Dietel, Fanny Alexandra; Zache, Carina; Burkner, Paul-Christian; Schulte, Johanna; Moebius, Martin; Bischof, Annika; Wilhelm, Sabine; Buhlmann, Ulrike

Internet-based interpretation bias modification for body dissatisfaction

Published in:
International Journal of Eating Disorders

DOI:
10.1002/eat.23280

Published: 01/06/2020

Document Version
Publisher's PDF, also known as Version of record

Received: 30 December 2019 | Revised: 7 April 2020 | Accepted: 14 April 2020

DOI: 10.1002/eat.23280

ORIGINAL ARTICLE

International Journal of EATING DISORDERS WILEY

Internet-based interpretation bias modification for body dissatisfaction: A three-armed randomized controlled trial

Fanny Alexandra Dietel MSc[1] | Carina Zache MSc[1] | Paul-Christian Bürkner PhD[2] | Johanna Schulte MSc[1] | Martin Möbius PhD[3] | Annika Bischof MSc[1] | Sabine Wilhelm PhD[4] | Ulrike Buhlmann PhD[1]

[1]Clinical Psychology and Psychotherapy, Department of Psychology, Westfälische Wilhelms-Universität, Muenster, Germany

[2]Department of Computer Science, Aalto University, Espoo, Finland

[3]Behavioral Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands

[4]Massachusetts General Hospital, Harvard Medical School, Boston

Correspondence
Ulrike Buhlmann, Department of Psychology, Clinical Psychology and Psychotherapy, Westfälische Wilhelms-Universität Münster, Fliednerstrasse 21, 48149 Münster, Germany.
Email: ulrike.buhlmann@uni-muenster.de

Action Editor: Tracey Wade

Abstract

Objective: Appearance-related interpretation bias is postulated to play a role in the maintenance of body dissatisfaction (BD), a risk factor for body dysmorphic disorder (BDD), and eating disorders (ED). Cognitive bias modification for interpretation (CBM-I) has been shown to reduce maladaptive interpretation bias and symptoms in various emotional disorders. This study investigated the acceptability and efficacy of an easily disseminable, web-based CBM-I program for BD.

Methods: Individuals with high BD (N = 318) were randomized to a multi-session CBM-I (Sentence Word Association Paradigm [SWAP] with feedback) vs. control (SWAP without feedback) versus waitlist condition. Interpretation bias, BD and associated symptoms were assessed at baseline and post-intervention. Symptoms were monitored up to 1-week and 4-week follow-up. We further investigated transference effects to stress reactivity, as predicted by cognitive-behavioral models, at post-intervention.

Results: Appearance-related CBM-I led to a differential pre–post increase in adaptive interpretation patterns, particularly for appearance-related and social situations (d = 0.65–1.18). Both CBM-I and control training reduced BD, BDD symptom severity, and depression. However, CBM-I (vs. control and waitlist) improved appearance-related quality of life (d = 0.51), self-esteem (d = 0.52), and maladaptive appearance-related beliefs (d = 0.47). State stress reactivity was overall reduced in the CBM-I condition (vs. waitlist). Intervention effects largely held stable up to follow-ups. Treatment satisfaction was comparable to other CBM-I studies, with low rates of adverse reactions.

Discussion: These findings support assumptions of cognitive-behavioral models for BD, BDD, and ED, and suggest that web-based CBM-I is an efficacious and acceptable intervention option.

KEYWORDS

body dissatisfaction, body dysmorphic disorder, cognitive bias modification, eating disorders, internet treatment, interpretation bias

1 | INTRODUCTION

Body dysmorphic disorder (BDD) and eating disorders (ED) present prevalent, impairing mental disorders that take a chronic, debilitating course if left untreated (Buhlmann et al., 2010; Hudson, Hiripi, Pope, & Kessler, 2007). While BDD is marked by an excessive preoccupation with perceived appearance-related defects, ED involve emaciation, and overvaluation of thinness in anorexia nervosa, uncontrollable binge eating in binge-eating disorder, and additional compensatory behaviors in bulimia nervosa (American Psychiatric Association, 2013). Research has identified high body dissatisfaction (BD), that is, the negative evaluation of one's own body (Stice & Shaw, 2002), as a risk factor for BDD and ED (Fang & Wilhelm, 2015; Stice & Shaw, 2002). BD is common (Bohne, Keuthen, Wilhelm, Deckersbach, & Jenike, 2002; Frederick, Peplau, & Lever, 2006) and has been linked to disordered eating, low self-esteem, decreased quality of life, and depression (Brechan & Kvalem, 2015; Mond et al., 2013; Paxton, Neumark-Sztainer, Hannan, & Eisenberg, 2006). Addressing BD within interventions might thus constitute an important avenue to augment existing treatment programs for BDD and ED, which demonstrate good, yet optimizable efficacy (Harrison, Fernández de la Cruz, Enander, Radua, & Mataix-Cols, 2016; Linardon, Wade, de la Piedad Garcia, & Brennan, 2017).

Maladaptive interpretation bias, i.e., the tendency to misinterpret ambiguous situations (Mathews & MacLeod, 2005), could represent a potential treatment target in BD interventions, as it is postulated to maintain BD, BDD, and ED. Specifically, cognitive-behavioral models posit that maladaptive interpretations, fostered by negative appearance-related schemas, exacerbate negative emotions—for example, under stress—and dysfunctional appearance-related behaviors (Cash, 2011; Fairburn, Cooper, & Shafran, 2003; Wilhelm, Phillips, & Steketee, 2013). Indeed, prior studies have demonstrated that negative appearance-related interpretation bias is associated with BD (Rodgers & DuBois, 2016) and malleable via brief, computerized interpretation retraining programs (Cognitive Bias Modification for Interpretation, CBM-I; e.g., MacLeod, 2012). CBM-I-based induction of appearance-related interpretation bias in mentally healthy individuals has been shown to produce congruent changes in bias patterns and, inconsistently, reactivity to BDD- and ED-relevant stressors (Dietel et al., 2018; Korn, Dietel, & Hartmann, 2019). These findings undermine hypotheses about the maintaining role of interpretation bias and the therapeutic relevance of CBM-I programs for BD.

Overall, CBM-I functions through contingency-based learning and automatization of adaptive interpretation patterns (Koster, Fox, & MacLeod, 2009; Kuckertz & Amir, 2017). Across emotional disorders, in-lab CBM-I programs have been shown to produce small- to medium-sized effects on interpretation patterns and small effects on mood (Hallion & Ruscio, 2011; Jones & Sharpe, 2017; Menne-Lothmann et al., 2014, for meta-analyses). Recent web-based adaptions of multi-session CBM-I have demonstrated comparable effects, for example, in anxiety disorders (Brettschneider, Neumann, Berger, Renneberg, & Boettcher, 2015; Hoppitt et al., 2014; Salemink, Kindt, Rienties, & van den Hout, 2014), although their magnitude varied, for instance, as a function of control conditions. Specifically, most studies have compared interventions against closely matched active control conditions, which produced attenuated beneficial effects. As the contribution of methodological factors to these effects (e.g., expectancy) remains largely unknown, randomized controlled trials (RCT), including a waitlist, represent an essential next step in differentiating CBM-I efficacy (Wilver & Cougle, 2019).

Recent studies have explored the effects of CBM-I in BDD (Premo, Sarfan, & Clerkin, 2016; Summers & Cougle, 2016) and ED (see Matheson, Wade, & Yiend, 2019, for a review). For instance, Summers and Cougle (2018) demonstrated that a four-session in-lab CBM-I program reduced overall symptoms, but not drive for thinness, in individuals with bulimic symptoms. Matheson, Wade, and Yiend (2018) found that a single-session appearance-related CBM-I intervention, versus self-worth-related CBM-I and placebo, improved interpretation biases and BD with a medium-sized effect. Last, Wilver and Cougle (2019) showed that a web-based, eight-session CBM-I program decreased BDD symptoms, but did not outperform progressive muscle relaxation in this aspect. While these studies highlight the potential utility of CBM-I for BD, to our knowledge, no study has investigated the effects of web-based appearance-related CBM-I within an RCT. A further critical, yet understudied feature concerns the acceptability of web-based CBM-I. To date, only one prior web-based CBM-I study in social anxiety reported indices on treatment satisfaction, which ranged below more guided web-based interventions (Brettschneider et al., 2015). Examining acceptability, including user feedback, marks a vital element in the prospective optimization and implementation of CBM-I, for example, within face-to-face or internet-based cognitive-behavioral interventions for BD (see Gentile et al., 2019; Heinicke, Paxton, McLean, & Wertheim, 2007).

Thus, extending upon previous research, the current study investigated the efficacy and acceptability of a web-based multi-session CBM-I intervention in a double-blind RCT. To this end, individuals with high BD were randomized to a CBM-I training (Interpretation Modification Program, IMP) versus active control (Interpretation Control Condition, ICC) versus waitlist condition (Wait-List Condition, WLC). We used a dismantling design (i.e., with the CBM-I group and active control group differing only in the feedback mechanism), whereby we aimed to identify the active ingredient of CBM-I training. We administered measures of BD, depression, BDD and ED symptoms, self-esteem, and quality of life at pre- and post-intervention, 1-week and 4-week follow-up to evaluate durability and generalizability of training effects. We further assessed transference onto stress reactivity, measuring emotional responding to distressing ambiguous videos.

We hypothesized that the IMP (vs. ICC and WLC) group would report lower maladaptive interpretations, BD, BDD symptom severity, as well as higher self-esteem and body image-related quality of life at posttreatment. Given prior evidence on placebo effects in CBM-I (Menne-Lothmann et al., 2014), we further predicted the ICC to show attenuated (vs. IMP), but substantial (vs. WLC) reductions in interpretation patterns, symptoms, and stress reactivity. We expected effects to be maintained up to follow-ups. We further explored treatment

expectancy, satisfaction, and adverse reactions to investigate acceptability under web-based dissemination conditions.

