven exercise on average on 12.0 days ($SD = 9.2$ days) within the last month.

CFA. The CFA of the four-function model (i.e., automatic positive reinforcement and automatic negative reinforcement, social positive reinforcement and social negative reinforcement) yielded an acceptable model fit slightly outside of the cutoff recommendations by Hu and Bentler (1999) for the following fit indices: CFI = .90, TLI = .89,

RMSEA = .09 (90% confidence interval [CI] = [.09, .10]), SRMR = .12. Although a five-function model with control functions loading on a separate factor showed a statistically significantly better fit compared with the four-function model, $\Delta\chi^2(4) = 46.44$, $p < .001$, model fit improved only marginally for relevant fit indices: CFI = .91, TLI = .90, RMSEA = .09 (90% CI = [.08, .09]), SRMR = .12. Therefore, we present further results for both models.

Differences in endorsement of functions. Table 1 depicts endorsement levels of each function. Significance pairwise differences in endorsement were found between all functions, $F(3, 799) = 437.19$, $p < .001$ (all comparisons $p < .001$); highest endorsements were for automatic positive reinforcement, followed by automatic negative reinforcement, social negative reinforcement, and social positive reinforcement. A similar pattern emerged for the five-function model, $F(4, 1062) = 338.77$, $p < .001$, in which endorsement of all functions differed ($p < .001$) except for endorsement of automatic positive reinforcement with control ($p = .180$).

Associations with clinical features. All functions of the four-function model were intercorrelated (all $r_s > .27$, $p_s < .001$) and significantly correlated with the number of driven-exercise episodes reported in the past month, compulsive-exercise severity, ED severity, and general psychopathology (depression and stress; all $r_s > .13$, $p < .001$; see Fig. 1). Automatic positive reinforcement, automatic negative reinforcement, and social negative reinforcement were also correlated with more emotion-regulation difficulties (all $r_s > .29$, $p_s \leq .008$). Perceived social support was negatively correlated with automatic negative reinforcement and social negative reinforcement (all $r_s < -.16$, $p_s \leq .019$). Similar results emerged for the five-function model (all $r_s > .19$, $p \leq .005$; see Fig. S1 in the Supplemental Material), with significantly greater associations of clinical features with control than automatic positive reinforcement ($t \geq 4.70$, $p < .001$).

Discussion

We found evidence supporting a four-function model of driven exercise in a German online sample of adolescents and adults. Similar to restrictive eating (Wang et al., 2021) but in contrast to purging and NSSI (Nock & Prinstein, 2004; Wedig & Nock, 2010), automatic-positive-reinforcement functions were the most strongly endorsed functions in this sample. A five-function model that included an additional control function instead of having control functions load onto automatic positive reinforcement also showed an acceptable model fit. However, endorsement of automatic positive reinforcement and control functions did not differ

Table 1. Endorsement of Different Functions of Driven Exercise Across the Three Studies

Study 1: German online sample				Study 2: German clinical sample			Study 3: U.S. online sample		
Four-function model				Four-function model			Four-function model		
	<i>M</i>	<i>SE</i>	95% CI	<i>M</i>	<i>SE</i>	95% CI	<i>M</i>	<i>SE</i>	95% CI
ANR	1.54 ^a	0.04	[1.47, 1.62]	1.88 ^a	0.06	[1.75, 2.00]	1.85 ^a	0.09	[1.68, 2.03]
APR	1.83 ^a	0.04	[1.75, 1.90]	2.11 ^a	0.06	[1.98, 2.23]	2.25 ^a	0.09	[2.08, 2.43]
SNR	0.90 ^a	0.04	[0.82, 0.97]	1.36 ^a	0.06	[1.23, 1.48]	1.09 ^a	0.09	[0.91, 1.26]
SPR	0.42 ^a	0.04	[0.34, 0.50]	0.67 ^a	0.06	[0.54, 0.79]	0.47 ^a	0.09	[0.29, 0.64]
Five-function model				Five-function model			Five-function model		
	<i>M</i>	<i>SE</i>	95% CI	<i>M</i>	<i>SE</i>	95% CI	<i>M</i>	<i>SE</i>	95% CI
ANR	1.54 ^{a,b}	0.04	[1.46, 1.63]	1.88 ^a	0.07	[1.74, 2.01]	1.85 ^{a,b}	0.09	[1.66, 2.04]
APR	1.89 ^b	0.04	[1.80, 1.97]	1.94 ^b	0.07	[1.81, 2.08]	2.32 ^a	0.09	[2.13, 2.51]
CTR	1.78 ^a	0.04	[1.69, 1.86]	2.23 ^{a,b}	0.07	[2.10, 2.36]	2.20 ^b	0.09	[2.01, 2.39]
SNR	0.89 ^{a,b}	0.04	[0.81, 0.98]	1.36 ^{a,b}	0.07	[1.22, 1.49]	1.09 ^{a,b}	0.09	[0.90, 1.28]
SPR	0.42 ^{a,b}	0.04	[0.33, 0.50]	0.67 ^{a,b}	0.07	[0.53, 0.80]	0.47 ^{a,b}	0.09	[0.28, 0.66]

Note: Estimated marginal means of different reinforcement functions. Superscripts indicate model-wise Tukey-corrected significant pairwise differences between functions with $p < .05$. CI = confidence interval; ANR = automatic negative reinforcement; APR = automatic positive reinforcement; CTR = control functions; SNR = social negative reinforcement; SPR = social positive reinforcement.

significantly in this model and were both endorsed to a greater degree than other functions. All four or five functions correlated with higher levels of psychopathology, and social negative reinforcement correlated with lower perceived social support. Control functions were significantly more strongly associated with clinical features than automatic positive reinforcement. Although our sample was composed of participants who reported at least one driven-exercise episode and a large proportion of the sample reported a current ED episode, the findings may not extend to a clinical sample.

Study 2

To test whether a similar functional model of driven exercise could be obtained in participants with severe EDs, we collected data in a sample of adolescents and adults currently receiving intensive inpatient treatment for EDs.

Method

Participants and procedures. We recruited patients with EDs who were currently receiving specialized ED inpatient treatment at the Schoen Clinic Roseneck, Prien am Chiemsee, Germany, between April 2021 and October 2022. Patients received a multimodal treatment program based on cognitive-behavior-therapy principles, including individual and group psychotherapy and high-calorie oral refeeding. All patients were admitted by self-admission or referral by a clinician. Patients were not admitted in case of acute suicidality or whenever involuntary treatment was required (for further treatment details, see e.g., Kolar,

Meule, et al., 2022). Patients were discouraged from exercising during treatment except for supervised exercise therapy sessions. However, they were generally able to engage in unmonitored exercise during breaks between sessions that required changing buildings, during unsupervised time on weekends (e.g., going for a walk in the city, to the nearby lake), and especially whenever patients were alone in their private rooms (e.g., during nighttime, when roommates were attending therapy sessions). Recruitment occurred via direct contact through research assistants. Participation was administered via paper-pencil or online using REDCap.

