2018

What Process Works For Whom: Individual Differences And The Impact Of Therapy Techniques And Treatment Mechanisms

John Raymond Keefe
University of Pennsylvania, jack.keefe@gmail.com

Follow this and additional works at: https://repository.upenn.edu/edissertations

Part of the Clinical Psychology Commons

Recommended Citation

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/edissertations/3432
For more information, please contact repository@pobox.upenn.edu.
What Process Works For Whom: Individual Differences And The Impact Of Therapy Techniques And Treatment Mechanisms

Abstract
While psychotherapy treatment manuals define the broad structure and targets of a therapy, therapists must decide how to implement treatments with a specific patient. Yet, patients are heterogeneous even within a disorder class, and there is little systematic research to guide a therapist to make principled adaptations. We examined the question of whether individual differences moderate the treatment effects of therapist interventions and both in-session and between-session processes of change, using data from a randomized controlled trial for panic disorder comparing panic-focused psychodynamic psychotherapy (PFPP) to cognitive-behavioral therapy. In Chapter 1, adherence to PFPP (n = 65) was observer rated in Sessions 2 and 10 to predict panic change after the rated session. Panic-specific interpretations predicted improvements, while non-panic-focused interventions did not. Concordant with dynamic theory, patients with more interpersonal problems benefitted especially from heightened focus on the interplay between interpersonal-emotional conflict and panic. In Chapter 2, we examined whether higher levels of observer-rated emotional expression—a marker of therapeutic engagement—across early PFPP sessions predicted subsequent panic improvements. We hypothesized that this relationship would be moderated by certain personality disorder traits related to emotionality: (1) borderline traits, which denote heightened, labile, dysregulated emotionality; and (2) obsessive-compulsive traits, related to muted, constrained emotionality. As predicted, borderline traits attenuated the otherwise positive relationship between emotional expression and symptom improvement, but obsessive-compulsive traits had inconsistent relationships between trial sites. Finally, in Chapter 3, we built on mediational analyses for both CBT and PFPP (n = 138), examining whether the presence of different psychological vulnerabilities moderated the symptomatic impact of improvement in two mediators of panic change: catastrophic, body-focused interpretation style, and panic-specific reflective functioning (PSRF). Patients beginning treatment with a more catastrophic style benefitted more from improvements in either mediator. Personality disorder traits blunted the impact of improvements in PSRF, but patients with no personality disorder evidenced high benefit from PSRF improvements. This set of findings suggests that personalization of psychotherapy can be empirically grounded. Clinical characteristics may inform how a therapist chooses to intervene, what in-session processes they focus on, and what types of psychological changes they aim to encourage.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Psychology

First Advisor
Robert J. DeRubeis

Second Advisor
Dianne L. Chambless

Keywords
mediation, personalized medicine, psychotherapy, psychotherapy process
WHAT PROCESS WORKS FOR WHOM: INDIVIDUAL DIFFERENCES AND THE
IMPACT OF THERAPY TECHNIQUES AND TREATMENT MECHANISMS

John Raymond Keefe

A DISSERTATION

in

Psychology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

2019

Supervisor of Dissertation

_______________________________

Robert J. DeRubeis, PhD, Samuel H. Preston Term Professor in the Social Sciences

Graduate Group Chairperson

_______________________________

John Trueswell, PhD, Professor of Psychology

Dissertation Committee

Dianne L. Chambless, PhD, Professor of Psychology

Sara R. Jaffee, PhD, Professor of Psychology
ACKNOWLEDGMENT

I must begin my acknowledgments by recognizing the incomparable contribution and influence of my dissertation advisor Robert DeRubeis on my career and my person. Like a securely attached parent, Rob steadfastly encouraged me to go out on my own intellectually and academically, but he was always there with Dutch M&Ms, espresso, and counsel when I got lost along the way. I greatly admire Rob’s understated inquisitiveness and flexibility, and his commitment to long-spanning, ambling intellectual dialogue, which deepened my thinking and work. If I ever have students of my own, I hope to emulate Rob.

I owe a special debt to Jacques Barber, who by recommendation of my undergraduate abnormal psychology professor Michelle Reimer adopted a young post-bac neuroscientist and retrained him as a clinical psychologist at the Center for Psychotherapy Research. Jacques, along with his then-postdoctoral fellow Kevin McCarthy, gave generous, warm, sagacious mentorship and great opportunities that made it even possible for me to apply to graduate school.

My dissertation chair Dianne Chambless modeled for me impeccable grace as a person and precision in thinking and research design, and I am a better person and researcher from knowing her. I also must thank Barbara Milrod for bringing me into the fold of her research team in New York, for downloading into me her gifted clinical acumen, for her sharp, direct, but always on point critique of my writing and ideas, and for her almost unbelievable faith in me. I also thank the great professors I have learned from at Penn, especially Sara Jaffee and Geoff Goodwin.
The extended research families of Rob and Jacques have served as amazing intellectual interlocutors and social raconteurs during my degree. I am particularly grateful for my fellow graduate students in the DeRubeis lab—Lorenzo Lorenzo-Luaces, Zachary Cohen, Colin Xu, and Thomas Kim. Ramaris German should be particularly thanked as a surrogate parent on behalf of Rob when he was out of the country my first semester at Penn. I have sincerely appreciated how Jacques has kept me connected to his active lab group, and particularly for the strong personal and professional relationships I formed with Kevin McCarthy, Sigal Zilcha-Mano, Ulrike Dinger, and Nili Solomonov.

I also must give unique thanks to my clinical supervisors at Penn, whose insights into my patients and my clinical work shaped not only my clinical skills but also my research: Sydney Pulver, Alan Goldstein, Rob DeRubeis (a twofer), and Melissa Hunt. I also have to thank the support staff at Penn, who have organized me and tolerated my bizarre, last-minute requests and queries with warmth, competence, and humor: Yuni Thornton, Joan Massimo, Pat Spann, Claire Ingulli, and Paul Newlon.

I also wish to express gratitude to past teachers who have shaped me: Allen Schneider, my undergraduate mentor in psychobiology; Sara Hiebert-Burch, who unwittingly talked me out of quitting the Honors Program at Swarthmore; and Margie Craig, my high-school journalism mentor.

To get through a doctorate a whole person, it is important to have a “life worth living” (as Marsha Linehan emphasized), and for making that possible I am in profound gratitude for the many people who fill my life with joy: my erstwhile friends back home in Virginia (Justin Belcher, Jessica Johnson, Caroline Mayen, Kelsey Sowers, Michael Terlecki); my friends in Philly (Hannah Blake, Jen Brown, Celia Caust-Ellenbogen, Lisa
Draco, Lori Felipe-Barkin, Lindsay Friedman, Elizabeth Jenkins, Elizabeth Laudenslager, Christiane Moore, Aaron Stock, Jessica Swaysland, Joshua Warner, Zachary Wiener); my Dungeons and Dragons group (!) (Mark Lewis, Lauren Lewis, Jasmine Erdener, Danielle Hanley); my friends in the diaspora (Toby Altman, Aaron Brecher, Suzanne van Bronswijk, Chelsea Davis, Jess Deutsch, Jess Engebretson, Alex Ho, Lotte Lemmens, John Maltempo, Charles Rabideau, James Robinson, Melvin Pelaez, Smith Koy Smith, Lauren Stokes, Kate Swisher, Laura Wang, Julia Wittes); and my friends here in the program (especially Kelly Allred, Corey Cusimano, Eliora Porter, and Rachel Schwartz).

I also must thank my loving family for everything they have done: my mother Joan Pease, who instilled in me great determination and whose optimism and pride sustain me, and my father Richard Keefe, who taught me to be gentle, curious, and odd; my Italian family, my dear cousins Chiara and Stephanie Ferraro and my aunt Eva Anne Ferraro who taught me kindness; my endlessly supportive maternal grandmother Madeline Pease; and my paternal grandmother Eloise Keefe, who did not live to see me enter graduate school but who is single-handedly the person most responsible for making me the psychologist I am today.