2 | METHODS

2.1 | Trial design

This was a double-blind, parallel-group trial conducted at the University of Muenster. The study protocol was approved by the local ethics committee. The outcome of this study is reported in accordance with the Consolidated Standards for Reporting Trials (CONSORT) statement (Moher et al., 2010).

2.2 | Participants

Participants were recruited via announcements posted on newspapers, the Internet (e.g., social networks), flyers, and participant pools. Participants received time-contingent course credit for participation. Inclusion criteria were: (a) aged between 18 and 40 years; (b) high BD, that is, a total score ≥19 on the Body Image Questionnaire, "Rejecting body evaluation" subscale (FKB-20-AKB; Clement & Löwe, 1996); (c) no acute suicidality, that is, a total score <3 on the Depressive Symptom Inventory-Suicidality Subscale (DSISS; von Glischinski, Teismann, Prinz, Gebauer, & Hirschfeld, 2016); (d) no current diagnosis of any mental disorder; (e) no current

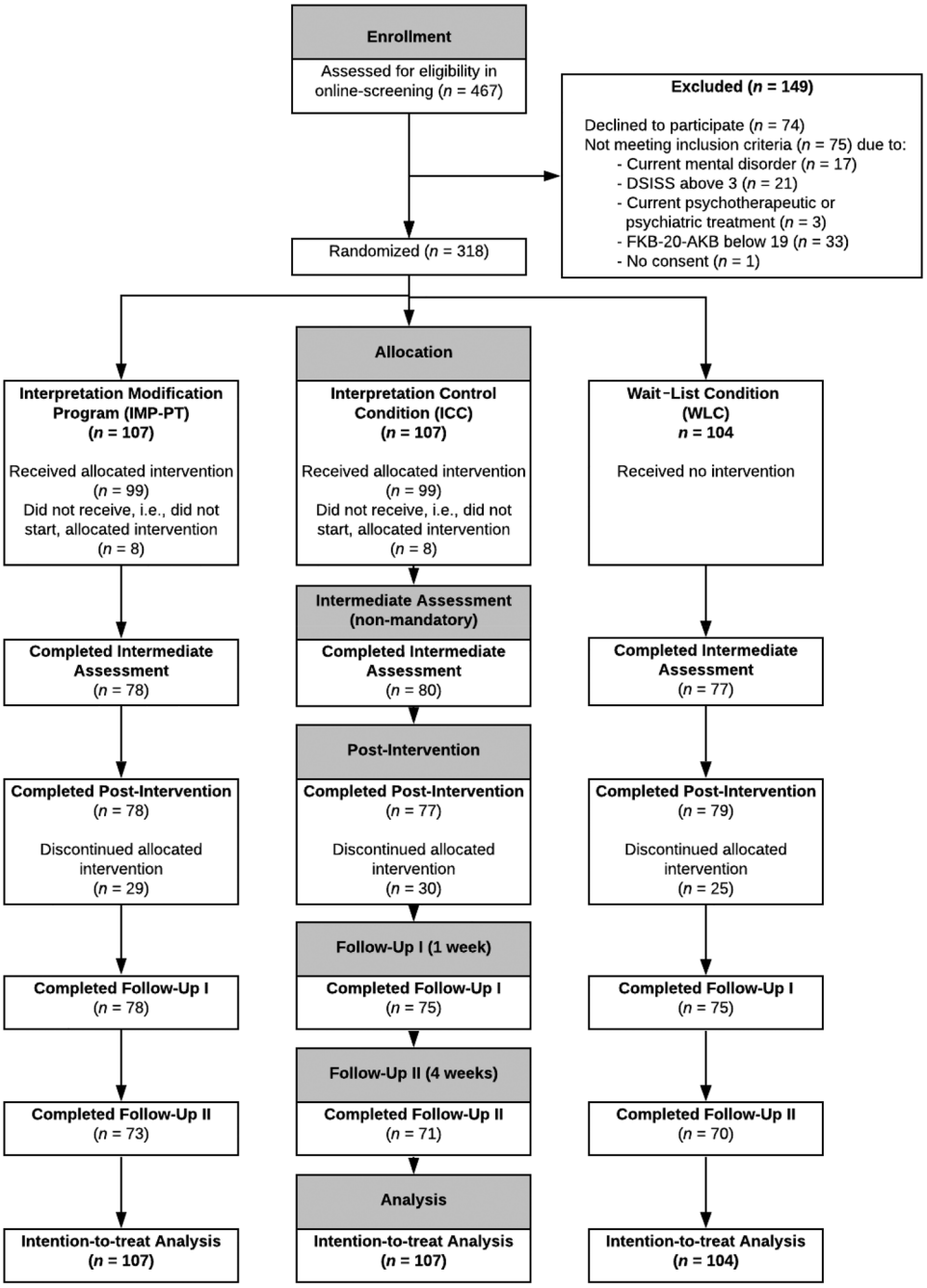

FIGURE 1 Flow of participants through phases of study according to the Consolidated Standards of Reporting Trials (CONSORT) guidelines. DSISS, Depressive Symptom Inventory-Suicidality Subscale; FKB-20-ABK, "Body Image Questionnaire," "Rejecting Body Image Subscale." *N* per assessment (post-intervention, follow-Up I, follow-Up II) represents total number of completers per time point, irrespective of prior completion

psychotherapeutic or psychiatric treatment. Participants not meeting inclusion criteria were redirected to a website containing contact information for mental health services. Three hundred eighteen individuals participated in this study (IMP: *N* = 107, ICC: *N* = 107, WLC: *N* = 104). Due to technical difficulties, baseline scores for *n* = 2 individuals were not recorded; however, cases were included in further multilevel analyses. Eighty-five (26.73%) participants dropped out before post-intervention assessment, resulting in *N* = 234 completers (73.58%) until post-intervention, and *N* = 182 (57.23%) until follow-up II. Figure 1 summarizes participant flow.

2.3 | Measures and materials

2.3.1 | Screening measures

To screen for elevated BD, we used the 10-item FKB-20-AKB (Clement & Löwe, 1996). Internal consistency was high in this sample (Cronbach's α = .88). To rule out suicidality, we employed the 4-item DSISS (von Glischinski et al., 2016).

2.3.2 | Primary outcome measures

Interpretation bias: Sentence Word Association Paradigm

We used the Sentence Word Association Paradigm (SWAP; Dietel et al., 2018), based on the Word Sentence Association Paradigm (Beard & Amir, 2008), to assess interpretation bias.

During assessment, 240 ambiguous sentence-word-combinations (50% positive vs. 50% negative word) were presented once, supplemented by 10 practice trials at pre-intervention. To measure transference, sentence-word-combinations consisted of 80 ambiguous appearance-related (e.g., "On the beach, you are wearing tight-fitting clothes."), 80 social (e.g., "You are giving a speech and everyone is laughing."), and 80 generally threatening scenarios (e.g., "There will be some changes in the company in the near future."). Situation sets for all categories were generated based on expert consensus and prevalidated (see Dietel, Möllmann, Bürkner, Wilhelm, & Buhlmann, 2019).

Trials started with a black fixation cross, displayed in the center of a white screen for 500 ms, replaced by an ambiguous sentence (e.g., "You see yourself in the mirror in bright lights."). After 3,500 ms, the interpretation (positive, e.g., "content," or negative, for example, "discontent") appeared centrally. Participants had to indicate as fast as possible whether sentence and word were related (pressing "L"/Yes vs. "S"/No on keyboard). The next trial was initiated upon button press. Decisions and reaction times were recorded.

Symptom measures

We assessed BD using the 7-item appearance evaluation (AE) and 12-item appearance orientation (AO) subscales of the Multidimensional Body-Self Relations Questionnaire-Appearance Scales (MBSRQ-AS; Cash, 2000; Vossbeck-Elsebusch et al., 2014). We further measured BDD symptom severity using the 18-item body dysmorphic symptoms inventory ("Fragebogen körperdysmorpher Symptome," Body Dysmorphic Symptoms Inventory, FKS; Buhlmann, Wilhelm, Glaesmer, Brähler, & Rief, 2009). To investigate ED and BDD symptoms, we rephrased the DSM-5 criteria (American Psychiatric Association, 2013) into dichotomous items ("Agree"/"Disagree"), employing a previously used 6-item version for BDD (Möllmann, Dietel, Hunger, & Buhlmann, 2017). For anorexia nervosa, criteria A (minimal body weight) and B (fear of gaining weight) were each represented in one, criterion C (body image disturbance) was represented in four items, and differentiation between restrictive versus binge-eating/purging type in two items. For bulimia nervosa, criteria B (compensatory behaviors) and D (shape- and weight-related self-evaluation) were rephrased to one, criteria A (episodes of binge eating) and C (frequency of bulimic behavior) to two items each. Internal consistencies for these measures were high across time points (AE: α = .87–.91; AO: α = .83–86; FKS: α = .86–.88).

2.3.3 | Secondary outcome measures

We assessed depression via the Patient Health Questionnaire, Depression Module (PHQ-9; Gräfe, Zipfel, Herzog, & Löwe, 2004; Kroenke & Spitzer, 2002), appearance-related beliefs via the 12-item Appearance-Schemas Inventory Revisited (ASI-R; Cash & Labarge, 1996; Grocholewski, Tuschen-Caffier, Margraf, & Heinrichs, 2011), body-image related quality of life via the Body Image Quality of Life Inventory (BIQLI; Cash & Fleming, 2002) and self-esteem via the 12-item Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965; von Collani & Herzberg, 2003). Internal consistencies in this study were high across time points (PHQ-9: α = .75–.83, ASI-R: α = .86–.89, BIQLI: α = .93–.94, RSES: α = .89–.91).