Of 174 interested patients, 118 patients fulfilled inclusion criteria (ages 16–50 and provided informed consent) and did not fail both attention checks or report fewer than one driven-exercise episode within the last month (exclusion criteria). Patients' mean age was 22.68 years old ($SD = 7.26$); all were assigned female sex at birth, and most identified as female gender ($n = 113$; 96%). Two patients (2%) each identified as male or nonbinary, and one (1%) identified as intersex. Ethnicity and sexual orientation were not recorded in this study because of German legislation limiting collection of these data. All participants that had any amount of partial missing data in the FAMB ($n = 11$; 9%) were excluded from CFA.

Measures. The same measures as in Study 1 were used.

Data analysis. We followed the same data-analytic procedure as in Study 1. In addition, we tested measurement invariance using the R *semTools* (Jorgensen et al.,

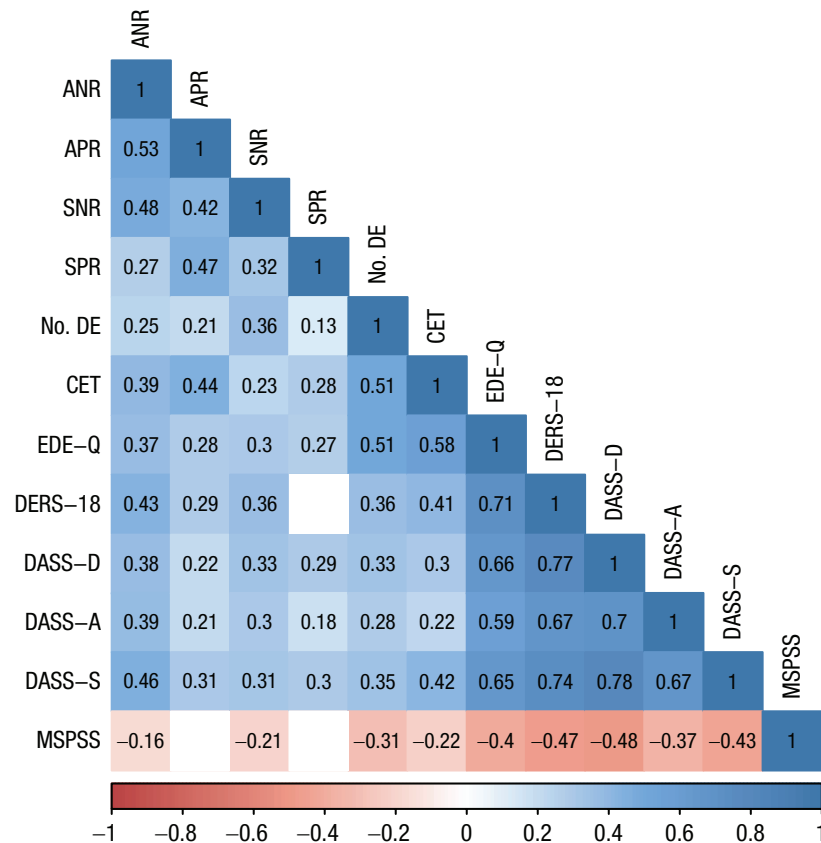


Fig. 1. Correlations among driven-exercise functions and clinical features in Study 1 ($N = 279$ adolescents and adults, oversampled for eating disorders). All presented correlations are statistically significant at $p < .05$. ANR = automatic negative reinforcement; APR = automatic positive reinforcement; SNR = social negative reinforcement; SPR = social positive reinforcement; No. DE = number of days with driven-exercise episodes in the last month; CET = Compulsive Exercise Test; EDE-Q = Eating Disorder Examination Questionnaire; DERS-18 = Difficulties in Emotion Regulation Scale-18, short form; DASS-D = Depression Anxiety Stress Scales depression subscale; DASS-A = Depression Anxiety Stress Scales, anxiety subscale; DASS-S = Depression Anxiety Stress Scales, stress subscale; MSPSS = Multidimensional Scale of Perceived Social Support.

2022) package to compare models between samples of Studies 1 and 2 to check whether obtained functional models differed between the online and the ED inpatient samples.

Results

Most patients self-reported a current diagnosis of anorexia nervosa ($n = 85$; 72%), subthreshold anorexia nervosa ($n = 21$; 18%), bulimia nervosa ($n = 8$; 7%), and subthreshold bulimia nervosa ($n = 1$; 1%), and three patients reported otherwise specified or unspecified EDs ($n = 3$; 3%). Most patients reported currently being underweight (BMI: $M = 18.37$ kg/m², $SD = 3.41$ kg/m²; range = 11.15 kg/m²–34.95 kg/m²; 58% BMI < 18.5 kg/m²). On average, participants reported exercising in a driven manner on 20.5 days ($SD = 10.0$) within

the last month, a time period that partially overlapped with inpatient treatment.

CFA. Again, we fitted a four- and a five-function model to the data. The four-function model fit was acceptable: CFI = .92, TLI = .91, RMSEA = .08 (90% CI = [.07, .09]), SRMR = .14. Similar to Study 1, the five-function model fitted statistically significantly better to the data, $\Delta\chi^2(4) = 14.10$, $p = .007$, but model-fit indices were only marginally better: CFI = .93, TLI = .92, RMSEA = .08 (90% CI = [.07, .09]), SRMR = .14. We again provide further results for both models.

Measurement invariance between online participants (Study 1) and ED inpatients (Study 2). Configural invariance (i.e., whether the same factorial structure held consistent across online participants and

inpatients with severe EDs) was obtained for the four- and five-function models (model fit for four-function model: CFI = .91, TLI = .91, RMSEA = .09, 90% CI [.08, .09], SRMR = .13; five-function model: CFI = .92, TLI = .91, RMSEA = .08, 90% CI [.08, .09], SRMR = .12). For the four-function model, metric invariance (i.e., similar factor loadings for both samples) was not achieved ($\Delta\chi^2 = 46.47$, $p = .009$), indicating that associations between functions may differ between online participants and inpatients with EDs, although differences in other model fit indices were very small ($\Delta \leq .01$). Metric invariance for the five-function model could also not be achieved because the model was not identified, most likely because of small sample size.