Finally, I would also like to thank the many, many cats who provided me unconditional positive regard during my doctoral process: Lenny & Belli at Fitzwater; and the full rescue litter my parents took care of—Momma, Willie, Rosie, Thelma, Louise, Right-Eye, Scooter, Blue, Bear, Princess, Abe, Mr. Two, Lily, and Bo.
ABSTRACT

WHAT PROCESS WORKS FOR WHOM: INDIVIDUAL DIFFERENCES AND THE IMPACT OF THERAPY TECHNIQUES AND TREATMENT MECHANISMS

John Raymond Keefe
Robert J. DeRubeis

While psychotherapy treatment manuals define the broad structure and targets of a therapy, therapists must decide how to implement treatments with a specific patient. Yet, patients are heterogeneous even within a disorder class, and there is little systematic research to guide a therapist to make principled adaptations. We examined the question of whether individual differences moderate the treatment effects of therapist interventions and both in-session and between-session processes of change, using data from a randomized controlled trial for panic disorder comparing panic-focused psychodynamic psychotherapy (PFPP) to cognitive-behavioral therapy. In Chapter 1, adherence to PFPP (n = 65) was observer rated in Sessions 2 and 10 to predict panic change after the rated session. Panic-specific interpretations predicted improvements, while non-panic-focused interventions did not. Concordant with dynamic theory, patients with more interpersonal problems benefitted especially from heightened focus on the interplay between interpersonal-emotional conflict and panic. In Chapter 2, we examined whether higher levels of observer-rated emotional expression—a marker of therapeutic engagement—across early PFPP sessions predicted subsequent panic improvements. We hypothesized that this relationship would be moderated by certain personality disorder traits related to emotionality: (1) borderline traits, which denote heightened, labile, dysregulated emotionality; and (2) obsessive-compulsive traits, related to muted, constrained
emotionality. As predicted, borderline traits attenuated the otherwise positive relationship between emotional expression and symptom improvement, but obsessive-compulsive traits had inconsistent relationships between trial sites. Finally, in Chapter 3, we built on mediational analyses for both CBT and PFPP \((n = 138)\), examining whether the presence of different psychological vulnerabilities moderated the symptomatic impact of improvement in two mediators of panic change: catastrophic, body-focused interpretation style, and panic-specific reflective functioning (PSRF). Patients beginning treatment with a more catastrophic style benefitted more from improvements in either mediator. Personality disorder traits blunted the impact of improvements in PSRF, but patients with no personality disorder evidenced high benefit from PSRF improvements. This set of findings suggests that personalization of psychotherapy can be empirically grounded. Clinical characteristics may inform how a therapist chooses to intervene, what in-session processes they focus on, and what types of psychological changes they aim to encourage.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENT</td>
<td>IV</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>VIII</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>X</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>XI</td>
</tr>
<tr>
<td>SPECIFIC PSYCHODYNAMIC TECHNIQUES IN PANIC TREATMENT</td>
<td>1</td>
</tr>
<tr>
<td>PATIENT PERSONALITY AND EMOTIONAL PROCESS</td>
<td>34</td>
</tr>
<tr>
<td>WHAT MECHANISM WORKS FOR WHOM IN PANIC</td>
<td>79</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>109</td>
</tr>
<tr>
<td>Table</td>
<td>Title</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.1</td>
<td>Descriptive data for baseline characteristics and for symptom change, in Study 1.</td>
</tr>
<tr>
<td>1.2</td>
<td>Descriptive statistics for technique process measurements (average 15 minutes per session), in Study 1.</td>
</tr>
<tr>
<td>1.3</td>
<td>Correlations between technique process measurements and baseline characteristics, in Study 1.</td>
</tr>
<tr>
<td>2.1</td>
<td>Descriptive data for baseline characteristics and for symptom change, in Study 2.</td>
</tr>
<tr>
<td>2.2</td>
<td>Example transcripts from sessions rated high/low in emotional expression, in Study 2.</td>
</tr>
<tr>
<td>3.1</td>
<td>Descriptive data for baseline characteristics and for symptom change, in Study 3.</td>
</tr>
<tr>
<td>S1</td>
<td>Examples of panic-focused clarification and interpretation, for Study 1 in Appendix C.</td>
</tr>
<tr>
<td>S2</td>
<td>Correlations between emotional expression ratings within the same 5-minute segment, for Study 2 in Appendix D.</td>
</tr>
<tr>
<td>S3</td>
<td>Average levels of emotional expression per 5-minute segment across sampled sessions, for Study 2 in Appendix D.</td>
</tr>
<tr>
<td>S4</td>
<td>Sources of variance in emotional expression per 5-minute segment, for Study 2 in Appendix D.</td>
</tr>
</tbody>
</table>
## LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Estimated effect sizes for the relationship between psychodynamic technique use at a given session and subsequent improvement in panic symptoms, in Study 1.</td>
<td>34</td>
</tr>
<tr>
<td>1.2</td>
<td>Estimated change in panic symptoms between Weeks 5 to 9, as a function of degree of interpersonal problems and their interaction with use of panic-focused interpretations at session 10, in Study 1.</td>
<td>35</td>
</tr>
<tr>
<td>2.1</td>
<td>Estimated effect sizes for the relationship between the average level of emotional expression and subsequent improvements in panic symptoms, in Study 2.</td>
<td>79</td>
</tr>
<tr>
<td>2.2</td>
<td>Estimated change in panic symptoms as a function of overall emotional expression in interaction with borderline personality disorder criteria, in Study 2.</td>
<td>80</td>
</tr>
<tr>
<td>3.1</td>
<td>Predicted degree of later symptom change attributable to early improvements in PSRF, as a function of baseline severity of catastrophic interpretive style, in Study 3.</td>
<td>109</td>
</tr>
<tr>
<td>3.2</td>
<td>Predicted degree of later symptom change attributable to early improvements in PSRF as a function of baseline personality disorder pathology, in Study 3.</td>
<td>110</td>
</tr>
</tbody>
</table>
Specific Psychodynamic Techniques in Panic Treatment

Panic-focused psychodynamic psychotherapy (PFPP; Busch, Milrod, Singer, & Aronson, 2012) is a 24-session, twice weekly brief psychodynamic psychotherapy (PDT) formulated for the treatment of panic disorder (PD) with and without agoraphobia. In several randomized controlled trials (RCTs), PFPP has been shown to be equivalent in efficacy to various forms of cognitive-behavioral psychotherapy (CBT; Beutel et al., 2013; Keefe, McCarthy, Dinger, Zilcha-Mano, & Barber, 2014; Milrod et al., 2016; Milrod et al., 2007). The exception comes from one site of a two-site study at which PFPP was inferior to CBT (Milrod et al., 2016). Further trials are underway (Sandell et al., 2015).

The PFPP treatment model is based on the assumption that the acute emergence of panic and the developmental vulnerability toward PD has underlying psychological meanings related to emotional-interpersonal conflicts and attachment dysregulation (Busch, Milrod, & Singer, 1999; Busch et al., 2012). For example, a patient who experiences ambivalence regarding a life-long romantic commitment to his partner may develop a PD shortly after announcing their engagement. The patient’s panic attacks may be triggered by anxious moments of unacknowledged intolerable ambivalence toward his partner. He may either be unaware of or perhaps frightened to recognize this ambivalence and its potentially frightening consequences, such as breaking up with or losing the support of his partner, which could result in a vicious cycle of unrecognized conflicted feelings and anxious arousal leading to panic. Consistent with PFPP’s etiological hypothesis, patients with PD report higher rates of alexithymia (i.e., difficulty verbalizing one’s emotions and motives), experiential avoidance, and lack of emotional acceptance,
compared to psychiatrically healthy controls and persons with simple phobias (Galderisi et al., 2008; Izci et al., 2014; Parker, Taylor, Bagby, & Acklin, 1993; Tull & Roemer, 2007).