2.3.4 | Stress reactivity: Video stressor task

To assess transference effects to stress reactivity, we presented seven video clips containing distressing, ambiguous situations. Videos were created based on expert consensus (Appendix A). After each video, participants rated state (a) valence (i.e., unpleasantness), (b) urge to check their appearance, (c) BD and (d) urge to avoid, based on previous stressor designs (Premo et al., 2016; Summers & Cougle, 2016).

Treatment acceptability: Expectancy, satisfaction, adverse reactions, and feedback

We used a modified 4-item version of the Credibility-Expectancy-Questionnaire (CEQ; Devilly & Borkovec, 2000) to measure treatment expectancy and credibility at baseline. Further, we assessed treatment satisfaction with the 8-item Client Satisfaction Questionnaire (CSQ-8; Schmidt, Lamprecht, & Wittmann, 1989). Internal consistencies were high in this study (CEQ: α = .81, CSQ-8: α = .93).

Additionally, we conceptualized a 7-item feedback questionnaire, whereby participants judged SWAP-related items on a visual analog scale ranging from 0 ("not at all") to 10 ("extremely," see Table 4 for items).

Participants could further provide written feedback. Adverse reactions (e.g., deterioration of mood, increased BD) and events (e.g., hospitalization) were investigated using a 10-item, self-conceptualized measure.

2.3.5 | Interventions

Interpretation Modification Program

To modify interpretation bias, participants in the IMP received the appearance-related SWAP including decision-contingent feedback. Endorsing positive and rejecting negative word-sentence combinations was reinforced (feedback: "CORRECT" displayed for 1,000 ms), all other reactions were corrected (feedback: "INCORRECT") at a 100% contingency. A success bar on the right side of the screen displayed participants' cumulative correct responses (Dietel et al., 2018).

During training sessions, participants saw 60 randomized appearance-related situations, presented twice, once with a positive and once with a negative interpretation. Ten filler trials (Dietel et al., 2018; Möbius, Tendolkar, Lohner, Baltussen, & Becker, 2015), presented twice, were added to prevent participants from exclusively reacting to word valence, resulting in 140 trials per session.

Interpretation Control Condition

Participants in the ICC received the SWAP without feedback.

Wait-List Condition

Participants in the WLC attended a waiting period of 2 weeks.

2.3.6 | Procedure

The procedure is depicted in Figure 2. The study was conducted via a Cake PHP-based online platform. Participants were randomly assigned to the IMP, ICC, or WLC following a block randomization plan created through http://www.randomization.com/. Throughout training, the IMP and ICC were instructed to complete as many sessions as possible and received reminder e-mails after 1, 3, and 5 days upon not

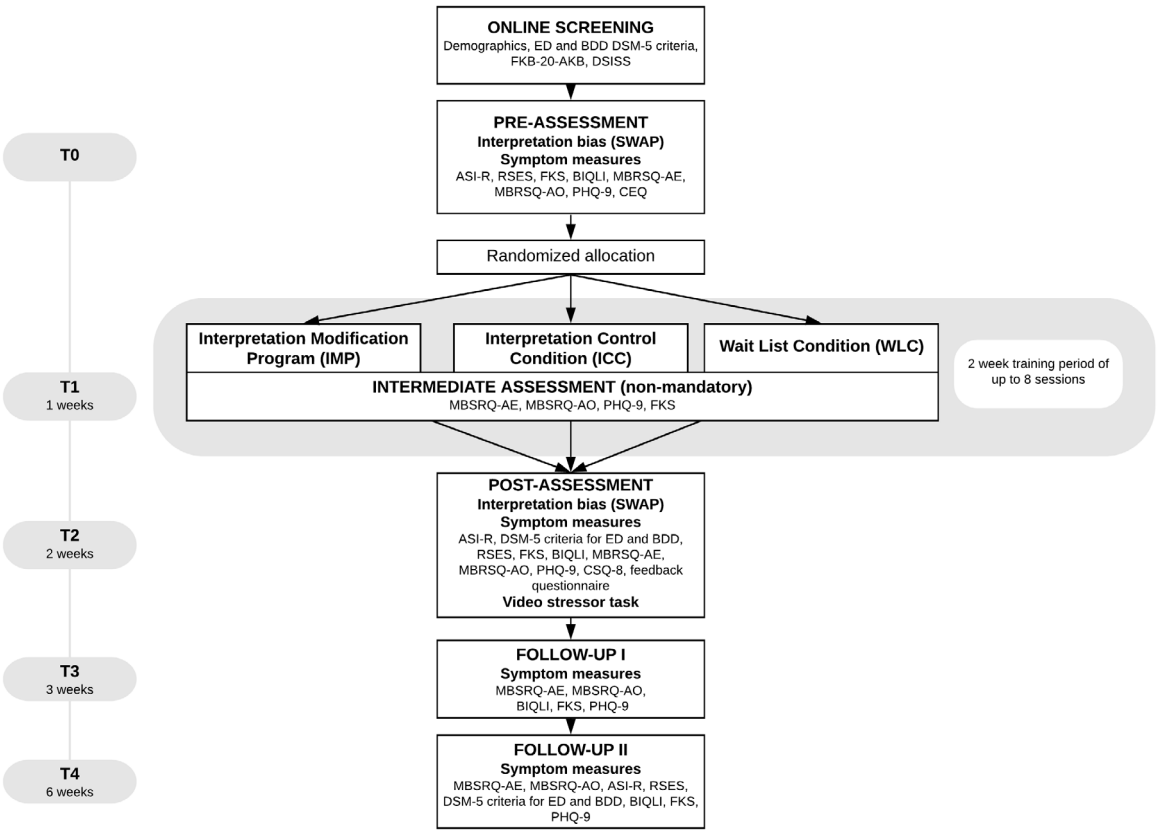

FIGURE 2 Study Procedure. ASI-R, Appearance Schemas Inventory; BDD, body dysmorphic disorder; BIQLI, Body Image Quality of Life Inventory; CEQ, Credibility Expectancy Questionnaire; CSQ-8, Client Satisfaction Questionnaire; DSISS, depressive symptom inventory-suicidality subscale; ED, eating disorders; FKB-20-AKB, Body Image Questionnaire, "Rejecting body evaluation subscale;" FKS, "Fragebogen körperdysmorpher Symptome," Body Dysmorphic Symptoms Inventory; MBSRQ-AO, Multidimensional Body-Self Relations Questionnaire, Appearance Orientation; MBSRQ-AE, Multidimensional Body-Self Relations Questionnaire, Appearance Evaluation; PHQ-9, Patient Health Questionnaire (Depression Module); RSES, Rosenberg Self-Esteem Scale

completing sessions. Upon not logging into their account for 10 consecutive days, participants were excluded and sent a drop out questionnaire, unless in waiting or training phase. Upon study completion, all participants were debriefed via e-mail and received time-contingent course credit.

2.4 | Design and statistical analyses

One-way analysis of variances, t-tests and chi-square analyses were conducted using IBM SPSS 25 to investigate psychometric or demographic baseline differences, treatment acceptability and session attendance.

TABLE 1 Baseline demographics and psychometrics

	IMP (*n* = 107) *M* (*SD*)/%	ICC (*n* = 105)[a] *M* (*SD*)/%	WLC (*n* = 104) *M* (*SD*)/%	Total (*N* = 316)[a] *M* (*SD*)/%
Age (years)	23.97 (10.63)	22.96 (5.28)	22.18 (3.82)	23.05 (7.24)
FKB-20-AKB	30.80 (6.99)	30.72 (7.06)	30.51 (7.23)	30.68 (7.08)
Completed sessions	4.57 (2.72)	4.28 (2.55)	—	—
Gender *N*				
Male	10.30%	7.50%	7.70%	8.50%
Female	89.70%	92.50%	92.30%	91.5%
Current level of education				
College student	86.0%	85.0%	93.3%	88.1%
High school student	0.0%	0.0%	1.0%	0.3%
Other[b]	14.0%	15.0%	5.8%	11.7%
Body areas of concern				
Skin tone/texture	43.9%	45.8%	41.3%	43.7%
Hair	20.6%	19.6%	22.1%	20.8%
Nose	26.2%	26.2%	26.9%	26.4%
Eyes	8.4%	5.6%	1.9%	5.3%
Ears	1.9%	4.7%	9.6%	5.3%
Mouth/teeth	7.5%	8.4%	5.8%	7.2%
Overall appearance of head/face	0.0%	0.9%	1.0%	0.6%
Breast/chest	29.9%	33.6%	41.3%	34.9%
Muscularity	22.4%	17.8%	19.2%	19.8%
Stomach	13.6%	15.8%	18.3%	16.2%
Hips	0.9%	1.8%	0.0%	0.9%
Bottom	0.9%	2.7%	2.0%	1.8%
Back	0.0%	0.0%	1.9%	0.6%
Legs/feet	14.7%	16.6%	11.7%	14.3%
Arms/hands	11.2%	7.4%	10.6%	9.7%
Genitals	8.4%	13.1%	10.6%	10.7%
Overall body build	0.0%	0.9%	0.0%	0.3%
Height	0.0%	0.0%	0.0%	0.0%
Weight	0.9%	0.0%	0.0%	0.3%
Ethnical features	0.0%	1.9%	1.9%	1.3%
Other[c]	34.6%	20.7%	13.9%	24.1%
Number of areas of concern	2.49 (1.43)	2.52 (1.23)	2.42 (1.18)	2.48 (1.28)

*Note: SD*s are in parentheses.

Abbreviations: FKB-20-ABK, "Body Image Questionnaire," "Rejecting Body Image Subscale"; ICC, Interpretation Control Condition; IMP, Interpretation Bias Modification Program; WLC, Wait-List Condition.

[a]Due to technical difficulties, questionnaire data of *n* = 2 participants were not recorded.

[b]"Other" includes: currently employed (full-time or part-time), unemployed.