Differences in endorsement of functions. Table 1 shows endorsement levels of each function. Again, significant pairwise differences in endorsement were found between all functions for the four-function model, $F(3, 339) = 159.97$, $p < .001$; all comparisons $p \leq .008$. Highest endorsements were reported for automatic-positive-reinforcement functions of driven exercise, closely followed by automatic negative reinforcement and then social negative reinforcement and social positive reinforcement. Differences in function endorsements were also found for the five-function model, $F(4, 452) = 133.81$, $p < .001$. In contrast to Study 1, differences were detected in frequency of endorsement of control (highest endorsement) and automatic-positive-reinforcement (second-highest endorsement) functions of driven exercise, $t(452) = 3.77$, $p = .002$, whereas automatic positive reinforcement and automatic negative reinforcement (third-highest endorsements) did not differ significantly, $t(452) = 0.90$, $p = .899$. All other endorsement comparisons were statistically significant ($p < .001$).

Associations with clinical features. All functions of the four-function model were intercorrelated (all $r_s > .27$, $p_s \leq .002$; see Fig. 2) except for social-positive-reinforcement and social-negative-reinforcement functions, which were not significantly correlated ($r = .09$, $p = .136$). All functions were correlated with ED severity (all $r_s > .18$, $p_s \leq .027$). Automatic positive reinforcement, automatic negative reinforcement, and social negative reinforcement were correlated with the reported number of driven-exercise episodes in the last month (all $r_s > .25$, $p_s \leq .004$), but only automatic positive reinforcement and automatic negative reinforcement were correlated with compulsive exercise severity, that is, CET total scores (all $r_s > .58$, $p_s \leq .018$). Automatic-negative-reinforcement and social-negative-reinforcement functions were associated with emotion-regulation difficulties, anxiety, and depression (all $r_s > .27$, $p_s \leq .045$). Social positive reinforcement was correlated only with emotion-regulation

difficulties and stress (all $r_s > .20$, $p_s \leq .047$). Surprisingly, automatic positive reinforcement was not statistically correlated with any of these measures (all $r_s < .31$, $p_s \geq .062$). Perceived social support was not statistically correlated with any function (all $r_s < .08$, $p_s \geq .448$). For the five-function model, control function correlated with all other functions (all $r_s > .37$, $p < .001$; see Fig. S2 in the Supplemental Material) and all other clinical features (all $r_s > .31$, $p \leq .011$) except for emotion-regulation difficulties and perceived social support. Correlations with non-ED-specific clinical features were significantly different between control and automatic positive reinforcement ($t \geq 3.19$; $p \leq .002$; higher for control with all features except social support).

Discussion

In general, findings from Study 1 were replicated in a sample of inpatients with severe EDs. The factor structure of maintenance functions was similar between Study 1 and Study 2 samples; however, metric and scalar invariance were not established. This suggests that comparisons of endorsement levels of functions should not be made between inpatient and nonclinical samples. Note that despite us finding a better fit of the five-function model compared with the four-function model, control functions showed higher associations with general clinical features (e.g., depression, emotion regulation) but similar associations with ED-specific features (e.g., ED symptoms, driven-exercise severity) and an inverse association with social support compared with automatic positive reinforcement. Associations between automatic positive reinforcement and clinical features were not maintained when modeling control as a separate factor.

Study 3

In this study, we tested whether a similar model of reinforcement functions could be obtained in a sample of U.S. adults oversampled for athletes/EDs. Given the high prevalence of driven exercise in athletes (e.g., Lejoyeux et al., 2008; Lichtenstein & Jensen, 2016), it is important to investigate whether similar functions can be found in this population.

Method

Participants and procedures. Participants were recruited online between July 2021 and May 2022. Recruitment occurred via a one-time online survey as part of a larger project on characterizing exercise and eating behavior across varied adult samples. Prescreening items (listed below) were used to oversample for active athletes and

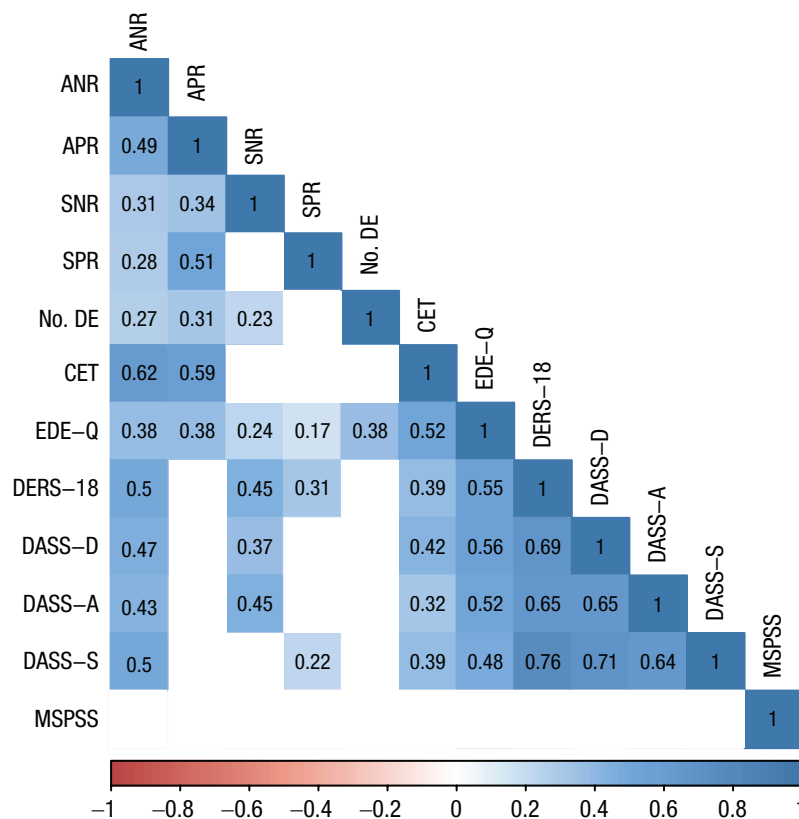


Fig. 2. Correlations among driven-exercise functions and clinical features in Study 2 ($N = 118$ adolescents and adults with eating disorders currently receiving inpatient treatment). All presented correlations are statistically significant at $p < .05$. ANR = automatic negative reinforcement; APR = automatic positive reinforcement; SNR = social negative reinforcement; SPR = social positive reinforcement; No. DE = number of days with driven-exercise episodes in the last month; CET = Compulsive Exercise Test; EDE-Q = Eating Disorder Examination Questionnaire; DERS-18 = Difficulties in Emotion Regulation Scale-18, short form; DASS-D = Depression Anxiety Stress Scales depression subscale; DASS-A = Depression Anxiety Stress Scales, anxiety subscale; DASS-S = Depression Anxiety Stress Scales, stress subscale; MSPSS = Multidimensional Scale of Perceived Social Support.

individuals with current or prior history of an ED. A total of 134 participants fulfilled inclusion criteria (18–64 years old and met criteria for at least one of the following categories: self-report of a current or prior diagnosis of an ED, currently identify as an athlete, are receiving treatment for an ED, or are a full-time college student—or fulfill any combination of these identifiers) and provided informed consent. Of those individuals, 47 participants were excluded because of missing data on the FAMB questionnaire. Another 35 participants were excluded because of not reporting at least one episode of driven exercise within the last 28 days via the EDE-Q.