One important intervention in PFPP, clarification, is the attempt by the therapist to help the patient become aware of the avoided and unconscious intrapsychic conflicts that give rise to panic (Busch et al., 2012). This process entails gathering information concerning the potential meanings of symptoms and actively helping the patient recognize, verbalize, and reflect on those meanings; the context of symptoms also helps to identify meanings. Successful clarifications along these lines lay the groundwork for accurate interpretations of conflicts. For example, clarification could help the aforementioned patient become aware that he habitually assigns frightful somatic sources (e.g., heart disease) to anxiety that emerges when he became engaged and works on wedding details, triggering intensely mixed feelings about his fiancée. This would enable the therapist to help the patient focus on specific causes for these feelings.

Another important PFPP intervention is interpretation—attempts to help the patient identify the specific dynamics and conflicts that underlie PD (Busch et al., 2012; Summers & Barber, 2010). Types of interpretations include: (a) defense (i.e., of a particular manner in which the patient avoids experiencing a particular distressing feeling or issue); (b) dynamic/conflict (i.e., of how a patient’s experiences are the result of a conflict between unacceptable wishes and the defenses against these wishes); (c) genetic (i.e., of how early, formative attachment relationships may have made the patient’s approach to interpersonal situations fraught with specific, identifiable vulnerabilities); and (d) transference (i.e., how the patient’s recurring underlying formative attachment
patterns and conflicts emerge in the relationship with the therapist) (Busch et al., 2012). PFPP therapists should attempt to link, whenever possible, the specifics of the patient’s panic symptoms to his/her specific underlying dynamics (Busch et al., 1999; Busch et al., 2012). A PFPP therapist who treats the engaged PD patient described above might interpret that he appears to be afraid of experiencing and expressing his anger with his partner. The therapist might cite the observation that the patient frequently talks about how much he loves his partner immediately after expressing his frustrations, thereby magically “undoing” his anger. The therapist could also point out that the unresolved and frightening feelings of anger and frustration seem to emerge as sensations of physical discomfort that trigger panic. By doing so, the therapist attempts to help the patient own and work through these conflicts, diminishing their power as panic triggers and helping the patient attribute any lingering anxiety to personal psychological meanings rather than, for example, somatic problems.

**Techniques Contributing to Efficacy in PDT for Anxiety**

Previous studies have demonstrated that use of specific psychodynamic techniques is predictive of outcomes in depression and personality disorder treatment (Barber, Crits-Christoph, & Luborsky, 1996; Barber, Muran, McCarthy, & Keefe, 2013; Hoglend, Dahl, Hersoug, Lorentzen, & Perry, 2011; Levy et al., 2006; McCarthy, Keefe, & Barber, 2016). However, there is a lack of empirical evidence supporting the efficacy of specific psychodynamic therapy (PDT) techniques for anxiety disorders, with only one small sampled study (n = 20) showing that patients whose therapists employed more interpretations across two sessions (3 and 9 out of an average of 20 sessions) of short-term PDT tended to have superior symptomatic outcomes at termination (Pitman, Slavin-
Mulford, & Hilsenroth, 2014). However, because the authors were unable to establish temporal precedence (i.e., that technique use preceded symptom change), a plausible account of the finding is that, for instance, patients with better prognoses pulled for more interpretations from their therapists in this setting. Moreover, disorders were not diagnosed with a reliable measure. Further research is clearly required to investigate the efficacy of specific PDT techniques for anxiety.

When considering the active ingredients that characterize effective PFPP, it should be noted PFPP is distinguished from generic short-term PDT through its explicit emphasis on panic and its associated dynamics (Busch et al., 2012). Clinical trials in which PDTs have been operationalized without a specific focus on the primary symptoms, or in which such a focus was proscribed or discouraged, have generated some of the most disappointing findings regarding the efficacy of PDTs (Durham et al., 1994; Garner et al., 1993; Gilboa-Schechtman et al., 2010; Poulsen et al., 2014). The rationale behind PFPP is that by focusing on experiences proximal to panic and by linking interpretations coherently to panic vulnerability, patients gain insight as to the specific underpinnings of their panic attacks (Rudden, Milrod, Target, Ackerman, & Graf, 2006). In PDT and other therapies, gains in insight have been found in different investigations to both predict further functional improvements post-treatment as well as protection against relapse across follow-up (Barber et al., 2013; Gibbons et al., 2009; Johansson et al., 2010; Kallestad et al., 2010). However, the specific hypothesis that psychodynamic focus on symptoms promotes greater symptom relief (Summers & Barber, 2010) has never been addressed directly in empirical investigations.
In the context of a randomized controlled trial of PD with and without agoraphobia evaluating manualized PFPP (Milrod et al., 2016), we measured the use of psychodynamic psychotherapy interventions in two sessions of psychotherapy taken from the early and middle phases of a 24-session treatment protocol (Sessions 2 and 10). Based on PFPP’s conceptual model (Busch et al., 2012), we hypothesized that therapists’ more frequent use of psychodynamic interpretations made with connections to symptoms of and/or vulnerabilities to panic, agoraphobia, or anxiety would predict greater subsequent symptom improvement in PFPP. In the PFPP manualization, therapists are expected to, on average, focus on clarification of panic meanings in the beginning of treatment, but to become progressively more interpretive as these meanings become clearer and the patient becomes socialized to psychodynamic treatment (Busch et al., 2012). We hypothesized that interpretations in the session taken from the middle phase of treatment (Session 10) would be more predictive of improvement than interpretations in the earlier session (Session 2), as by the middle phase of treatment the therapist has sufficient information about the patient to make meaningful and accurate interpretations (Andrusyna, Luborsky, Pham, & Tang, 2006; Crits-Christoph, Cooper, & Luborsky, 1988). We anticipated that interpretations made without reference to panic would not predict subsequent symptom improvement, when accounting for panic-focused interpretations.

In addition, we hypothesized that panic-focused clarifications made early on in therapy (Session 2) would contribute to prediction of psychotherapy outcomes, as clarification in this phase can help the therapist gather information and can encourage initial exploration of dynamics underlying the patient’s panic symptoms.
Moderators of technique-outcome relationships. Like many manualized psychotherapies, PFPP allows for therapist flexibility in focus and application of therapeutic techniques (Busch et al., 2012), but relatively little empirical data exist to help guide moment-to-moment judgments of how to respond to an individual patient. As specific therapeutic techniques can be more or less conducive of change among particular patients (Keefe, Webb, & DeRubeis, 2016; Sasso, Strunk, Braun, DeRubeis, & Brotman, 2015), we furthermore hypothesized that specific patients would be likely to evince apparent benefit from particular psychodynamic interventions.

Specifically, we hypothesized that panic-focused interpretations would be more important to outcomes when patients entered the trial with more interpersonal problems as measured by the Inventory of Interpersonal Problems (IIP; Horowitz, Alden, Wiggins, & Pincus, 2000). Interpersonal conflicts and transitions are frequently stressors surrounding the onset of PD (Klass et al., 2009; Scocco, Barbieri, & Frank, 2007), and couples in which one patient has PD often exhibit relational distress and avoidant conversational and cognitive styles (Chambless, 2010). As unresolved and/or unconscious interpersonal conflicts such as intolerable dependency and anger can be viewed as a trigger for panic in PFPP, we hypothesized that individuals with more interpersonal problems might be more likely to have such conflicts or to have relatively more pervasive conflicts, and thus would benefit relatively more from panic-focused interpretations addressing those conflicts. In addition, the dynamics associated with panic, including difficulties with separation, dependency and anger could contribute to both panic and interpersonal problems. In other words, our hypothesis was based on the notion that patients with more interpersonal problems may be more likely to have the
dynamic-interpersonal contributions to panic that panic-focused interpretations may specifically help reveal and resolve.