[c]"Other includes": chin, cheeks, forehead, shoulders, and individual combinations of body parts.

To examine effects of condition on outcomes, we conducted multilevel analyses via Bayesian Markov chain Monte Carlo (MCMC) methods with a two-level hierarchy using the *brms* package (Bürkner, 2017), which is based on the probabilistic programing language Stan (Carpenter et al., 2017) for *R* (R Core Team, 2018). Within this approach, outcome variables (Level 1) are nested within different participants (Level 2). Multilevel modeling accounts for sample hierarchy, interdependencies between data, and some missing data patterns (Hoffman & Rovine, 2007; Quené & Van den Bergh, 2004). Bayesian MCMC methods estimate uncertainty both for overall and varying effects across participants (Bürkner, 2017). We conducted intention-to-treat (ITT) and per-protocol analysis (PP). ITT was performed using available data, including participants who did not complete follow-up assessments.

The a priori power analysis using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) and assuming a small effect size for pre–post symptom reduction (see Menne-Lothmann et al., 2014; analysis parameters: $1 - \beta = .95$, $\alpha = .05$, $f_2 = 0.125$, $r_{between} = 0.5$) yielded a total sample size of 210 to observe these effects.

3 | RESULTS

The demographic properties of this sample are shown in Table 1. Groups did not differ on baseline demographic or psychometric variables (all *p*s > .18, $\eta_p^2 = 0.01$).

3.1 | Session compliance and dropout

Of the 318 randomized participants, 78 (72.8%) in the IMP, 77 (72.0%) in the ICC, and 79 (74.0%) in the WLC completed the post-intervention assessment. Dropout rates were nearly equal in all three arms (27.2% in IMP vs. 28.0% in ICC vs. 26.0% in WLC; $\chi^2[8] = 6.56$, $p = .59$, $V = 0.10$). Early terminators conducted on average 1.51 (1.50) sessions in the IMP, versus 1.57 (1.65) sessions in the ICC. Seventy three (68.22%) participants in the IMP, 71 (66.36%) in the ICC, and 70 (67.31%) in the WLC completed follow-up II assessments. Completers in IMP and ICC did not differ in session attendance, making the results less attributable to differential engagement (Table 1).

3.2 | Intention-to-treat analysis

3.2.1 | Primary outcome measures

Interpretation bias: Endorsement rates

As evident in Table 2, for appearance-related situations, the IMP and ICC showed substantial pre-post reductions in negative and pre–post increases in positive interpretations, with no such changes in the WLC. For social situations, the pattern of results was identical. For general situations, there were reductions in negative and increases in positive interpretations largely irrespective of condition.

TABLE 2 Effects of conditions on endorsement rates on the Sentence Word Association Paradigm

		Pre-intervention			Post-intervention			Effect sizes[a]		
Endorsement rates (%)		IMP (*N* = 107) *M* (*SD*)	ICC (*N* = 107) *M* (*SD*)	WLC (*N* = 104) *M* (*SD*)	IMP (*N* = 78) *M* (*SD*)	ICC (*N* = 77) *M* (*SD*)	WLC (*N* = 79) *M* (*SD*)	IMP Pre–post	ICC Pre–post	WLC Pre–post
Appearance-related situations	Positive	60.16 (16.01)	60.61 (15.24)	62.09 (15.28)	83.97 (17.99)	66.14 (19.67)	59.43 (18.23)	1.18	0.28	0.19
	Negative	42.66 (21.78)	42.80 (19.82)	39.86 (19.37)	15.35 (16.61)	33.38 (23.23)	41.27 (21.97)	−1.12	−0.46	−0.20
Social situations	Positive	60.70 (14.82)	60.93 (13.15)	60.58 (13.11)	84.29 (14.00)	72.08 (15.71)	63.16 (14.31)	1.51	0.70	0.17
	Negative	46.80 (18.48)	49.46 (17.23)	47.60 (16.82)	15.03 (15.04)	36.43 (21.10)	42.59 (20.23)	−1.61	−0.84	−0.20
General situations	Positive	69.04 (13.47)	70.14 (12.43)	68.97 (12.46)	86.99 (10.85)	82.63 (13.79)	81.80 (14.79)	1.34	0.74	0.72
	Negative	56.61 (17.46)	55.05 (17.87)	56.99 (17.16)	45.77 (16.81)	43.64 (17.38)	46.17 (13.40)	−0.65	−0.72	−0.67

Abbreviations: ICC, Interpretation Control Condition; IMP, Interpretation Bias Modification Program; WLC, Wait-List Condition.
[a]Cohen's *d* values for pre- to post-intervention (pre–post): < 0.20 = negligible effect size, 0.20 = small effect size, 0.50 = medium effect size, and 0.80 = large effect size.

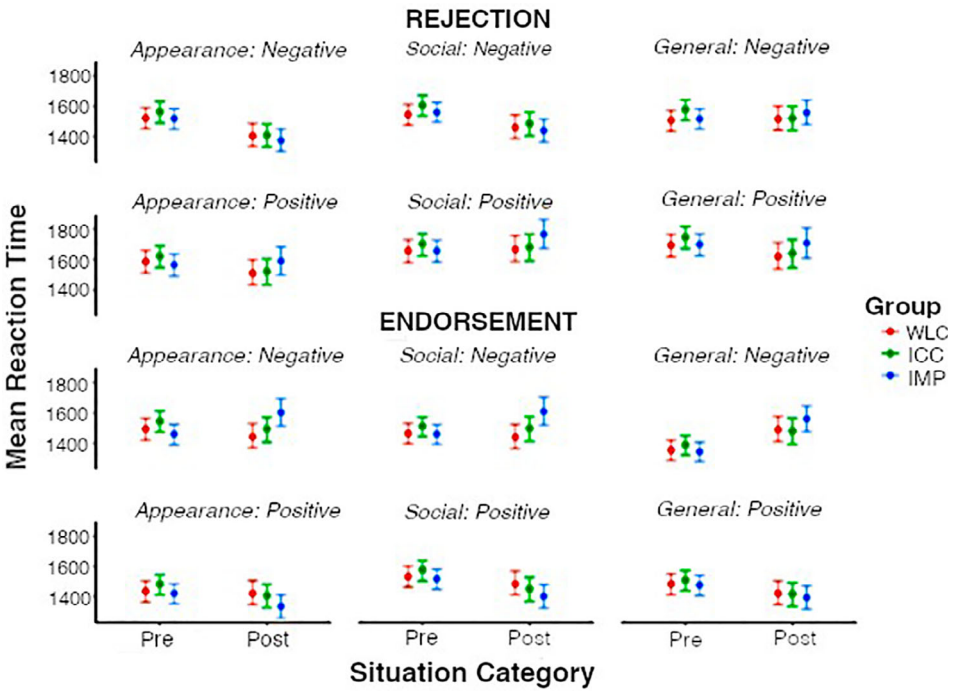

FIGURE 3 Mean reaction times (in ms) per situation and response category. Error bars represent 95% credibility intervals of the mean. ICC, Interpretation Control Condition; IMP, Interpretation Bias Modification Program; WLC, Wait-List Condition [Color figure can be viewed at wileyonlinelibrary.com]

Interpretation bias: Reaction times

As shown in Figure 3, for endorsements of *positive* interpretations, RT substantially decreased in the IMP (vs. ICC and WLC) in appearance-related and social, but not general situations. However, RT in the ICC decreased from pre-to post-intervention only in social situations. For endorsements of *negative* interpretations, RT increased in the IMP (vs. ICC and WLC) across all situation categories—however, RT also increased for the ICC and WLC for general situations.

For rejections of *positive* interpretations, reaction times remained overall stable from pre-to post-intervention in appearance-related situations. However, the IMP (vs. ICC and WLC) exhibited slower rejection of positive social interpretations from pre- to post-intervention. For general situations, RT decreased from pre-to post-intervention for the ICC and WLC (vs. IMP). For rejections of *negative* interpretations, there was a pre–post decrease in RT across groups for appearance-related and social, but not general situations.

Symptom measures

As shown in Table 3, concerning changes from pre-intervention to intermediate assessment, no substantial differences were found across all groups and measures (MBSRQ-AO, MBSRQ-AE, FKS, PHQ-9; see OSF supplements for results).

Concerning changes from pre- to post-intervention, analyses revealed a substantial pre–post reduction in appearance orientation (MBSRQ-AO) for the IMP (b = −0.12, 95%–CI = [−0.23, −0.02]), and ICC (b = −0.13, 95%–CI = [−0.23, −0.03]), but not the WLC. However, appearance evaluation (MBSRQ-AE) increased between pre- and post-intervention in the IMP (b = 0.14, 95%–CI = [0.02, 0.26]), but not in the ICC (b = 0.01, 95%–CI = [−0.11, 0.13]) or WLC. Regarding BDD symptom severity (FKS), there was a substantial pre–post reduction for the IMP (b = −2.77, 95%–CI = [−4.52, −1.11]) and ICC (b = −2.10, 95%–CI = [3.83, −0.32]), which was absent in the WLC. Maladaptive appearance-related beliefs (ASI-R) decreased substantially between pre–post intervention in the IMP (b = −0.17, 95%–CI = [−0.28, −0.05]), but not in the ICC (b = −0.08, 95%–CI = [−0.19, 0.03]) and WLC. Regarding self-reported DSM-5 criteria for BDD, anorexia nervosa (AN) and bulimia nervosa (BN), no substantial changes were observed from pre- to post-intervention for the IMP (BDD: b = −1.42, 95%–CI = [−3.70, 0.81]; AN: b = 1.80, 95%–CI = [−0.94, 4.63], BN: b = 3.19, 95%–CI = [−4.66, 14.56]) or ICC (BDD: b = 0.37, 95%–CI = [−1.64, 2.58]; AN: b = 1.17, 95%–CI = [−1.20, 3.73], BN: b = 4.73, 95%–CI = [−3.09, 16.00]).