The final sample comprised 52 participants (age: $M = 30.59$ years, $SD = 8.65$). Most participants identified as female gender ($n = 36$; 69%), one identified as male (2%), and three (6%) identified as nonbinary/other gender ($n = 11$; 21% did not report or preferred not to disclose their gender). Of the individuals reporting,

most participants identified race/ethnicity as White ($n = 39$; 75%), followed by Asian ($n = 3$; 6%) and Black ($n = 1$; 2%); 17% did not report race/ethnicity. Three participants (6%) reported more than one race/ethnicity. Two (4%) participant had any amount of partial missing data in the FAMB and was thus excluded from CFA.

Measures. We used the English versions of the following three measures that were already used in Study 1 and Study 2: FAMB (Wedig & Nock, 2010), EDE-Q (Fairburn & Beglin, 1994), and CET (Taranis & Meyer, 2011).

Data analysis. Again, we conducted CFAs for the four-function and five-function models and compared model fit. We also tested differences in endorsement of functions and calculated correlations of the different functions with ED and driven-exercise severity. Measurement

invariance between German and U.S. samples was not tested because measurement variance could not be attributed to sample or translation characteristics.

Results

Twenty-two participants (42%) self-reported currently having an ED diagnosed by a medical professional; anorexia nervosa ($n = 16$; 73%) was the most prominent ED, followed by bulimia nervosa ($n = 3$; 14%) and otherwise specified/unspecified EDs ($n = 3$; 14%). Half of participants, $n = 26$ (50%), considered themselves currently an athlete; 30% reported regular running activity, and 17% reported that their current athletic participation was competitive. On average, participants reported exercising in a driven manner on 12.6 days ($SD = 9.5$) within the last 28 days. Of the total analytic sample, 14 (16%) identified both as an athlete and as having a current ED diagnosis.

CFA. We fitted the same four- and five-function models. Both model fits were similar as in Studies 1 and 2; there was a slightly better fit for the five-function model, $\Delta\chi^2(4) = 19.72$, $p < .001$; four-function model: CFI = .95, TLI = .94, RMSEA = .09 (90% CI = [.07, .11]), SRMR = .18; five-function model: CFI = .95, TLI = .95, RMSEA = .09 (90% CI = [.07, .11]), SRMR = .17). Again, we provide further results for both models.

Differences in endorsement of functions. Endorsement levels for each function are reported in Table 1. Endorsements of all functions of the four-function model differed significantly, $F(3, 153) = 152.67$, $p < .001$; all comparisons' $ps < .001$. Highest endorsements were reported for automatic positive reinforcement, followed by automatic negative reinforcement, social negative reinforcement, and social positive reinforcement. Differences between endorsements of functions were also found for the five-function model, $F(4, 204) = 113.40$, $p < .001$. All differences were statistically significant with $p \leq .011$ except for the difference in endorsements of automatic-positive-reinforcement and control functions ($p = .763$). This contrasts with Study 1 (German online sample: automatic positive reinforcement significantly higher endorsement) and Study 2 (German inpatient sample: control functions significantly higher endorsement).

Associations with clinical features. Again, all functions of the four-function model were intercorrelated (all $rs \geq .36$, $ps \leq .009$; see Fig. 3). Only social negative reinforcement correlated with ED severity ($r = .28$, $p = .044$), and automatic positive reinforcement correlated with frequency of driven-exercise episodes in the past month ($r = .29$, $p = .039$). None of the functions were associated with

compulsive exercise severity (all $rs \leq |.12|$, $ps \geq .416$). A similar correlation pattern emerged for the five-function model, with statistically significant intercorrelations ($rs \geq .30$, $ps \leq .009$; see Fig. S3 in the Supplemental Material), except for automatic positive reinforcement with automatic negative reinforcement ($r = .18$, $p = .198$). Associations with clinical features emerged only for control with frequency of driven-exercise episodes ($r = .33$, $p = .017$) and social negative reinforcement with ED severity ($r = .28$, $p \leq .044$; all other correlations $rs \leq .27$, $p \geq .058$). Associations with clinical features were not significantly distinct between control functions and automatic positive reinforcement ($t \leq 1.33$; $p \geq .190$).

Discussion

Both four-function and five-function structures could be replicated in a U.S. adult sample with a large proportion of athletes and individuals with self-reported EDs; however, the models performed worse on some model-fit statistics because of the smaller sample size. When looking at endorsements of different functions, automatic-positive-reinforcement functions again emerged as the most highly endorsed motivations for driven exercise. In contrast to Study 1 and Study 2, no difference between endorsements of automatic-positive-reinforcement and control functions was observed. However, control functions were associated with frequency of driven-exercise episodes in the previous month, indicating a potential clinical significance of these functions.

General Discussion

We investigated driven-exercise functions in three studies with samples including adolescents and adults and individuals with different ED presentations, ED-severity levels, and athletic engagement. In general, the four-function structure previously found for NSSI, binge eating, purging, and restrictive eating was also obtained for driven exercise in all samples. Similar to restrictive eating (Wang et al., 2021), automatic positive reinforcement was the function most highly endorsed across all samples, followed by automatic-negative-, social-negative-, and social-positive-reinforcement functions. In contrast to previous studies, a five-function model including control as a separate maintenance function for driven exercise yielded a slightly better but only marginally different model fit compared with the four-function model. Because model comparison tests based on χ^2 differences are rather sensitive and the more parsimonious four-function model was also found in other maladaptive behaviors, we suggest the four-function model as the preferred model for automatic and social