**Method**

**Participants**

**Patients.** The present study is a secondary analysis of patients randomized to the PFPP condition \((N = 80)\) of a two-site randomized controlled trial comparing PFPP, CBT, and applied relaxation training among patients with primary DSM-IV panic disorder with or without agoraphobia. Treatment took place twice a week for 12 weeks. Patients were recruited at New York Presbyterian/Weill-Cornell (hereafter, “Site A”) and the University of Pennsylvania (hereafter, “Site B”). Participants received study treatment gratis. Participants gave informed written consent. Both sites’ institutional review boards approved the protocol, and the study is registered with ClinicalTrials.gov (identifier: NCT00353470).

Patients were included in the trial if they had the spontaneous occurrence of one or more panic attacks for the month before trial entry, and qualified for DSM-IV panic disorder diagnosis determined as per the ADIS-IV (DiNardo, Brown, & Barlow, 1995). Cross-site agreement on ADIS ratings for panic severity (with a “4” indicating the diagnostic threshold) was excellent \((ICC = 1.00)\). See Milrod et al., 2016 for further details.

Non-study psychotherapy was prohibited. Medications were permitted if stable for at least two months at presentation, and were recorded, held constant, and monitored during the trial. Exclusion criteria were active substance dependence (less than 6 months’ remission), a history of psychosis or bipolar disorder, acute suicidality, and organic
mental syndrome. Additional details on trial design, independent evaluator training, and therapy adherence can be consulted in the primary outcome paper (Milrod et al., 2016).

Of the 80 randomized patients, 65 ($n = 30$ Site A, $n = 35$ Site B) provided the necessary data to be included in the present study (see section on missing data). Descriptive data on patients’ demographics may be found in Table 1.

**Therapists.** Sixteen doctoral-level therapists (11 M.D., 5 Ph.D.) administered PFPP across the two sites. Therapists had an average of 15 years of post-graduate experience ($SD = 8.2$), and an average of 5 years’ experience in some form of time-limited psychodynamic therapy ($SD = 6.3$). The average total caseload was four (median = 3.5, range = 1 to 11) for therapists whose patients were included in the present analyses. All therapists were experienced therapists who were specifically trained in PFPP over the span of a 2-day, 10-hour course. Therapists participated in monthly group supervision and received regular individual supervision from senior clinicians. For the primary outcome paper, basic adherence to PFPP was established (Milrod et al., 2016).

**Outcome Index**

**Panic Disorder Severity Scale (PDSS; Shear et al., 1997).** The PDSS is a widely used diagnosis-based, composite, global rating of panic disorder severity, with acceptable psychometric properties. The PDSS was administered by trained, master’s-level independent evaluators who were uninformed as to treatment condition. Interrater reliability on the PDSS was excellent ($IC[2,1] = 0.95$). The PDSS was administered five times during treatment: at baseline (Week 0), Week 1, Week 5, Week 9, and termination (Week 12).

**Moderator Measures**
Inventory of Interpersonal Problems—Circumplex (IIP; Horowitz et al., 2000)). The IIP-Circumplex is a 64-item self-report measure of maladaptive interpersonal problems that is a shorter version of the full 127-item IIP. The sum score of the IIP reflects an individual’s degree of interpersonal distress. The IIP exhibits adequate internal reliability and 10-week test-retest reliability (Alden, Wiggins, & Pincus, 1990), and exhibited excellent internal reliability in this sample (alpha = 0.95).

Process Measure

Panic-Focused Psychodynamic Psychotherapy Rating Scale (PFPP-RS; Keefe, Phillips, Busch, & Milrod, 2016). The PFPP-RS is an observer-rated scale developed to assess the degree to which therapists used general psychodynamic interpretive techniques, as well as more specific panic-focused psychodynamic techniques. It was developed by the author JRK in conjunction with PFPP developers Fred Busch and Barbara Milrod. The use of each technique was rated on a 5-point Likert scale ranging from 0 (technique not present in section) to 2 (at least one clear example of the technique in section) to 4 (technique applied fully and comprehensively in section). Scores reflected the degree to which a technique was prototypic to the rated segment (i.e., adherence), not whether the raters believed the therapist applied the technique in a particularly apt way (i.e., competence). For items measuring the use of interpretation, a score of 2 would indicate that the therapist made at least one clear, identifiable interpretation. A score of 4 could represent either multiple individual interpretations made within a rating segment, or a single, continually developed interpretation over an extended period of time. Sessions were divided into 15-min segments, with each segment
rated for each technique. A single rating per session was calculated by averaging ratings made across the three 15-min segments.

Scores of the PFPP-RS items were summed to produce the following subscales:
(a) Psychodynamic Interpretations – the degree to which the therapist used any of the four types of interpretations (defense, genetic, dynamic, and transference); and (b) Panic-Focus: the degree to which the therapist focused on panic, agoraphobia, and anxiety symptoms and made connections between these symptoms and panic dynamics, including underlying conflicts. For Panic-Focus items, to score a 2 or higher on any given item, a therapist had to make a clear, unambiguous reference to panic, agoraphobia, or anxiety in relation to the intervention. This subscale includes two items assessing use of clarification with regard to panic symptoms and their personal meaning (PF-Clarification), and three items assessing interpretations that address the emergence of and vulnerability to experiences of panic, agoraphobia, and anxiety (PF-Interpretation). Examples of therapist interventions qualifying as clarifications and interpretations can be found in Supplemental Table 1. To separate panic-focused interpretations from non-panic-focused interpretations, the PF-Interpretation subscore was subtracted from the total Psychodynamic Interpretations score, as PF-Interpretations reflect a subset of Psychodynamic Interpretations. This score will be referred to as Non-Panic-Focused Interpretations.

Video-recorded sessions were rated using the PFPP-RS by six advanced undergraduate psychology majors at the University of Pennsylvania who each received approximately 20 hours of training by the developers of the scale. Two authors who were PFPP developers (FB & BR), and a graduate student author familiar with the model (NS)
provided additional consultation as to the validity and reliability of ratings during training. Training included a review of the rater manual, the PFPP therapy manual (Busch et al., 2012), and rating of several training tapes of PFPP. Training was continued until raters consistently rated single items within a point (+/-) of their graduate student trainer. Additionally, raters met approximately every other week with the study leader JRK to rate a tape collectively and discuss rating challenges and questions.

All available Sessions 2 and 10 were rated for each PFPP patient. Sessions 3 and 9 were rated in cases wherein Session 2 or 10 (respectively) was not available. Sessions were randomly assigned to raters, who were uninformed of the outcome data. Two raters rated each tape, and ratings were averaged across raters.

Random effects ICCs were calculated using variance estimates from an REML mixed model in the R package “lme4” (Bates et al., 2016; Shrout & Fleiss, 1979). Interrater reliability per 15-minute segment was good for all Psychodynamic Interpretations ($ICC[2,2] = 0.80$) and adequate for Panic-Focused Clarification ($ICC[2,2] = 0.71$), Panic-Focused Interpretations ($ICC[2,2] = 0.70$), and the difference score reflecting Non-Panic-Focused Interpretations ($ICC[2,2] = 0.68$).

**Statistical Analyses**

All analyses were conducted using the R statistical programming language (R Core Team, 2017) and run using robust regressions as implemented in the R package “Robustbase” (Maechler et al., 2016). Given the effective sample size ($n = 65$), robust regression was selected over standard regression for its superior properties of robustness against multivariate outliers and deviation from homoscedasticity (Huber & Ronchetti, 2009). A robust regression (a) retains full information on all observations in an initial
estimate of parameters; (b) iteratively determines weights for each observation based on a particular estimator function from this initial estimate, such that points much farther from model predictions in the previous iteration are given lower weight; and (c) recalculates final parameter estimates based on the final weighting when the values of the coefficients converge within a specified tolerance (Koller & Stahel, 2011). Semi-partial correlation effect sizes (sr) were estimated for parameters of interest from linear regressions.