3.2.2 | Secondary outcome measures

We found a significant pre-post decrease in depression (PHQ-9) in the IMP (b = −0.95, 95%–CI = [−1.88, −0.03]), but not in the ICC (b = −0.79, 95%–CI = [−1.77, 0.13]) and WLC. Considering quality of life (BIQLI), the IMP showed a differential pre–post increase (b = 6.81, 95%–CI = [2.58, 10.75]), which was absent in the ICC (b = 2.17, 95%–CI = [−1.78, 6.07]), and WLC. Self-esteem (RSES) increased substantially between pre- and post-intervention in the IMP (b = 4.80, 95%–CI = [1.50, 8.07]), but not in the ICC (b = 1.05, 95%–CI = [−2.09, 4.24]) or WLC.

3.2.3 | Stress reactivity

As shown in Figure 4, only the IMP group (vs. WLC) reported lower state unpleasantness (b = −7.00, 95%–CI = [−13.00, −0.72]), BD

TABLE 3 Effects of conditions on symptom outcomes

	Pre-intervention			Mid-intervention			Post-intervention			Follow-Up I (1 week)			Follow-Up II (4 weeks)			Effect sizes[a]					
	IMP (n = 107)	ICC (n = 105)[b]	WLC (n = 104)	IMP (n = 78)	ICC (n = 80)	WLC (n = 77)	IMP (n = 78)	ICC (n = 77)	WLC (n = 79)	IMP (n = 75)	ICC (n = 72)	WLC (n = 73)	IMP (n = 72)	ICC (n = 70)	WLC (n = 70)	IMP			ICC		
Symptom severity	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	Pre–post	Pre-FUI	Pre-FUII	Pre–post	Pre-FUI	Pre-FUII
FKS	23.05 (8.63)	24.16 (8.69)	22.00 (8.89)	19.47 (8.15)	20.88 (9.02)	19.20 (8.61)	19.00 (7.78)	20.45 (9.39)	20.72 (9.72)	16.23 (7.30)	17.97 (9.53)	18.96 (9.85)	16.47 (8.19)	17.73 (9.21)	17.90 (9.47)	−0.44	−0.57	−0.34	−0.36	−0.39	−0.35
MSRQ-AO[c]	3.62 (0.60)	3.57 (0.57)	3.70 (0.58)	3,55 (0.59)	3.52 (0.49)	3.68 (0.49)	3.48 (0.54)	3.41 (0.60)	3.68 (0.55)	3.46 (0.53)	3.41 (0.54)	3.69 (0.55)	3.44 (0.59)	3.38 (0.57)	3.64 (0.51)	−0.34	−0.31	−0.26	−0.38	−0.36	−0.37
MSRQ-AE[c]	2.79 (0.73)	2.79 (0.73)	2.79 (0.73)	2.94 (0.76)	2.89 (0.76)	2.96 (0.75)	3.02 (0.78)	2.95 (0.86)	3.00 (0.79)	3.00 (0.72)	2.96 (0.81)	2.92 (0.80)	3.13 (0.72)	3.06 (0.90)	3.02 (0.76)	0.34	0.44	0.36	0.02	0.10	0.12
PHQ-9	6.64 (3.59)	7.37 (3.90)	7.07 (4.78)	6.46 (4.28)	6.80 (3.79)	6.48 (3.86)	5.48 (3.97)	6.05 (3.68)	6.31 (4.48)	4.92 (3.63)	5.59 (3.91)	5.86 (4.15)	5.21 (4.09)	5.70 (4.18)	5.60 (4.29)	−0.41	−0.47	−0.28	−0.35	−0.34	−0.26
BIQLI	2.29 (18.59)	1.59 (19.21)	6.16 (20.07)	—	—	—	7.53 (17.46)	3.01 (19.30)	4.61 (20.90)	6.12 (15.97)	2.04 (18.88)	3.05 (19.85)	3.45 (0.57)	3.41 (0.58)	3.56 (18.12)	0.51	0.60	0.39	0.12	0.10	0.14
ASI-R	3.64 (0.50)	3.57 (0.47)	3.65 (0.50)	—	—	—	3.46 (0.49)	3.44 (0.51)	3.61 (0.54)	—	—	—	3.44 (0.52)	3.35 (0.54)	3.62 (0.53)	−0.47	—	−0.48	−0.20	—	−0.44
RSES	65.58 (16.03)	65.58 (16.03)	67.69 (16.34)	—	—	—	70.19 (14.91)	68.55 (15.74)	69.51 (16.33)	—	—	—	69.56 (16.70)	69.84 (16.47)	70.00 (14.58)	0.52	—	0.40	0.11	—	0.17
Disorder criteria	%	%	%				%	%	%				%	%	%						
BDD	13.1	8.4	8.7	—	—	—	3.7	5.6	4.8	—	—	—	6.5	6.5	6.7						
AN	4.7	8.4	10.6	—	—	—	3.7	4.7	2.9	—	—	—	3.7	5.6	4.8						
BN	1.9	2.8	1.9	—	—	—	1.3	2.6	1.3	—	—	—	0.0	2.8	1.9						

Abbreviations: AN, anorexia nervosa; ASI-R, Appearance Schemas Inventory; BDD, body dysmorphic disorder; BIQLI, Body Image Quality of Life Inventory; BN, bulimia nervosa; FKS, "Fragebogen körperdysmorpher Symptome," Body Dysmorphic Symptoms Inventory; ICC, Interpretation Control Condition; IMP, Interpretation Bias Modification Program; MBSRQ-AE, Multidimensional Body-Self Relations Questionnaire, Appearance Evaluation; MBSRQ-AO, Multidimensional Body-Self Relations Questionnaire, Appearance Orientation; PHQ-9, Patient Health Questionnaire (Depression Module); RSES, Rosenberg Self-Esteem Scale; WLC, Wait-List Condition.

[a]Cohen's d values for pre- to post-intervention (pre–post), pre-intervention to 1-week follow-up (Pre-FUI), versus 4-week follow-up (Pre-FUII): < 0.20 = negligible effect size, 0.20 = small effect size, 0.50 = medium effect size, 0.80 = large effect size.

[b]Due to technical difficulties, questionnaire data of n = 2 participants were not recorded.

[c]Higher AE scores reflect more positive body evaluation; higher AO scores index more body-influencing behavior.

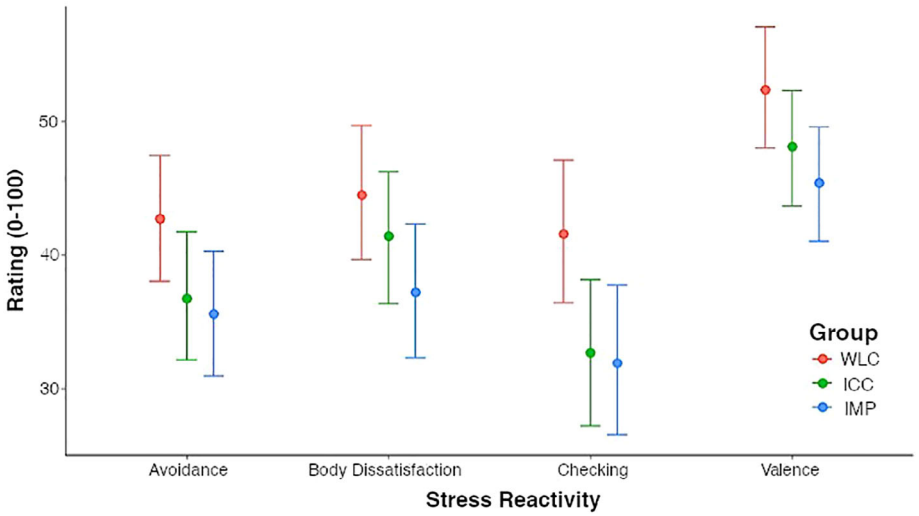

FIGURE 4 Effects of conditions on state stress reactivity (range: 0–100). Error bars represent 95% credibility intervals of the mean. ICC, Interpretation Control Condition; IMP, Interpretation Bias Modification Program; WLC, Wait-List Condition. Valence = mean valence rating of stressor videos (0–100), body dissatisfaction = mean BD rating of stressor videos (0–100), rituals = mean rituals rating of stressor videos (0–100), avoidance = mean avoidance rating of stressor videos (0–100) [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 4 Credibility and treatment satisfaction for treatment completers

Treatment expectancy and satisfaction[a]	IMP (n = 78) M (SD)	ICC (n = 77) M (SD)
CEQ	21.08 (6.04)	21.27 (6.25)
CSQ-8	18.24 (4.96)	15.90 (4.52)
Feedback[b]	M (SD)	M (SD)
Comprehendible	9.12 (1.81)	9.13 (1.82)
Difficult to use	2.08 (1.65)	2.13 (1.48)
Compatible with schedule	7.34 (2.54)	7.03 (2.64)
Entertaining	3.73 (2.29)	3.72 (2.62)
Helpful	4.64 (2.66)	1.89 (2.92)
Time-consuming	7.44 (2.20)	6.21 (2.21)
Demanding	4.95 (2.51)	5.28 (2.66)
Self-reported adverse reactions and events[c]	%	%
Deterioration of mood	1.28	5.19
Increase in anxiety	1.28	1.30
Increase in body dissatisfaction	3.85	5.19
Increase in suicidal ideation	0.00	1.30
Hospitalization	0.00	0.00

Abbreviations: CEQ, Credibility Expectancy Questionnaire; CSQ-8, Client Satisfaction Questionnaire; ICC, Interpretation Control Condition; IMP, Interpretation Bias Modification Program.

[a]Treatment expectancy assessed at pre-intervention, treatment satisfaction assessed at post-intervention.

[b]Feedback questions were conceptualized specifically for this study. Theoretical range: 0 = not at all; 10 = extremely.