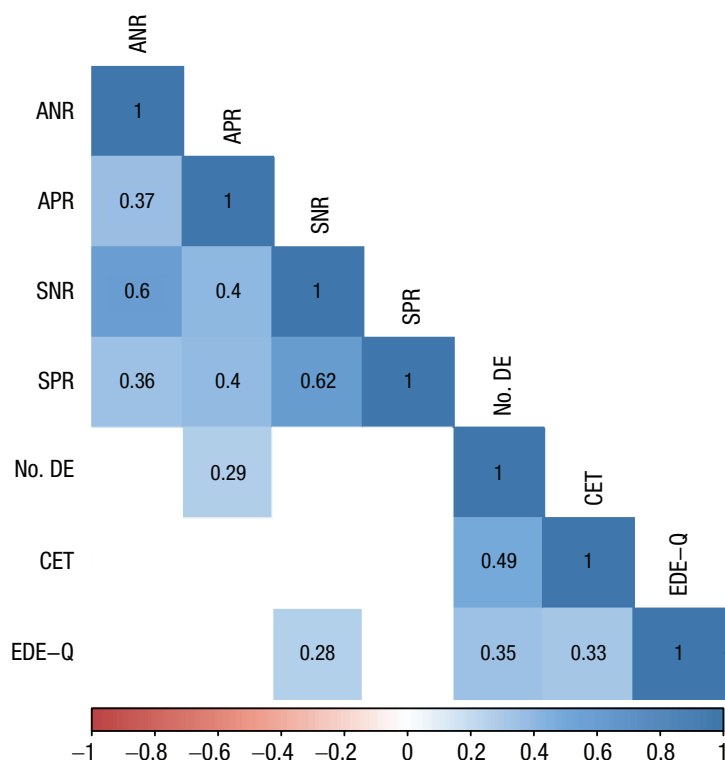


Fig. 3. Correlations among driven exercise functions and clinical features in Study 3 ($N = 52$ adults, oversampled for athletes). All presented correlations are statistically significant at $p < .05$. ANR = automatic negative reinforcement; APR = automatic positive reinforcement; SNR = social negative reinforcement; SPR = social positive reinforcement; No. DE = number of days with driven exercise episodes in the last month; CET = Compulsive Exercise Test; EDE-Q = Eating Disorder Examination Questionnaire.

functions of driven exercise. This is further underlined by a recent simulation study indicating that relying on model-fit indices for model selection results in overfactored solutions (Montoya & Edwards, 2021), which is why we do not recommend using model-fit indices to decide on the number of factors. We therefore generally recommend using a four-function model when investigating and comparing reinforcement functions of maladaptive behaviors except when control functions are of specific research interest. Note that none of the models were below the specified threshold for SRMR model-fit statistics (which may, however, perform especially poorly with categorical data; see Montoya & Edwards, 2021), although other model-fit statistics indicated sufficient fit. Based on these data, there was some evidence that control functions were important, especially in the inpatient ED sample, because they were the most endorsed when modeled as an additional factor. Higher endorsement of all functions of driven exercise, especially the automatic-reinforcement functions, was associated with indicators of driven-exercise severity (i.e., global score of a dimensional driven-exercise measure, frequency of driven-exercise episodes in the previous month) across samples. Associations of functions,

especially negative-reinforcement functions, with general clinical features (e.g., emotion regulation, depression, or perceived social support) were found consistently across online participants and inpatients with EDs. Control functions were more strongly associated with emotion regulation, depression, anxiety, and stress than automatic positive reinforcement, further corroborating the clinical importance of control functions in driven exercise.

Identifying maintenance functions may inform both conceptualization and treatment of driven exercise. Previously described conceptualizations of driven exercise as exercise addiction or compulsive exercise were highly specific and did not replicate well across different contexts and stages of illness severity (Scharmer et al., 2020; Schlegel et al., 2022). However, in this study, very similar maintenance functions for driven exercise emerged as those that have been identified for binge eating, purging, restrictive eating, and other maladaptive behaviors, such as NSSI. This might indicate that automatic and social reinforcement are relevant transdiagnostic maintenance functions across problematic behaviors. In ED treatment, driven exercise has predominantly been considered to serve simply for

compensation of caloric intake. Our findings provide an empirical basis for potential additional treatment targets, including replacing driven exercise by finding alternative, healthy behaviors that increase positive affect or teaching people to engage in exercise related to long-term values (e.g., team sports), as opposed to short-term affect regulation. Furthermore, the considerable overlap in highly endorsed functions across different maladaptive behaviors may inform developing transdiagnostic interventions. Reiterating Wang et al. (2021), intervening on shared functions instead of specific behaviors may help in reducing the frequency of several maladaptive behaviors simultaneously.

The identified endorsement distribution of maintenance functions of driven exercise more closely resembled the pattern observed for restrictive eating than for other maladaptive behaviors, with automatic positive reinforcement being highest endorsed. Given that most of the participants with EDs in our studies were diagnosed with anorexia nervosa, this is in line with recent research highlighting the role of dysregulated positive affect and its treatment for this ED (Coniglio et al., 2019; Haynos et al., 2021). In addition, previous EMA studies have also highlighted the importance of increased positive affect after physical activity or exercise as a maintenance factor for driven exercise (Kolar et al., 2020; Reichert et al., 2020). Furthermore, preliminary experimental research shows that women who report driven exercise experience exercise as more rewarding than women who do not engage in driven exercise (Holland, 2016), indicating potential alterations in reward systems. In contrast to other forms of maladaptive behaviors, engaging in balanced and nonexcessive exercise is inherently adaptive, and several of the functions that maintain driven exercise, specifically, automatic-reinforcement functions, are even promoted in exercise-based treatments for other mental disorders, including depression or anxiety (Strohle, 2009). Thus, in contrast to treatment for restrictive eating, a function-oriented treatment of driven exercise may focus on promoting and implementing adaptive forms of exercise that serve these functions without the long-term negative consequences. For example, engaging in team sports could simultaneously serve automatic and social-positive functions, whereas social norms within and around team sports (sporting rules, clear time settings, socializing after training/games) might prevent individuals from “going too far” and turning to engagement in driven exercise. There is already preliminary evidence from a randomized controlled trial showing that implementing healthy exercise behaviors in treatment of EDs reduces the severity of driven exercise after treatment (Dittmer et al., 2020).

Regarding control functions, our data suggest that considering control as a separate function of driven exercise may be most useful in clinical populations given that the more parsimonious four-function model showed only marginally worse model fit and emerged for other maladaptive behaviors. In the four-function model, control functions consistently loaded onto automatic positive reinforcement, in accordance with the Wang et al. (2021) study on functions of restrictive eating. Similar to other studies (Nock & Prinstein, 2004; Wang et al., 2021; Wedig & Nock, 2010), social functions were less endorsed than automatic functions, especially social positive reinforcement. This suggests that despite the fact that some people engage in driven exercise for social reasons, the widespread belief that people engage in various problematic behaviors “to get attention” is probably a misconception.

Several strengths of our research can be noted. We investigated maintenance functions of driven exercise in three separate studies including adolescents and adults, individuals with severe EDs, individuals residing in the United States and Germany, and a high proportion of athletes. However, because of strict data-protection laws in Germany and the design of Study 3, which was conducted in the United States, several demographic data have not been recorded in all studies, including ethnicity, sexual orientation, and socioeconomic status. Thus, generalization to minority groups might not be applicable. In addition, only participants in Studies 1 and 2 were selected on reporting driven exercise using an extensive definition of driven exercise, whereas in Study 3, an item of the EDE-Q was used. Identifying driven exercise in athletes may be generally more difficult because of their unique relationship with exercising. In combination, inclusion could have been too broad in Study 3. However, participants were given a short definition of driven exercise directly before filling in the FAMB to support athletes’ understanding of how driven exercise differs from other forms of exercising. Our analyses for the first two studies were preregistered, and analyses of the third study closely followed the analytic plan of the former studies. Nonetheless, our investigations relied on cross-sectional data of self-report of the functions of driven exercise, and the assessment period of driven-exercise episodes might have had an overlap with inpatient treatment for participants in Study 2. Future research may benefit from investigating maintenance functions using intensive longitudinal designs to capture developmental and dynamic changes in these functions. In addition, we asked for intended functions, or the purpose of the maladaptive behavior, but not whether individuals effectively achieved a desired outcome via the intended function

(e.g., exercising in a driven manner to feel proud about something might not result in feeling prouder at the end of a workout). Thus, further research is needed to investigate contingencies of maintenance functions of driven exercise in experimental or EMA studies. For example, understanding whether driven exercise is maintained through intermittent or consistent schedules of contingencies and whether these schedules vary across illness stages may provide insight to prevent driven exercise among individuals who are most at risk.