**Missing data.** In the primary trial, a not missing at random (NMAR) pattern of treatment dropout was detected, such that patients with worse PDSS symptom trajectories were more likely to terminate from treatment prematurely (Milrod et al., 2016). When outcomes for treatment noncompleters are imputed in the NMAR context, imputation and other missing data methods can lead to biased estimates and confidence intervals (Graham, 2009). As such, only individuals who provided data up to the Week 9 assessment point were included in our analyses (n = 65; 81.3% of the intention-to-treat sample).

Several trial completers were missing videotapes of one of the two sessions due to technical issues or therapist/research assistant error (n = 27), but not treatment dropout. Process ratings for completers missing a video-recording of a session can be presumed to be missing at random in relation to panic symptom outcomes (Rubin & Little, 2002). This degree of missingness is not considered prohibitive in the missing data literature (White, Royston, & Wood, 2011), and process ratings have been successfully imputed in the past (e.g., Forand et al., 2018; Lorenzo-Luaces et al., 2017). Accordingly, random forest imputation (Stehkoven & Bühlmann, 2011) was used to impute missing data for completers, using all baseline data, in addition to all PDSS scores and termination and
pre-to-post treatment change scores on the Sheehan Disability Scale, Inventory of Interpersonal Problems, and Hamilton Rating Scale for Depression. Furthermore, all rated technique process ratings from observed sessions (the non-missing session and Session 5, which was rated but not used in this manuscript) were included in the imputation model, such that all patients had observed session process during their therapy contributing to the imputation of their missing session ratings. For this data set, a normalized root mean square error of prediction was estimated at 0.28, indicating that imputation accuracy was adequate (Stekhoven & Bühlmann, 2011). Analyses for patients with complete data for a given session were also run and compared to imputed data. No changes in patterns of statistical significance were detected.

Analytic strategy. Two sets of analyses were performed. Within a robust linear regression framework (Koller & Stahel, 2016), technique use at Session 2 (end of Week 1) was used to predict PDSS symptom change between Weeks 1 and 5 of treatment, while technique use at Session 10 (end of Week 5) was used to predict symptom change between Weeks 5 and 9 of treatment. For two reasons, we examined change in the symptom measurement interval following the sampled session rather than the entire remainder of the therapy: (a) This permitted establishment of closer temporal precedence between techniques and outcomes than is often performed in “long reach” studies that sample from an early session to predict change throughout the entire treatment (e.g., Keefe et al., 2016); and (b) as improvement in the trial was linear and technique use following the sampled session may also influence symptom change, using too large a prediction interval may obfuscate the signal of how technique use occurring in the sampled session per se relates to subsequent symptom change.
Three process terms were included in each regression unless otherwise specified: PF-Interpretation, PF-Clarification, and Non-PF-Interpretation. In each analysis, baseline PDSS symptom score and PDSS change prior to the measured session were included as covariates. Prior panic symptom levels were included as a covariate to allay the possibility that patients who were low severity or getting better could have “pulled” for more or fewer techniques, generating an epiphenomenal relationship. Due to a site by treatment interaction reported in the primary outcome paper across the three tested treatments (Milrod et al., 2016), we also examined whether any process measures interacted with site to predict outcomes, and we planned to report any such interactions if they were found at least at trend level ($p < .10$). However, no such interactions with site were detected, suggesting that process relationships were not detectably different across sites.

Furthermore, IIP scores were examined as a moderator of the relationship of Session 10 panic-focused interpretations to subsequent change, tested by specifying an interaction between the two variables. The Johnson-Neyman technique was applied to probe the regions of significance of the interaction (Johnson & Fay, 1950).

Given that we conducted seven statistical tests, we adjusted $p$-values using the Benjamini-Hochberg correction to control for the false discovery rate at an alpha of 0.05 (Benjamini & Hochberg, 1995), employing the core $R$ function “p.adjust.” These are reported as adjusted $p$-values.

In addition, we conducted two secondary, post-hoc statistical checks on the robustness of our obtained findings. In the first, we employed a mixed model to estimate therapist-level variance simultaneously with our model estimates, which did not result in
substantively different conclusions. In the second, we analyzed whether the technique variables related to pre-to-post treatment functional and interpersonal outcomes. Both analyses are reported in Appendix B.

**Results**

**Descriptive Statistics**

The average patient in this study had a baseline PDSS score of 13.9 (range 7 to 20), considered to be in the moderately-ill severity range for patients with comorbid agoraphobia (Furukawa et al., 2009). More than three-quarters of patients ($n = 53$; 81.5%) qualified for a co-morbid DSM-IV diagnosis of agoraphobia. The average patient reported interpersonal problems in the high-normal range of severity (approximately +0.64 SD over the normative mean). Other baseline demographic and clinical information can be found in Table 1.

Technique scores at Sessions 2 and 10, as well as indices of their consistency over time, are presented in Table 2. Reliable within-patient stability between sessions was observed only for panic-focused interpretations. Within a given case, mean levels of panic-focused and non-panic focused interpretations increased from Session 2 to Session 10.

Prior to data analysis, the correlations between every process measurement and baseline PDSS severity and our proposed moderator variable (IIP scores) were examined. As displayed in Table 3, there were no significant correlations between baseline panic and interpersonal problem severity and any of the technique variables at either time point.

**Early Panic Symptom Change (Weeks 1 to 5)**
There was no significant relationship between panic-focused clarifications at Session 2 (end of Week 1) and symptom change between Weeks 1 to 5 ($B = 1.16$ [95% CI: -0.20, 2.52], $SE = 0.68$, $t[59] = 1.71$, $p = 0.092$, adjusted $p = 0.184$, $sr = 0.21$).

Neither panic-focused interpretations ($B = -1.34$ [95% CI: -3.32, 0.64], $SE = 0.99$, $t[59] = -1.35$, $p = 0.181$, adjusted $p = 0.290$, $sr = -0.15$) nor non-panic focused interpretations ($B = -0.60$ [95% CI: -2.88, 1.68], $SE = 1.14$, $t = -0.53$, $p = 0.599$, adjusted $p = 0.599$, $sr = -0.09$) yielded significant predictions of symptom change in the subsequent measurement interval.

**Later Panic Symptom Change (Weeks 5 to 9)**

Higher levels of panic-focused interpretations at mid-treatment (Session 10, end of Week 5) predicted greater panic symptom improvement between Weeks 5 and 9 ($B = 1.79$ [95% CI: 0.61, 2.97], $SE = 0.59$, $t = 3.04$, $p = 0.004$, adjusted $p = 0.016$, $sr = 0.37$), subsequent to the measured session. By contrast, non-panic focused interpretations were unrelated to subsequent outcomes ($B = -0.47$ [95 CI: -1.54, 0.60], $SE = 0.54$, $t[59] = -0.88$, $p = 0.382$, adjusted $p = 0.437$, $sr = -0.09$). Panic-focused clarification at this session was also unrelated to outcomes ($B = -0.56$ [95% CI: -1.37, 0.25], $SE = 0.41$, $t[59] = -1.38$, $p = 0.175$, adjusted $p = 0.351$, $sr = -0.20$). Figure 1 summarizes all technique-outcome relationships for early and later panic symptom change.

**Interpersonal Problems as a Moderator of Technique-Outcome Relationships**

The higher the score on the IIP at intake, the stronger the relation of panic-focused interpretations at Session 10 was to subsequent symptom change ($B = 3.65$ [95% CI: 1.24, 6.05], $SE = 1.20$, $t[57] = 3.04$, $p = 0.004$, adjusted $p = 0.016$, $sr = 0.29$). The Johnson-Neyman technique identified an IIP score of 1.0 as the cutoff for exhibiting a
significant, positive relationship between panic-focused interpretations and subsequent improvement \((n\ of\ individuals\ with\ IIP\ \geq\ 1.0 = 38,\ 58.5\%;\ see\ Figure\ 2)\). For patients with scores greater than or equal to 1.0, indicating higher levels of interpersonal problems at baseline, there was a significant relation between panic-focused interpretations and subsequent symptom improvement \((sr = 0.41,\ p = 0.001)\), whereas for patients with less interpersonal distress the association was not significant \((sr = 0.11,\ p = 0.255)\), in large part because patients with lower interpersonal distress did well symptomatically regardless of interpretation level (see Figure 2).