[c]Self-reported adverse reactions as related to training contents, administered at post-intervention.

(b = −7.13, 95%-CI = [−13.95, −0.50]) and urge to avoid (b = −6.97, 95%-CI = [−13.66, −0.08]) during the stressor. However, both IMP (b = 9.74, 95%-CI = [−17.18, −2.13]) and ICC (b = −9.04, 95%-CI = [−16.71, −0.93]) exhibited fewer urges to check their appearance, compared to the WLC. All other differences between the ICC and WLC remained non-substantial (unpleasantness: b = −4.32, 95%-CI = [−10.32, 1.81], BD: b = −2.95, 95%-CI = [−9.78, 4.23], urge to avoid: b = −5.91, 95%-CI = [−12.63, 0.68]).

3.2.4 | Follow-up effects

Substantial pre–post effects were maintained for outcome measures in the IMP and ICC at follow-ups, except for the PHQ-9 at 4-week follow-up (Table 3).

3.2.5 | Per-protocol analysis

All effects were replicable in the PP analysis.

3.2.6 | Treatment credibility, expectancy, satisfaction, and adverse reactions

There were no baseline differences in treatment expectancy and credibility (see Table 4; t[153] = 0.20, p = .84, d = 0.03). At post-intervention, the IMP (vs. ICC) was overall more satisfied with the intervention (t[153] = 3.80, p = .002, d = 0.61), deeming it more helpful (t[147] = 4.45, p < .001, d = 0.73), but also more time-consuming (t[147] = 3.41, p = .001, d = 0.56). Rates of self-reported adverse reactions were low in both active groups, with the IMP reporting transient decreased mood, increased anxiety and BD. No major adverse events were reported.

Feedback regarding optimization of the IMP program (N = 52) concerned item wording (34.6%), task monotony (23.1%), task duration (17.3%), technical difficulties (11.5%), subjective impression that the training had no positive effect (9.6%), and lack of feedback concerning one's time remaining (3.8%).

4 | DISCUSSION

Extending prior findings, this study investigated the effects of a multi-session, web-based, appearance-related CBM-I program to explore its therapeutic potential, and acceptability.

Consistent with predictions and prior studies (e.g., Premo et al., 2016; Summers & Cougle, 2016), the IMP enhanced adaptive interpretation patterns for appearance-related and social scenarios. That is, endorsement rates for positive interpretations increased and those for negative interpretations decreased with training in these categories. Considering reaction times, positive endorsements occurred more rapidly and negative endorsements occurred more slowly in the IMP (vs. ICC and WLC) at post-intervention, while changes in rejection times were overall less modulated by condition. Interpretation patterns for general situations changed irrespective of group, which can likely be attributed to generic expectancy and practice effects via the baseline assessment. Overall, these findings suggest that multi-session CBM-I successfully modified interpretation bias for both trained (i.e., appearance-related) and non-trained (i.e., social), categorically proximal stimulus material, which is in line with results from a prior in-lab CBM-I study for BDD (Premo et al., 2016).

As expected, the IMP further led to overall small- to medium-sized pre-post decrease in appearance orientation, BDD severity, maladaptive appearance-related beliefs and depression, while boosting appearance evaluation, body-related quality of life and self-esteem. With the exception of depression, effects were maintained at both follow-ups and within PP analyses. These results are consistent with other multi-session CBM-I studies, demonstrating transference of interpretation bias changes to symptoms (Matheson et al., 2019; Salemink et al., 2014).

Regarding stress reactivity, the IMP (vs. WLC) was further associated with overall lower valence (i.e., unpleasantness), state BD and urge to check ratings during exposure to video stressors. Notably, the ICC did not differ substantially from the WLC on these variables. Given the overall small effects and lack of a baseline stress reactivity assessment, this finding is suggestive, albeit not conclusive, of CBM-I-related transference to stress reactivity. However, this is one of the first studies to identify such transference within the full sample, given prior inconsistent results (Premo et al., 2016; Summers & Cougle, 2016).

In sum, these results corroborate assumptions of cognitive-behavioral models whereby changes in appearance-related interpretation bias modulate emotions and behavioral tendencies, for example, under stress (Cash, 2011; Fairburn et al., 2003; Wilhelm et al., 2013). Effect sizes for pre–post bias and symptom change, as well as for stress reactivity, were largely in line with meta-analytic findings (Hallion & Ruscio, 2011; Jones & Sharpe, 2017; Menne-Lothmann et al., 2014). Importantly, appearance-related CBM-I effects generalized to symptoms associated with BD, for example, self-esteem and appearance-related maladaptive beliefs. Further, changes in clinical outcomes, for example, BD, occurred first at post-intervention, not intermediate assessment. While this finding may be partially explained by insufficient sensitivity for change for self-report measures, it is in line with meta-analyses indicating that multi-session training is more efficacious than brief CBM-I (Jones & Sharpe, 2017; Menne-Lothmann et al., 2014). The generalization and durability of gains illustrate the potential therapeutic value of appearance-related CBM-I. However, it should be noted that, overall, pre–post intervention effect sizes were mostly small, and training did not modulate the prevalence of self-report BDD or ED symptoms. Further, effects on depression were not maintained on follow-up. Overall, these findings underline the necessity to integrate CBM-I into a broader treatment context or within stepped care approaches (Beard, Rifkin, Silverman, & Björgvinsson, 2019).

For most measures, we found similar, albeit attenuated beneficial effects for the ICC. As discussed, such effects have been previously demonstrated in in-lab and web-based CBM-I studies (Menne-Lothmann et al., 2014). As illustrated in this RCT, effects are more substantial in the ICC than in the WLC, and thus unlikely exclusively generated by expectancy. However, as argued in previous studies, recurrent presentation of valent interpretations at a 50:50 ratio might per se induce contingency-based learning and higher-order reappraisal (Dietel et al., 2018). As the present study is not able to further discern the contribution of these processes, it appears promising to investigate underlying processes and predictors in CBM-I. Given the ICC effects, research might also explore other control groups, for example, yoked controls, depending on the research question under study (Blackwell, Woud, & MacLeod, 2017). Moreover, given the enhanced effects in the IMP (vs. ICC), results suggest that contingencies established through feedback might be the active ingredients of SWAP-based CBM-I, driving associative learning (Gonsalves, Whittles, Weisberg, & Beard, 2019). Nevertheless, contrasting different versions of such feedback might be next steps in optimizing appearance-related CBM-I for clinical use.

Regarding acceptability, the IMP was more satisfied with the intervention, deeming it to be more helpful than the ICC. Given no between-group baseline differences in expectancy, these results reflect the relative perceived efficacy of the IMP design. These findings are also consistent with prior acceptability ratings for web-based CBM-I (Brettschneider et al., 2015). However, the IMP (vs. ICC) group did not evaluate the intervention to be more entertaining and criticized item wording, task duration, and monotonicity. Future studies should explore ways to address these issues, for example, by using idiosyncratic sentence-word-combinations and gamification (Boendermaker, Prins, & Wiers, 2015). Improving engagement might additionally prove useful in lowering dropout, particularly in web-based settings (Melville, Casey, & Kavanagh, 2010). Last, the low rate of adverse reactions in this study suggests that web-based appearance-related CBM-I is safe to use, which is relevant for prospective clinical use.

This study is not without limitations. First, to create a scalable solution, participants of this non-preregistered trial were assessed online, which may limit the reliability and validity of some diagnostic and experimental tasks. However, numerous studies have demonstrated comparable outcomes for experimental tasks in lab- versus web-based settings (Dandurand, Shultz, & Onishi, 2008; Hilbig, 2016). Future studies might examine clinical status and further relevant

endpoints - for example, attention bias - under more controlled in-lab conditions. Second, in this study, the assessment and training paradigm (SWAP) were identical, as, to our knowledge, there is currently no alternative to investigating decision-based and RT-based components of interpretation bias simultaneously. However, future studies might assess interpretation bias using additional instruments, for example, appearance-related implicit tasks (Buhlmann, Teachman, & Kathmann, 2011). In this respect, longer follow-up assessments appear critical to evaluate the long-term effects of CBM-I on BD. Third, the absence of therapeutic guidance, while important to test standalone treatment effects, might have adversely affected treatment efficacy and satisfaction. Indeed, research indicates that guidance is beneficial in web-based interventions (Palmqvist, Carlbring, & Andersson, 2007), even if differential effects are small (Baumeister, Reichler, Munzinger, & Lin, 2014). Thus, it appears worthwhile to investigate the augmentative value of appearance-related CBM-I within more guided treatment rationales.

In conclusion, this study provides evidence for the efficacy and acceptability of appearance-related CBM-I in BD, emphasizing its potential as a treatment module. As the first three-armed RCT in this field, it further differentiates feedback-induced effects, mere associative processing effects, and expectancy effects, thus informing prospective research on underlying processes of CBM-I. This study also proposes alternative avenues for future research, for instance, investigating long-term follow-up and broader effects of appearance-related CBM-I, as well as its potential as an augmentation to CBT. Such investigations could further clarify the benefits of web-based appearance-related CBM-I within different settings, designs, and populations.

ACKNOWLEDGMENTS

We would like to thank Xenia Hartleb, Annika Junge, Raphael Rupprecht, Helena Loges, Julia Wichtrup, and Marieke Meier for their assistance in data collection and analysis. We would further like to thank Maik Hagenbruch for his assistance in programming the online platform for this study.

CONFLICT OF INTEREST

We wish to confirm that there are no known conflicts of interest associated with this publication. This study was supported by internal funding and did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions. SWAP stimuli, self-conceptualized measures and data analyses underlying this study are available online from our OSF repository (https://osf.io/4fpwx/?).