Functional assessment of driven exercise can help in increasing understanding of how this maladaptive behavior is maintained over time. Given the serious long-term consequences of this behavior, developing more precise models of driven exercise across illness stages and diagnoses can ultimately inform personalized treatments that target individual functions of driven exercise. Such adaptive treatments can aid in integrating exercise behaviors that are healthy and adaptive, as opposed to problematic and harmful, to improve physical- and mental-health outcomes of individuals with EDs.

Transparency

Action Editor: Kelsie T. Forbush

Editor: Jennifer L. Tackett

Author Contributions

David R. Kolar: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Resources; Writing – original draft; Writing – review & editing.

Ann F. Haynos: Conceptualization; Methodology; Resources; Writing – review & editing.

Shirley B. Wang: Conceptualization; Methodology; Resources; Writing – review & editing.

Theresa Lask: Conceptualization; Data curation; Investigation; Writing – review & editing.

Stuart B. Murray: Data curation; Investigation; Resources; Writing – review & editing.

Ulrich Voderholzer: Investigation; Resources; Writing – review & editing.

Sasha Gorrell: Conceptualization; Data curation; Investigation; Methodology; Project administration; Resources; Writing – review & editing.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

S. B. Wang is supported by the National Science Foundation Graduate Research Fellowship under Grant DGE-1745303 and the National Institute of Mental Health under Grant F31MH125495. S. Gorrell and S. B. Murray are supported by the National Institute of Mental Health under Grants K23MH126201 (S. Gorrell), R21MH131787 (S. Gorrell), and K23MH115184 (S. B. Murray). S. Gorrell is also

supported by the Brain & Behavior Research Foundation (NARSAD Young Investigator Award).

Open Practices

This article has received the badges for Open Data and Preregistration. More information about the Open Practices badges can be found at <http://www.psychologicalscience.org/publications/badges>



ORCID iDs

David R. Kolar  <https://orcid.org/0000-0002-8649-5467>

Shirley B. Wang  <https://orcid.org/0000-0002-8583-3014>

Acknowledgments

We are thankful to Eva Zisler and Nina Kölbl for data entry. Preliminary results have been presented as posters at the XXVIIIth Annual Meeting of the Eating Disorders Research Society in Philadelphia and the 1st German Psychotherapy Congress in Berlin. The master's thesis of T. Lask was based on partial data from Study 1. We are grateful to the participants and patients who took part in this study and the staff at the Schoen Clinic Roseneck who helped during data collection. A preprint version of this article is available at <https://psyarxiv.com/ad3mq/>. Supplemental material is available at <https://osf.io/cymzn>.

Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/suppl/10.1177/21677026241290455>

Notes

1. Small changes have been made to preregistered hypotheses. We analyzed correlations with general emotion-regulation problems instead of emotional suppression. Originally, we planned to use the Emotion Regulation Questionnaire (which has an emotion suppression subscale) following Wang et al. (2021) but mistakenly used the Difficulties in Emotion Regulation Scale (which does not assess emotion suppression). Differences in endorsement of functions were not compared exploratorily between participants that self-report an eating disorder and/or report a global Eating Disorder Examination Questionnaire score above the cutoff of 2.3. At the time of preregistration, we registered to compare groups within Study 1 as a redundancy in case of not being able to recruit a clinical sample for Study 2 because of, again, strict COVID regulations in late 2020 through early 2021. Given the extensive comparison between Study 1 and Study 2 in this article, we considered comparing groups based on self-reported diagnoses and questionnaire cut-off redundant. Interested readers can conduct this analysis with the data provided at <https://osf.io/bfwej>. The preregistration also registers analyzing (a) functions of restrictive eating and (b) how functions and disordered-eating behaviors relate to gender roles, which has not been conducted for this study. We

considered analyzing functions of restrictive eating too extensive for this article (because it focuses on driven exercise), and how functions and disordered-eating behaviors relate to gender roles was registered and (partially) analyzed for T. Lask's master's thesis in Study 1 only.

2. Missing data were handled slightly differently in all studies. We considered implementing multiple imputation but ultimately decided not to do so for all studies because there is currently no method for pooling model-fit estimates for ordinal data estimated with diagonally weighted least squares. However, we assessed potential bias in estimates that were due to pairwise deletion, which was considered low. Supplementary Table S1 reports how we handled missing data for all studies and results of the bias assessment.