**Discussion**

We investigated the relation between use of specific PFPP techniques and symptomatic outcomes in the treatment of panic disorder. Our first important finding suggests that at mid-therapy (Session 10), panic patients receiving a high level of panic-focused interpretations exhibited greater subsequent symptom improvement. However, non-panic-focused interpretations did not predict subsequent symptom improvement during either the earlier or later periods of treatment. Moreover, panic-focused interpretations at Session 10 also predicted to pre-to-post improvements in interpersonal functioning (see Appendix B). These findings lend support to the importance of taking a symptom-focused approach in short-term psychodynamic therapies for anxiety (Busch et al., 2012; Tasca, Hilsenroth, & Thompson-Brenner, 2014). Past process findings demonstrating a positive relationship between interpretations and symptom change (Pitman et al., 2014) could reflect that many interpretations in short-term psychotherapy in fact are symptom-focused, even when that is not the specific intent of the study.
On the other hand, interpretations in an early session, whether or not they were related to panic symptoms, did not predict subsequent panic symptom improvement over the following four weeks. Possibly, early interpretations are not as accurate as those made after the therapist has learned more about the patient. In previous psychodynamic process studies, observer-rated accuracy of interpretations derived from themes coded from early session transcripts or pre-treatment history interviews predicted greater symptom change and likelihood of having a “sudden gain” (Andrusyna et al., 2006; Crits-Christoph et al., 1988). Alternatively, interpretation at a very early stage may sometimes be experienced by the patient as overwhelming (McCarthy et al., 2016). At an early stage of therapy, it is also potentially less likely that patients have been fully socialized to the both the structure and tasks and goals of psychodynamic therapy (Luborsky, 1984), which may make it harder for them to build on therapists’ interpretations with further personal exploration and development of insight.

Taken together, it is possible that early interpretations are both less accurate to the patient’s dynamics, and less likely to be perceived by the patient as well-timed, which may be considered matters of intervention competence. Future process studies on interpretation accuracy and the role of supportive techniques and alliance in PFPP would help distinguish between these and other hypotheses explaining our findings. However, our findings do not support the conclusion that the other types of dynamic techniques assessed are necessarily without use or are counterproductive, but rather that panic-focused interpretations are the only statistically reliable signal of positive process for the average patient in this sample. It is also plausible that more complex process relationships exist (e.g., interactions between early and mid-technique use; levels of clarification and
interpretation), but we did not explore these possibilities due to our need to limit the number of tests performed in this small sample.

As we hypothesized, patients with more interpersonal problems at baseline exhibited a stronger relationship between mid-therapy panic-focused interpretations and subsequent change. This finding is consistent with PFFP’s theoretical model, which proposes that unconscious conflict in the context of relationships may contribute to experiences of panic, such that patients with more interpersonal distress may need a more intense focus on the emotions and conflicts underlying this distress (Busch et al., 2012). Panic dynamics are typically interwoven with the interpersonal problems that these patients struggle with. For example, many panic patients are prone to being in relationships where they struggle to assert their own needs (Zilcha-Mano et al., 2015). This linkage may allow for more readily identifiable dynamics and conflicts, and more opportunities to identify them in relation to interpersonal difficulties, compared to patients with relatively fewer interpersonal problems. Ergo, it may be the case that therapists were more accurate in their interpretations for patients with stronger interpersonal issues.

However, patients with lower levels of interpersonal problems had good symptom improvement in this interval (i.e., between Weeks 5 to 9) regardless of panic-focused interpretations, whereas increasing interpersonal distress was more predictive of poor symptom improvement in less panic-focused interpretive therapies (see Figure 2). We would argue that this pattern of results is concordant with the perspective that patients with more interpersonal problems particularly need a more panic-focused, interpretive therapy. Past psychodynamic process-outcome research has rarely sought to identify
beneficial matches between techniques and patient characteristics, with one exception being the body of literature suggesting an important role for transference interpretations in treating personality disorder (Hoglend et al., 2011; Keefe & DeRubeis, 2018). Overall, our finding is consistent with a perspective wherein patients with more complicated or treatment-resistant presentations may require more active or skillful approaches, and may reveal more about process-outcome relationships than those DeRubeis and colleagues (2014) have called “easier” (or more straightforward) patients (DeRubeis, Gelfand, German, Fournier, & Forand, 2014; Keefe et al., 2016).

**Limitations and future directions**

Fifteen cases (18.8%) were unable to be used for the present analyses due to dropout, which was nonrandom and related to poorer symptom trajectories in the parent trial (Milrod et al., 2016). The remaining patients included in our study represent a subsample of individuals who improved relatively more symptomatically. It is possible that the observed technique-outcome relationships would not be obtained among the dropout patients; alternatively, relatively less efficacious therapy process may have led to worse symptom trajectories, promoting dropout. Examining the relationship between technique use and treatment dropout would be an interesting way to disambiguate these possibilities, although this effort would be poorly powered in our sample due to the low base rate of dropout and lack of early therapy tapes due to dropout. In addition, several sessions were not available to be rated due to protocol error. However, we employed a standard, validated method for imputing the missing ratings (Stekhoven & Bühlmann, 2011), and results obtained with data only from cases with complete data mirrored those obtained when the imputations were included.
The patterns obtained using ratings from early and middle sections of therapy were observed using only one session from each phase. A better approach would be to sample multiple sessions from each phase (Dennhag, Gibbons, Barber, Gallop, & Crits-Christoph, 2012). Unfortunately, at one of the two treatment sites, only recordings of Sessions 2, 5, and 10 were available on a systematic basis. Future studies that include ratings of multiple recordings within the same interval would allow for the investigation of more complex patterns of the relation between process and symptom change, such as variability in technique use across sessions (Owen & Hilsenroth, 2014). In addition, the average therapy was not intensely interpretive (e.g., at session 10, less than one panic-focused interpretation per fifteen minutes), indicating that we could not meaningfully examine hypotheses that that intensities of interpretations in between the extremes is more effective than very low or very high intensities (see McCarthy et al., 2016).

Finally, our findings suggest but cannot confirm the presence of a causal relationship between the intensity of panic interpretations and symptom change. A stronger test of the causal hypothesis would require experimental manipulation of the causal variable, in a manner like Hoglend et al.’s (2008) randomized comparison of psychodynamic therapy with versus without transference interpretations.

**Conclusions**

Psychodynamic therapists implementing PFPP should focus on interpretation of the possible conflicts underlying panic as they enter the middle phase of therapy. Particularly tying the patient’s dynamics to experiences of panic, anxiety, and agoraphobia—rather than making general interpretations concerning relational or personal patterns—may be especially important for effective short-term treatment of
panic. For patients presenting with higher levels of interpersonal distress, an emphasis on panic-focused interpretations may be especially important in promoting remission from panic disorder.

References


DeRubeis, R. J., Gelfand, L. A., German, R. E., Fournier, J. C., & Forand, N. R. (2014). Understanding processes of change: how some patients reveal more than others-
and some groups of therapists less-about what matters in psychotherapy.