ORCID

Fanny Alexandra Dietel https://orcid.org/0000-0002-9911-6155
Paul-Christian Bürkner https://orcid.org/0000-0001-5765-8995
Johanna Schulte https://orcid.org/0000-0001-9795-1168
Sabine Wilhelm https://orcid.org/0000-0002-5316-2711
Ulrike Buhlmann https://orcid.org/0000-0003-4997-3236

ENDNOTES

[1]Items are available from our OSF repository.

[2]Unpublished web-based validation data (N = 110) in an unselected student sample showed a pre-post increase in state distress after viewing all videos, $t(109) = 3.19$, $p = .02$, $d = 0.30$. MBSRQ subscales were significantly correlated with mean state valence (AE: $r = -.39$, AO: $r = .36$), body dissatisfaction (AE: $r = -.58$, AO: $r = .26$), urge to check (AE: $r = -.30$, AO: $r = .40$), and urge to avoid (AE: $r = -.39$, AO: $r = .34$). Internal consistencies of the state scales across videos were high (discomfort: $\alpha = .79$; body dissatisfaction: $\alpha = .89$, urge to check: $\alpha = .83$, urge to avoid: $\alpha = .77$).

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Publishing. https://doi.org/10.1176/appi.books.9780890425596

Baumeister, H., Reichler, L., Munzinger, M., & Lin, J. (2014). The impact of guidance on internet-based mental health interventions—A systematic review. *Internet Interventions*, *1*, 205–215. https://doi.org/10.1016/j.invent.2014.08.003

Beard, C., & Amir, N. (2008). A multi-session Interpretation Modification Program: Changes in interpretation and social anxiety symptoms. *Behavior Research and Therapy*, *46*, 1135–1141. https://doi.org/10.1016/j.brat.2008.05.012

Beard, C., Rifkin, L. S., Silverman, A. L., & Björgvinsson, T. (2019). Translating CBM-I into real-world settings: Augmenting a CBT-based psychiatric hospital program. *Behavior Therapy*, *50*, 515–530. https://doi.org/10.1016/j.beth.2018.09.002

Blackwell, S. E., Woud, M. L., & MacLeod, C. (2017). A question of control? Examining the role of control conditions in experimental psychopathology using the example of cognitive bias modification research. *The Spanish Journal of Psychology*, *20*, E54. https://doi.org/10.1017/sjp.2017.41

Boendermaker, W. J., Prins, P. J. M., & Wiers, R. W. (2015). Cognitive bias modification for adolescents with substance use problems—Can serious games help? *Journal of Behavior Therapy and Experimental Psychiatry*, *49*, 13–20. https://doi.org/10.1016/j.jbtep.2015.03.008

Bohne, A., Keuthen, N. J., Wilhelm, S., Deckersbach, T., & Jenike, M. A. (2002). Prevalence of symptoms of body dysmorphic disorder and its correlates: A cross-cultural comparison. *Psychosomatics*, *43*, 486–490. https://doi.org/10.1176/appi.psy.43.6.486

Brechan, I., & Kvalem, I. L. (2015). Relationship between body dissatisfaction and disordered eating: Mediating role of self-esteem and depression. *Eating Behaviors*, *17*, 49–58. https://doi.org/10.1016/j.eatbeh.2014.12.008

Brettschneider, M., Neumann, P., Berger, T., Renneberg, B., & Boettcher, J. (2015). Internet-based interpretation bias modification for social anxiety: A pilot study. *Journal of Behavior Therapy and Experimental Psychiatry*, *49*, 21–29. https://doi.org/10.1016/j.jbtep.2015.04.008

Buhlmann, U., Glaesmer, H., Mewes, R., Fama, J. M., Wilhelm, S., Brähler, E., & Rief, W. (2010). Updates on the prevalence of body dysmorphic disorder: A population-based survey. *Psychiatry Research*, *178*, 171–175. https://doi.org/10.1016/j.psychres.2009.05.002

Buhlmann, U., Teachman, B. A., & Kathmann, N. (2011). Evaluating implicit attractiveness beliefs in body dysmorphic disorder using the Go/No-go Association Task. *Journal of Behavior Therapy and Experimental Psychiatry*, *42*, 192–197. https://doi.org/10.1016/j.jbtep.2010.10.003

Buhlmann, U., Wilhelm, S., Glaesmer, H., Brähler, E., & Rief, W. (2009). Fragebogen körperdysmorpher Symptome (FKS): Ein Screening-instrument [Body Dysmorphic Symptoms Inventory: A screening instrument]. *Verhaltenstherapie*, *19*, 237–242. https://doi.org/10.1159/000246278

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*, 1–28.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*. https://doi.org/10.18637/jss.v076.i01

Cash, T. F. (2000). *MBSRQ users' manual* (3rd Rev. ed.). Retrieved from www.body-images.com

Cash, T. F. (2011). Cognitive-behavioral perspectives on body image. In *Body image: A handbook of science, practice, and prevention* (2nd ed., pp. 39–47). New York, NY: Guilford Press.

Cash, T. F., & Fleming, E. C. (2002). The impact of body image experiences: Development of the body image quality of life inventory. *International Journal of Eating Disorders*, *31*, 455–460. https://doi.org/10.1002/eat.10033

Cash, T. F., & Labarge, A. S. (1996). Development of the appearance schemas inventory: A new cognitive body-image assessment. *Cognitive Therapy and Research*, *20*, 37–50. https://doi.org/10.1007/BF02229242

Clement, U., & Löwe, B. (1996). Validation of the FKB-20 as scale for the detection of body image distortions in psychosomatic patients. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, *46*, 254–259.

Dandurand, F., Shultz, T. R., & Onishi, K. H. (2008). Comparing online and lab methods in a problem-solving experiment. *Behavior Research Methods*, *40*, 428–434. https://doi.org/10.3758/BRM.40.2.428

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, *31*, 73–86. https://doi.org/10.1016/S0005-7916(00)00012-4

Dietel, F. A., Möbius, M., Steinbach, L., Dusend, C., Wilhelm, S., & Buhlmann, U. (2018). Effects of induced appearance-related interpretation bias: A test of the cognitive-behavioral model of body dysmorphic disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *61*, 180–187. https://doi.org/10.1016/j.jbtep.2018.07.003

Dietel, F. A., Möllmann, A., Bürkner, P., Wilhelm, S., & Buhlmann, U. (2019). Interpretation bias across body dysmorphic, social anxiety and generalized anxiety disorder—A multilevel, diffusion model account. *PsyArXiv*. https://doi.org/10.31234/osf.io/3hnpu

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment. *Behaviour Research and Therapy*, *41*, 509–528. https://doi.org/10.1016/S0005-7967(02)00088-8

Fang, A., & Wilhelm, S. (2015). Clinical features, cognitive biases, and treatment of body dysmorphic disorder. *Annual Review of Clinical Psychology*, *11*, 187–212. https://doi.org/10.1146/annurev-clinpsy-032814-112849

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

Frederick, D. A., Peplau, L. A., & Lever, J. (2006). The swimsuit issue: Correlates of body image in a sample of 52,677 heterosexual adults. *Body Image*, *3*, 413–419. https://doi.org/10.1016/j.bodyim.2006.08.002

Gentile, A. J., Lima, C. L., Flygare, O., Enander, J., Wilhelm, S., Mataix-Cols, D., & Rück, C. (2019). Internet-based, therapist-guided, cognitive–behavioural therapy for body dysmorphic disorder with global eligibility for inclusion: An uncontrolled pilot study. *BMJ Open*, *9*, e024693. https://doi.org/10.1136/bmjopen-2018-024693

Gonsalves, M., Whittles, R. L., Weisberg, R. B., & Beard, C. (2019). A systematic review of the word sentence association paradigm (WSAP). *Journal of Behavior Therapy and Experimental Psychiatry*, *64*, 133–148. https://doi.org/10.1016/j.jbtep.2019.04.003

Gräfe, K., Zipfel, S., Herzog, W., & Löwe, B. (2004). Screening psychischer Störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D)" [Screening for psychiatric disorders with the Patient Health Questionnaire (PHQ): Results from the German validation study]. *Diagnostica*, *50*(4), 171–181. https://doi.org/10.1026/0012-1924.50.4.171

Grocholewski, A., Tuschen-Caffier, B., Margraf, J., & Heinrichs, N. (2011). Überzeugungen über das Erscheinungsbild: Eine Fragebogenvalidierung [Beliefs about appearance. Validation of a questionnaire]. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *40*, 85–93. https://doi.org/10.1026/1616-3443/a000078

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin*, *137*, 940–958. https://doi.org/10.1037/a0024355

Harrison, A., Fernández de la Cruz, L., Enander, J., Radua, J., & Mataix-Cols, D. (2016). Cognitive-behavioral therapy for body dysmorphic disorder: A systematic review and meta-analysis of randomized controlled trials. *Clinical Psychology Review*, *48*, 43–51. https://doi.org/10.1016/j.cpr.2016.05.007

Heinicke, B. E., Paxton, S. J., McLean, S. A., & Wertheim, E. H. (2007). Internet-delivered targeted group intervention for body dissatisfaction and disordered eating in adolescent girls: A randomized controlled trial. *Journal of Abnormal Child Psychology*, *35*, 379–391. https://doi.org/10.1007/s10802-006-9097-9

Hilbig, B. E. (2016). Reaction time effects in lab-versus web-based research: Experimental evidence. *Behavior Research Methods*, *48*, 1718–1724. https://doi.org/10.3758/s13428-015-0678-9

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods*, *39*, 101–117. https://doi.org/10.3758/BF03192848

Hoppitt, L., Illingworth, J. L., MacLeod, C., Hampshire, A., Dunn, B. D., & Mackintosh, B. (2014). Modifying social anxiety related to a real-life stressor using online cognitive bias modification for interpretation. *Behaviour Research and Therapy*, *52*, 45–52. https://doi.org/10.1016/j.brat.2013.10.008