References

- Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment, 10*(2), 176–181.
- Berg, K. C., Peterson, C. B., Frazier, P., & Crow, S. J. (2012). Psychometric evaluation of the eating disorder examination and eating disorder examination-questionnaire: A systematic review of the literature. *International Journal of Eating Disorders, 45*(3), 428–438. <https://doi.org/10.1002/eat.20931>
- Coniglio, K. A., Christensen, K. A., Haynos, A. F., Rienecke, R. D., & Selby, E. A. (2019). The posited effect of positive affect in anorexia nervosa: Advocating for a forgotten piece of a puzzling disease. *International Journal of Eating Disorders, 52*(9), 971–976. <https://doi.org/10.1002/eat.23147>
- Dalle Grave, R., Calugi, S., & Marchesini, G. (2008). Compulsive exercise to control shape or weight in eating disorders: Prevalence, associated features, and treatment outcome. *Comprehensive Psychiatry, 49*(4), 346–352. <https://doi.org/10.1016/j.comppsy.2007.12.007>
- Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLOS ONE, 10*(4), Article e0121945. <https://doi.org/10.1371/journal.pone.0121945>
- Dittmer, N., Voderholzer, U., Monch, C., Cuntz, U., Jacobi, C., & Schlegel, S. (2020). Efficacy of a specialized group intervention for compulsive exercise in inpatients with anorexia nervosa: A randomized controlled trial. *Psychotherapy and Psychosomatics, 89*(3), 161–173. <https://doi.org/10.1159/000504583>
- Edlund, K., Johansson, F., Lindroth, R., Bergman, L., Sundberg, T., & Skillgate, E. (2022). Body image and compulsive exercise: Are there associations with depression among university students? *Eating and Weight Disorders—Studies on Anorexia, Bulimia and Obesity, 27*(7), 2397–2405. <https://doi.org/10.1007/s40519-022-01374-x>
- Engel, S. G., Wonderlich, S. A., Crosby, R. D., Mitchell, J. E., Crow, S., Peterson, C. B., Le Grange, D., Simonich, H. K., Cao, L., Lavender, J. M., & Gordon, K. H. (2013). The role of affect in the maintenance of anorexia nervosa: Evidence from a naturalistic assessment of momentary behaviors and emotion. *Journal of Abnormal Psychology, 122*(3), 709–719. <https://doi.org/10.1037/a0034010>
- Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders, 16*(4), 363–370. [https://doi.org/10.1002/1098-108X\(199412\)16:4<363::AID-EAT2260160405>3.0.CO;2-#](https://doi.org/10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-#)
- Ganson, K. T., Mitchison, D., Rodgers, R. F., Cunningham, M. L., Murray, S. B., & Nagata, J. M. (2022). Compulsive exercise among college students: 5-year time trends in prevalence and demographic, substance use, and mental health correlates. *Eating and Weight Disorders—Studies on Anorexia, Bulimia and Obesity, 27*(2), 717–728. <https://doi.org/10.1007/s40519-021-01210-8>
- Gjestvang, C., Bratland-Sanda, S., & Mathisen, T. F. (2021). Compulsive exercise and mental health challenges in fitness instructors: presence and interactions. *Journal of Eating Disorders, 9*(1), Article 107. <https://doi.org/10.1186/s40337-021-00446-0>
- Goodwin, H., Haycraft, E., & Meyer, C. (2012). The relationship between compulsive exercise and emotion regulation in adolescents. *British Journal of Health Psychology, 17*(4), 699–710. <https://doi.org/10.1111/j.2044-8287.2012.02066.x>
- Goodwin, H., Haycraft, E., & Meyer, C. (2014). Emotion regulation styles as longitudinal predictors of compulsive exercise: A twelve month prospective study. *Journal of Adolescence, 37*(8), 1399–1404. <https://doi.org/10.1016/j.adolescence.2014.10.001>
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment, 26*(1), 41–54. <https://doi.org/10.1023/B:JOBA.0000007455.08539.94>
- Greimel, E., Kato, Y., Müller-Gartner, M., Salchinger, B., Roth, R., & Freidl, W. (2016). Internal and external resources as determinants of health and quality of life. *PLOS ONE, 11*(5), Article e0153232. <https://doi.org/10.1371/journal.pone.0153232>
- Grolemund, G., Hayes, A., Henry, L., & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software, 4*(43), Article 1686. <https://doi.org/10.21105/joss.01686>
- Gutzweiler, R., & In-Albon, T. (2018). Überprüfung der Gütekriterien der deutschen Version der Difficulties in Emotion Regulation Scale in einer klinischen und einer Schülerstichprobe Jugendlicher [Validity and Reliability of the German Version of the DERS in a Clinical and a School Sample of Adolescents]. *Zeitschrift für klinische Psychologie und Psychotherapie, 47*(4), 274–286. <https://doi.org/10.1026/1616-3443/a000506>
- Harris, A., Hay, P., & Touyz, S. (2020). Psychometric properties of instruments assessing exercise in patients with eating disorders: A systematic review. *Journal of Eating Disorders, 8*, Article 45. <https://doi.org/10.1186/s40337-020-00315-2>
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap): A metadata-driven methodology and workflow

- process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42(2), 377–381. <https://doi.org/10.1016/j.jbi.2008.08.010>
- Hausenblas, H. A., & Downs, D. S. (2002). How much is too much? The development and validation of the exercise dependence scale. *Psychology & Health*, 17(4), 387–404. <https://doi.org/10.1080/0887044022000004894>
- Haynos, A. F., Anderson, L. M., Askew, A. J., Craske, M. G., & Peterson, C. B. (2021). Adapting a neuroscience-informed intervention to alter reward mechanisms of anorexia nervosa: A novel direction for future research. *Journal of Eating Disorders*, 9(1), Article 63. <https://doi.org/10.1186/s40337-021-00417-5>
- Hepp, J., Carpenter, R. W., Störkel, L. M., Schmitz, S. E., Schmah, C., & Niedtfeld, I. (2020). A systematic review of daily life studies on non-suicidal self-injury based on the four-function model. *Clinical Psychology Review*, 82, Article 101888. <https://doi.org/10.1016/j.cpr.2020.101888>
- Hilbert, A., Tuschen-Caffier, B., Karwautz, A., Niederhofer, H., & Munsch, S. (2007). Eating Disorder Examination-Questionnaire. *Diagnostica*, 53(3), 144–154. <https://doi.org/10.1026/0012-1924.53.3.144>
- Holland, L. A. (2016). *Mechanisms underlying pathological exercise in women with eating disorders*. Florida State University. <https://search.proquest.com/docview/1842421581>
- Hotelling, H. (1940). The selection of variates for use in prediction with some comments on the general problem of nuisance parameters. *The Annals of Mathematical Statistics*, 11(3), 271–283.
- Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). *semTools: Useful tools for structural equation modeling* (Version R Package Version 0.5-6). <https://CRAN.R-project.org/package=semTools>
- Kolar, D. R., & Gorrell, S. (2020a). A call to experimentally study acute affect-regulation mechanisms specific to driven exercise in eating disorders. *International Journal of Eating Disorders*, 54(3), 280–286. <https://doi.org/10.1002/eat.23427>
- Kolar, D. R., & Gorrell, S. (2020b). *Driven exercise screener*. OSF. <https://osf.io/hmt7w/>
- Kolar, D. R., Kaurin, A., Meule, A., Schlegl, S., Dittmer, N., & Voderholzer, U. (2022). Interpersonal, affective and compulsive features of driven exercise in anorexia nervosa. *Journal of Affective Disorders*, 307, 53–61. <https://doi.org/10.1016/j.jad.2022.03.044>
- Kolar, D. R., Meule, A., Naab, S., & Voderholzer, U. (2022). Early within-person weight gain and variability during inpatient treatment for anorexia nervosa: Age-dependent effects on treatment outcome. *European Eating Disorders Review*, 30(4), 328–340. <https://doi.org/10.1002/erv.2893>
- Kolar, D. R., Neumayr, C., Roth, M., Voderholzer, U., Perthes, K., & Schlegl, S. (2020). Testing an emotion regulation model of physical activity in adolescents with anorexia nervosa: A pilot ecological momentary assessment. *European Eating Disorders Review*, 28(2), 170–183. <https://doi.org/10.1002/erv.2706>
- Lampe, E. W., Gorrell, S., Smith, K., Payne-Reichert, A. M., Juarascio, A. S., & Manasse, S. M. (2023). Divergent trajectories of positive affect following maladaptive and non-maladaptive exercise among individuals with binge-spectrum eating disorders. *International Journal of Eating Disorders*, 56(5), 1001–1010. <https://doi.org/10.1002/eat.23901>
- Lejoyeux, M., Avril, M., Richoux, C., Embouazza, H., & Nivoli, F. (2008). Prevalence of exercise dependence and other behavioral addictions among clients of a Parisian fitness room. *Comprehensive Psychiatry*, 49(4), 353–358. <https://doi.org/10.1016/j.comppsy.2007.12.005>
- Lekkas, D., Gyorda, J. A., & Jacobson, N. C. (2022). A machine learning investigation into the temporal dynamics of physical activity-mediated emotional regulation in adolescents with anorexia nervosa and healthy controls. *European Eating Disorders Review*, 31, 147–165. <https://doi.org/10.1002/erv.2949>
- Lichtenstein, M. B., & Jensen, T. T. (2016). Exercise addiction in CrossFit: Prevalence and psychometric properties of the Exercise Addiction Inventory. *Addictive Behaviors Reports*, 3, 33–37. <https://doi.org/10.1016/j.abrep.2016.02.002>
- Lin, C. J., DeRoo, L. A., Jacobs, S. R., & Sandler, D. P. (2012). Accuracy and reliability of self-reported weight and height in the Sister Study. *Public Health Nutrition*, 15(6), 989–999. <https://doi.org/10.1017/s1368980011003193>
- Meyer, C., Taranis, L., Goodwin, H., & Haycraft, E. (2011). Compulsive exercise and eating disorders. *European Eating Disorders Review*, 19(3), 174–189. <https://doi.org/10.1002/erv.1122>
- Mikschl, S., Schlegl, S., Külz, A. K., Rief, W., & Voderholzer, U. (2016). Funktionalitäten von Essstörungen - Ergebnisse einer Analyse im stationären Setting [Functionalities of eating disorders - Results of an analysis in an inpatient setting]. *Verhaltenstherapie*, 26(3), 159–169. <https://doi.org/10.1159/000448496>
- Montoya, A. K., & Edwards, M. C. (2021). The poor fit of model fit for selecting number of factors in exploratory factor analysis for scale evaluation. *Educational and Psychological Measurement*, 81(3), 413–440.
- Nilges, P., & Essau, C. (2015). Die depressions-angst-stress-skalen. *Der Schmerz*, 29(6), 649–657.
- Nock, M. K., & Prinstein, M. J. (2004). A functional approach to the assessment of self-mutilative behavior. *Journal of Consulting and Clinical Psychology*, 72(5), 885–890.
- Noetel, M., Dawson, L., Hay, P., & Touyz, S. (2017). The assessment and treatment of unhealthy exercise in adolescents with anorexia nervosa: A Delphi study to synthesize clinical knowledge. *International Journal of Eating Disorders*, 50(4), 378–388. <https://doi.org/10.1002/eat.22657>
- R Core Team. (2022). *R: A language and environment for statistical computing* (Version 4.2.0). R Foundation for Statistical Computing.
- Reichert, M., Schlegel, S., Jagau, F., Timm, I., Wieland, L., Ebner-Priemer, U. W., Hartmann, A., & Zeeck, A. (2020).