Table 1.1
Descriptive Data for Baseline Characteristics and for Symptom Change

<table>
<thead>
<tr>
<th>Baseline Measure</th>
<th>Mean (SD) or # (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline PDSS</td>
<td>13.9 (3.2)</td>
</tr>
<tr>
<td>PDSS Change Week 1 to 5</td>
<td>-1.5 (3.6)</td>
</tr>
<tr>
<td>PDSS Change Week 5 to 9</td>
<td>-1.2 (3.4)</td>
</tr>
<tr>
<td>SDS</td>
<td>16.1 (6.3)</td>
</tr>
<tr>
<td>HAM-D</td>
<td>10.6 (4.8)</td>
</tr>
<tr>
<td>IIP</td>
<td>1.2 (0.5)</td>
</tr>
<tr>
<td>Agoraphobia Diagnosis</td>
<td>53 (81.5%)</td>
</tr>
<tr>
<td>Age</td>
<td>39.5 (14.0)</td>
</tr>
<tr>
<td>Gender (Female)</td>
<td>44 (67.7%)</td>
</tr>
<tr>
<td>Ethnicity (Hispanic)</td>
<td>8 (12.3%)</td>
</tr>
<tr>
<td>Race (Black, Other Non-Caucasian)</td>
<td>11 (16.9%), 4 (6.2%)</td>
</tr>
<tr>
<td>Concurrent Psychopharmacology</td>
<td>16 (24.6%)</td>
</tr>
<tr>
<td>Age of Panic Onset (years)</td>
<td>27.5 (11.4)</td>
</tr>
<tr>
<td>SCID-II PersD Diagnosis</td>
<td>32 (49.2%)</td>
</tr>
<tr>
<td>Cluster A PersD Traits</td>
<td>1.4 (1.9)</td>
</tr>
<tr>
<td>Cluster B PersD Traits</td>
<td>2.7 (3.2)</td>
</tr>
<tr>
<td>Cluster C PersD Traits</td>
<td>4.0 (3.1)</td>
</tr>
<tr>
<td>Total PersD Traits</td>
<td>8.0 (6.3)</td>
</tr>
</tbody>
</table>

*Note.* HAM-D = Hamilton Rating Scale for Depression; IIP = Inventory of Interpersonal Problems; PDSS = Panic Disorder Severity Scale; SCID-II = Structured Clinical Interview for the Diagnosis of Axis-II Disorders; SDS = Sheehan Disability Scale
Table 1.2
**Descriptive Statistics for Technique Process Measurements (Average per 15-minute Segment)**

<table>
<thead>
<tr>
<th>Process Measurement</th>
<th>Session 2 M (SD)</th>
<th>Session 10 M (SD)</th>
<th>Stability Coefficient</th>
<th>Change in Technique Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panic-Focused Interpretation</td>
<td>0.9 (0.6)</td>
<td>1.2 (0.7)</td>
<td>( r = 0.50^{***} )</td>
<td>( d = 0.55^{***} )</td>
</tr>
<tr>
<td>Panic-Focused Clarification</td>
<td>2.3 (1.0)</td>
<td>2.3 (1.1)</td>
<td>( r = 0.09 )</td>
<td>( d = 0.11 )</td>
</tr>
<tr>
<td>Non-Panic-Focused Interpretations</td>
<td>0.2 (0.4)</td>
<td>0.5 (0.6)</td>
<td>( r = 0.03 )</td>
<td>( d = 0.53^{**} )</td>
</tr>
</tbody>
</table>

*Note.* * = \( p < .05 \); ** = \( p < .01 \); *** = \( p < .001 \); Cohen’s \( d \) for paired \( t \) test calculated using formula \( t_c \) (Dunlap, Cortina, Vaslow, & Burke, 1996)

Table 1.3
**Correlations between Technique Process Measurements and Baseline Characteristics**

<table>
<thead>
<tr>
<th>Baseline Characteristic</th>
<th>Panic Focus—Interpretation (S2/S10)</th>
<th>Panic Focus—Clarification (S2/S10)</th>
<th>Non-Panic Focused Interpretations (S2/S10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDSS</td>
<td>( r = -0.06 / -0.00 )</td>
<td>( r = -0.13 / 0.08 )</td>
<td>( r = -0.12 / 0.07 )</td>
</tr>
<tr>
<td>IIP</td>
<td>( r = 0.19 / 0.06 )</td>
<td>( r = -0.07 / -0.02 )</td>
<td>( r = -0.03 / 0.24 )</td>
</tr>
</tbody>
</table>

*Note.* All \( ps > .05 \). IIP = Inventory of Interpersonal Problems, PDSS = Panic Disorder Severity Scale
Figure 1.1. Estimated effect sizes for the relationship between psychotherapy technique use at a given session and subsequent improvement in panic symptoms as measured by the PDSS. Positive semipartial correlations indicate that higher levels of the intervention are associated with more subsequent symptom improvement. Bars are 95% confidence intervals.
Figure 1.2. Estimated change in panic symptoms as measured by the PDSS between Weeks 5 to 9, as a function of degree of interpersonal problems as measured by the IIP and their interaction with use of panic-focused interpretations at session 10. Positive values represent predicted symptom worsening, while negative values represent predicted symptom improvement. All regression variables not displayed in the figure were set to the sample means.
Patient Personality and Emotional Process

Panic-focused psychodynamic psychotherapy (PFPP) (Busch, Milrod, Singer, & Aronson, 2012) is a 24-session, 12-week evidence-based treatment for panic disorder (PD; Beutel et al., 2013; Keefe, McCarthy, Dinger, Zilcha-Mano, & Barber, 2014; Milrod et al., 2016; Milrod et al., 2007) that focuses on understanding and working through unrealized or disavowed meanings surrounding the onset of acute attacks and associated anxiety in patients with panic disorder. One common meaning might be real or imagined loss of attachment figures, as suggested by epidemiological data showing high comorbidity or past history of separation anxiety among panic patients (Kossowsky et al., 2013; Milrod et al., 2014), and the fact that emotional stressors, such as relationship conflict and interpersonal loss, frequently precede the development of panic disorder (Klass et al., 2009; Scocco, Barbieri, & Frank, 2007). Recognizing the sometimes-conflicted emotions and fantasies connected with experiences of panic and anxiety (such as unacknowledged rage at attachment figures) is hypothesized to improve panic-specific reflective functioning (PSRF; Rudden, Milrod, Target, Ackerman, & Graf, 2006). PSRF is an interview-based measured intended to tap into the degree to which patients can identify and discuss potential psychological meanings surrounding and triggers to experiences of panic and anxiety (Rudden et al., 2006). Early improvement in PSRF has been shown to predict subsequent improvement in panic symptoms in both PFPP and cognitive-behavioral therapy (CBT) (Barber et al., under review).

The theory that informs CBTs for panic disorder focuses on patients’ tendencies to catastrophically misinterpret bodily sensations (Clark et al., 1997) and thereby to fear
those sensations (Boswell et al., 2013). However, specific problems in emotional recognition and acceptance have also been observed in studies of patients with panic disorder, suggesting that these factors may also contribute to panic experiences. Patients with PD report higher rates of alexithymia, experiential avoidance, and lack of emotional acceptance, compared to psychiatrically healthy controls and persons with simple phobias (Galderisi et al., 2008; Izci et al., 2014; Parker, Taylor, Bagby, & Acklin, 1993; Tull & Roemer, 2007). Relative to non-psychiatric controls, panic patients have also been observed to use relatively more emotional avoidance strategies in response to viewing negatively or positively-valenced film clips (Tull & Roemer, 2007). Experimentally instructing use of such strategies in healthy controls (relative to allowing emotional experience) promotes subjective distress, heightened physiological arousal in the moment, and also physiological reactivity in a subsequent stressful interpersonal task (Tull, Jakupcak, & Roemer, 2010).

Avoidance of emotions in day-to-day life may promote development of panic attacks if emotional contents are not addressed—for example, a person may get strongly physiologically aroused due to unacceptable emotions, but be unable to dissipate that arousal due to lack of emotional awareness and a consequent inability to acknowledge what is upsetting him/her. A patient may also attribute emotional arousal to frightening somatic causes that can psychologically stand in for conflicted feelings. In support of this conceptualization, a recent process study of CBT for panic disorder found that reductions in patient reports of emotional suppression preceded improvements in catastrophic, body-focused cognitions and panic symptoms (Strauss, Kivity, & Huppert, 2018).