Hudson, J. I., Hiripi, E., Pope, H. G., Jr., & Kessler, R. C. (2007). The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biological Psychiatry*, *61*, 348–358. https://doi.org/10.1016/j.biopsych.2006.03.040

Jones, E. B., & Sharpe, L. (2017). Cognitive bias modification: A review of meta-analyses. *Journal of Affective Disorders*, *223*, 175–183. https://doi.org/10.1016/j.jad.2017.07.034

Korn, J., Dietel, F. A., & Hartmann, A. S. (2019). An experimental study on the induction of an eating disorder-specific interpretation bias in healthy individuals: Testing the interpretation modification paradigm for eating disorders (IMP-ED). *Cognitive Therapy and Research*, *43*, 1097–1108. https://doi.org/10.1007/s10608-019-10026-x

Koster, E. H. W., Fox, E., & MacLeod, C. (2009). Introduction to the special section on cognitive bias modification in emotional disorders. *Journal of Abnormal Psychology*, *118*, 1–4. https://doi.org/10.1037/a0014379

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals*, *32*, 509–515. https://doi.org/10.3928/0048-5713-20020901-06

Kuckertz, J. M., & Amir, N. (2017). Cognitive bias modification. In S. G. Hofmann & G. J. G. Asmundson (Eds.), *The Science of Cognitive Behavioral Therapy* (pp. 463–491) Cambridge, MA: Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-803457-6.00019-2

Linardon, J., Wade, T. D., de la Piedad Garcia, X., & Brennan, L. (2017). The efficacy of cognitive-behavioral therapy for eating disorders: A

systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology*, *85*, 1080–1094. https://doi.org/10.1037/ccp0000245

MacLeod, C. (2012). Cognitive bias modification procedures in the management of mental disorders. *Current Opinion in Psychiatry*, *25*, 114–120. https://doi.org/10.1097/YCO.0b013e32834fda4a

Matheson, E., Wade, T. D., & Yiend, J. (2018). A new cognitive bias modification technique to influence risk factors for eating disorders. *International Journal of Eating Disorders*, *51*, 959–967. https://doi.org/10.1002/eat.22938

Matheson, E., Wade, T. D., & Yiend, J. (2019). Utilising cognitive bias modification to remedy appearance and self-worth biases in eating disorder psychopathology: A systematic review. *Journal of Behavior Therapy and Experimental Psychiatry*, *65*, 101482. https://doi.org/10.1016/j.jbtep.2019.101482

Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual Review of Clinical Psychology*, *1*, 167–195. https://doi.org/10.1146/annurev.clinpsy.1.102803.143916

Melville, K. M., Casey, L. M., & Kavanagh, D. J. (2010). Dropout from internet-based treatment for psychological disorders. *British Journal of Clinical Psychology*, *49*, 455–471. https://doi.org/10.1348/014466509X472138

Menne-Lothmann, C., Viechtbauer, W., Höhn, P., Kasanova, Z., Haller, S. P., Drukker, M., … Lau, J. Y. F. (2014). How to boost positive interpretations? A meta-analysis of the effectiveness of cognitive bias modification for interpretation. *PLoS One*, *9*, e100925. https://doi.org/10.1371/journal.pone.0100925

Möbius, M., Tendolkar, I., Lohner, V., Baltussen, M., & Becker, E. S. (2015). Refilling the half-empty glass – Investigating the potential role of the interpretation modification paradigm for depression (IMP-D). *Journal of Behavior Therapy and Experimental Psychiatry*, *49*, 37–43. https://doi.org/10.1016/j.jbtep.2015.03.002

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gotzsche, P. C., Devereaux, P. J., … Altman, D. G. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMJ (Clinical Research Ed)*, *340*, c332. https://doi.org/10.1136/bmj.c332

Möllmann, A., Dietel, F. A., Hunger, A., & Buhlmann, U. (2017). Prevalence of body dysmorphic disorder and associated features in German adolescents: A self-report survey. *Psychiatry Research*, *254*, 263–267. https://doi.org/10.1016/j.psychres.2017.04.063

Mond, J., Mitchison, D., Latner, J., Hay, P., Owen, C., & Rodgers, B. (2013). Quality of life impairment associated with body dissatisfaction in a general population sample of women. *BMC Public Health*, *13*, 920. https://doi.org/10.1186/1471-2458-13-920

Palmqvist, B., Carlbring, P., & Andersson, G. (2007). Internet-delivered treatments with or without therapist input: Does the therapist factor have implications for efficacy and cost? *Expert Review of Pharmacoeconomics and Outcomes Research*, *7*, 291–297. https://doi.org/10.1586/14737167.7.3.291

Paxton, S. J., Neumark-Sztainer, D., Hannan, P. J., & Eisenberg, M. E. (2006). Body dissatisfaction prospectively predicts depressive mood and low self-esteem in adolescent girls and boys. *Journal of Clinical Child and Adolescent Psychology*, *35*, 539–549. https://doi.org/10.1207/s15374424jccp3504_5

Premo, J. E., Sarfan, L. D., & Clerkin, E. M. (2016). Training interpretation biases among individuals with body dysmorphic disorder symptoms. *Body Image*, *16*, 54–62. https://doi.org/10.1016/j.bodyim.2015.11.004

Quené, H., & Van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication*, *43*, 103–121. https://doi.org/10.1016/j.specom.2004.02.004

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rodgers, R. F., & DuBois, R. H. (2016). Cognitive biases to appearance-related stimuli in body dissatisfaction: A systematic review. *Clinical Psychology Review*, *46*, 1–11. https://doi.org/10.1016/j.cpr.2016.04.006

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Salemink, E., Kindt, M., Rienties, H., & van den Hout, M. (2014). Internet-based cognitive bias modification of interpretations in patients with anxiety disorders: A randomised controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry*, *45*, 186–195. https://doi.org/10.1016/j.jbtep.2013.10.005

Schmidt, J., Lamprecht, F., & Wittmann, W. W. (1989). Zufriedenheit mit der stationären Versorgung. Entwicklung eines Fragebogens und erste Validitätsuntersuchungen [Satisfaction with inpatient care: Development of a questionnaire and first validity assessments]. *Psychotherapie Psychosomatik Medizinische Psychologie*, 39, 248–255.

Stice, E., & Shaw, H. E. (2002). Role of body dissatisfaction in the onset and maintenance of eating pathology: A synthesis of research findings. *Journal of Psychosomatic Research*, *53*, 985–993. https://doi.org/10.1016/S0022-3999(02)00488-9

Summers, B. J., & Cougle, J. R. (2016). Modifying interpretation biases in body dysmorphic disorder: Evaluation of a brief computerized treatment. *Behaviour Research and Therapy*, *87*, 117–127. https://doi.org/10.1016/j.brat.2016.09.005

Summers, B. J., & Cougle, J. R. (2018). Effects of an appearance-focused interpretation training intervention on eating disorder symptoms. *Behavioural and Cognitive Psychotherapy*, *46*, 676–689. https://doi.org/10.1017/S1352465818000164

von Collani, G., & Herzberg, P. Y. (2003). Eine revidierte Fassung der deutschsprachigen Skala zum Selbstwetgefühl von Rosenberg [A Revised Version of the German Adaptation of Rosenberg's Self-Esteem Scale]. *Zeitschrift Für Differentielle Und Diagnostische Psychologie*, *24*, 3–7. https://doi.org/10.1024//0170-1789.24.1.3

von Glischinski, M., Teismann, T., Prinz, S., Gebauer, J. E., & Hirschfeld, G. (2016). Depressive symptom inventory suicidality subscale: Optimal cut points for clinical and non-clinical samples. *Clinical Psychology & Psychotherapy*, *23*, 543–549. https://doi.org/10.1002/cpp.2007

Vossbeck-Elsebusch, A. N., Waldorf, M., Legenbauer, T., Bauer, A., Cordes, M., & Vocks, S. (2014). German version of the Multidimensional Body-Self Relations Questionnaire–Appearance Scales (MBSRQ-AS): Confirmatory factor analysis and validation. *Body Image*, *11*, 191–200. https://doi.org/10.1016/j.bodyim.2014.02.002

Wilhelm, S., Phillips, K. A., & Steketee, G. (2013). *Cognitive behavioral therapy for body dysmorphic disorder—A treatment manual*. New York, NY: Guilford.

Wilver, N. L., & Cougle, J. R. (2019). An internet-based controlled trial of interpretation bias modification versus progressive muscle relaxation for body dysmorphic disorder. *Journal of Consulting and Clinical Psychology*, *87*(3), 257–269. https://doi.org/10.1037/ccp0000372

How to cite this article: Dietel FA, Zache C, Bürkner P-C, et al. Internet-based interpretation bias modification for body dissatisfaction: A three-armed randomized controlled trial. *Int J Eat Disord*. 2020;53:972–986. https://doi.org/10.1002/eat.23280

APPENDIX

TABLE A1 Description of stressor videos

Video title	Description	Duration
The conversation	Four people are having a conversation. While a fifth person is approaching (POV), people stop talking and look at POV with a critical facial expression	00:22
The group photo	A woman is taking a photo of three other people (one of them is POV). The woman is then taking a second photo of the POV only and is looking at it with a critical facial expression	00:27
The presentation	POV is starting to give a presentation. Six people are sitting in the same room. They start whispering and giggling during the presentation	00:25
The hairdresser	POV is entering a hairdressing salon. The hairdresser is welcoming POV while eyeing POV up and down.	00:23
The coffee shop	A woman is entering a coffee shop and approaches POV. While looking at POV, the woman turns away	00:20
The elevator	The elevator door opens, and a man is entering the elevator. He directly looks at the other person in the elevator (POV)	00:25
The job interview	POV is knocking on a door and is entering an office. A woman in a business outfit shakes hands with POV and offers a seat. The woman starts talking about the application while frowning	00:27

Abbreviation: POV, point-of-view (first person) perspective within video.