- Mood and dysfunctional cognitions constitute within-subject antecedents and consequences of exercise in eating disorders. *Psychotherapy and Psychosomatics*, 89(2), 119–121. <https://doi.org/10.1159/000504061>
- Revelle, W. (2022). *psych: Procedures for personality and psychological research* (Version 2.2.5). Northwestern University. <https://CRAN.R-project.org/package=psych>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- RStudio Team. (2022). *RStudio: Integrated development for R* (Version 2022.07.1). RStudio Inc. <https://rstudio.com/>
- Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in multivariate statistical analysis. A Festschrift for Heinz Neudecker* (pp. 233–247). Kluwer Academic Publishers.
- Scharmer, C., Gorrell, S., Schaumberg, K., & Anderson, D. (2020). Compulsive exercise or exercise dependence? Clarifying conceptualizations of exercise in the context of eating disorder pathology. *Journal of Clinical Sport Psychology*, 46, Article 101586. <https://doi.org/10.1016/j.psychsport.2019.101586>
- Schlegl, S., Vierl, L., Kolar, D. R., Dittmer, N., & Voderholzer, U. (2022). Psychometric properties of the Compulsive Exercise Test in a large sample of female adolescent and adult inpatients with anorexia nervosa and bulimia nervosa. *International Journal of Eating Disorders*, 55(4), 494–504. <https://doi.org/10.1002/eat.23694>
- Slade, P. (1982). Towards a functional analysis of anorexia nervosa and bulimia nervosa. *British Journal of Clinical Psychology*, 21(3), 167–179. <https://doi.org/10.1111/j.2044-8260.1982.tb00549.x>
- Solenberger, S. E. (2001). Exercise and eating disorders: A 3-year inpatient hospital record analysis. *Eating Behaviors*, 2(2), 151–168. [https://doi.org/10.1016/s1471-0153\(01\)00026-5](https://doi.org/10.1016/s1471-0153(01)00026-5)
- Stiles-Shields, C., Bamford, B., Lock, J., & Le Grange, D. (2015). The effect of driven exercise on treatment outcomes for adolescents with anorexia and bulimia nervosa. *International Journal of Eating Disorders*, 48(4), 392–396. <https://doi.org/10.1002/eat.22281>
- Strohle, A. (2009). Physical activity, exercise, depression and anxiety disorders. *Journal of Neural Transmission*, 116(6), 777–784. <https://doi.org/10.1007/s00702-008-0092-x>
- Taranis, L., & Meyer, C. (2011). Associations between specific components of compulsive exercise and eating-disordered cognitions and behaviors among young women. *International Journal of Eating Disorders*, 44(5), 452–458. <https://doi.org/10.1002/eat.20838>
- Victor, S. E., & Klonsky, E. D. (2016). Validation of a Brief Version of the Difficulties in Emotion Regulation Scale (DERS-18) in five samples. *Journal of Psychopathology and Behavioral Assessment*, 38(4), 582–589. <https://doi.org/10.1007/s10862-016-9547-9>
- Wang, S. B., Fox, K. R., Boccagno, C., Hooley, J. M., Mair, P., Nock, M. K., & Haynos, A. F. (2021). Functional assessment of restrictive eating: A three-study clinically heterogeneous and transdiagnostic investigation. *Journal of Abnormal Psychology*, 130(7), 761–774. <https://doi.org/10.1037/abn0000700>
- Wedig, M. M., & Nock, M. K. (2010). The functional assessment of maladaptive behaviors: A preliminary evaluation of binge eating and purging among women. *Psychiatry Research*, 178(3), 518–524. <https://doi.org/10.1016/j.psychres.2009.05.010>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (2nd ed.). Springer.
- Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment*, 52(1), 30–41.