**Emotions and Outcome in Psychodynamic Therapies**
Psychodynamic frameworks conceptualize attempts to avoid affects and particular affectively charged contents as contributing to the emergence of psychiatric symptoms and psychosocial dysfunction (Subic-Wrana et al., 2016). Within these frameworks, defense mechanisms work to maintain lack of awareness (Perry & Bond, 2012), which occurs to protect the person from experiencing psychic danger, yet precludes the individual’s ability to process and address the relevant conflicts or wishes. Panic patients have been observed to have heightened use of so-called neurotic and immature defenses relative to healthy controls (Busch, Shear, Cooper, Shapiro, & Leon, 1995; Calati, Oasi, De Ronchi, & Serretti, 2010; Kipper et al., 2004), and to exhibit a unique defensive profile compared to depression patients (Busch et al., 1995).

Therapeutic focus on difficult-to-express or disavowed affect has been commonly considered to be a feature distinguishing psychodynamic therapies (PDT) from cognitive-behavioral approaches (Blagys & Hilsenroth, 2000; McCarthy & Barber, 2009). In PDT, such focus can be achieved via supportive interventions encouraging expression of affect, clarification/confrontation highlighting important areas of affective exploration, or interpretations of affectively-laden meanings. In a meta-analytic examination, psychodynamic therapies in which therapists were coded by observers as being especially affect-focused were more successful in symptomatic outcomes, with a medium effect size (Diener, Hilsenroth, & Weinberger, 2007), although temporal precedence was not established in most studies and some studies in fact involved patient expression. In two recent studies assessing psychodynamic therapist technique specifically in anxiety disorder therapies (Pitman, Slavin-Mulford, & Hilsenroth, 2014) and in a transdiagnostic sample examining anxiety symptom improvement (Pitman, Hilsenroth, Weinberger,
Conway, & Owen, 2017), higher focus by the therapist on unexpressed/avoided affects was related to more pre-to-post treatment symptom improvement.

In PDTs, links (often temporally sequenced) have been found between the degree to which a patient engages in emotional experiencing or processing and positive outcomes (Abbass, Town, Ogrodniczuk, Joffres, & Lilliengren, 2017; Fisher, Atzil-Slonim, Bar-Kalifa, Rafaeli, & Peri, 2016; Friederich et al., 2017; Johansson, Town, & Abbass, 2014; Kramer, Pascual-Leone, Despland, & de Roten, 2015; Town, Abbass, & Bernier, 2013; Town, Salvadori, Falkenström, Bradley, & Hardy, 2017), convergent with findings in humanistic-experiential therapies (Pascual-Leone & Yeryomenko, 2016). A recent meta-analysis suggested that, across psychotherapies and psychiatric disorders, increased expression of emotion by the patient although many of the included studies had unclear temporal precedence between the affective measurement and outcome (Peluso & Freund, 2018). In their transtheoretical conception of the importance of emotional expression and experiencing in psychotherapy, Lane, Ryan, Nadel, & Greenburg (2015) propose that emotional activation of episodic and semantic memory content facilitates the reconsolidation of those memories into new, potentially more adaptive forms. From a psychodynamic perspective, heightened affective experiences in therapy may indicate that a patient is tolerating more affectively charged material, allowing for working through of conflicts and ultimately improvements in reflective functioning. Heightened patient emotional experiencing has been examined as a positive predictor of improvement in a study of intensive short-term psychodynamic therapy for patients with generalized anxiety disorder (Lilliengren, Johansson, Town, Kisely, & Abbass, 2017), but we are not aware of any other such study in anxiety.
Study Hypotheses: Main Effects and Moderators

PFPP’s clinical theory predicts that patient affective engagement is important to exploration of meanings and conflicts surrounding panic and anxiety experiences (Busch et al., 2012). To examine the relationships between early in-session emotional engagement and subsequent changes in symptoms and PSRF, we developed a measure of in-session emotional expression. Emotional expression can be thought of as a broad process marker of engaged emotions, incorporating basic aspects of both emotional experience (i.e., does the patient exhibit non-verbal signs of emotional activation?; McCullough et al., 2003) and processing (i.e., does the patient speak in an identifiable manner about specific, current emotional experiences; Klein, Mathieu, Gendlin, & Kiesler, 1969; Pascual-Leone & Greenburg, 2005)? Levels of emotional expression were assessed in early sessions (2, 5, and 10) of a 24-session PFPP protocol from a two-site randomized controlled trial comparing PFPP to CBT and applied relaxation training for panic (Milrod et al., 2016). We hypothesized that patients with higher levels of emotional expression in early PFPP sessions would experience greater symptom and PSRF improvement subsequent to the process-measured sessions. Moreover, we examined both overall patient emotional expression and expression of specific emotional states, to help determine whether patient engagement with particular affects is especially important to treatment process (e.g., grief; anger) or potentially deleterious to treatment (e.g., anxiety in-session, as might be predicted by psychodynamic conflict models).

Moderators. Although many psychotherapy processes are thought of as universally positive (e.g., the therapeutic alliance), it may be that some processes are more important in some psychotherapies, relative to others. Also, the same process,
depending on traits that can capture an aspect of specific functioning relating to affect, may yield strong positive effects in some patients, little effect in others, and negative effects in still others (cf. Lorenzo-Luaces et al., 2017). We hypothesized that patient personality disorder traits related to constrained (obsessive-compulsive personality disorder; OCPD) versus heightened (borderline personality disorder; BPD) affect would moderate the relationship between emotional expression and symptom improvement.

A psychodynamic conception of OCPD might focus on OCPD patients’ tendency to intellectualize emotional experiences (Caligor, Kernberg, & Clarkin, 2007; Summers & Barber, 2010). The typically theorized repertoire of defense for OCPD patients emphasizes focus on cognition and circumstance over affects (e.g., intellectualization; isolation of affect; rationalization); a strong need for control, order, and perfection defends against the dangers of experiencing and acting on affect and wishes that feel destabilizing. Relative to psychiatrically healthy individuals, OCPD patients report being less accepting of their emotions, less clear about what their emotions mean, and more distressed by feeling emotional (Steenkamp, Suvak, Dickstein, Shea, & Litz, 2015). In treatment with OCPD patients, focusing on affect may help to counteract defenses that represent avoidance of emotional conflicts (Barber & Muenz, 1996). We thus hypothesized that patients with more OCPD personality traits as indicated by the SCID-II (First, Gibbon, Spitzer, Williams, & Benjamin, 1997) would show a stronger relationship between early levels of emotional expression and later symptomatic improvements, as for these patients emotional expression may particularly indicate a more flexible use of defense in-session, relative to their typical profile.
In contrast, we hypothesized that patients with increasingly more BPD traits would show no or even negative relationships between early emotional expression and later improvements. While BPD is a relatively uncommon comorbidity to PD (Friborg, Martinussen, Kaiser, Overgard, & Rosenvinge, 2013), many patients exhibit elevated BPD pathology. In general and clinical populations, qualifying for even one DSM-defined BPD criterion is uniquely prognostic of significant interpersonal dysfunction and psychosocial disability (Ellison, Rosenstein, Chelminski, Dalrymple, & Zimmerman, 2016; Zimmerman, Chelminski, Young, Dalrymple, & Martinez, 2012), indicating that even non-diagnostic BPD may be clinically relevant. Affective dysregulation is a common if not defining feature of BPD. BPD patients show stronger reactivity to interpersonal events (Santangelo, Bohus, & Ebner-Priemer, 2012), as well as more labile mood around a more negative baseline (Ebner-Priemer et al., 2015), relative to healthy controls and MDD patients. Accordingly, the therapist’s efforts to help patients regulate, understand, and usefully work with difficult-to-comprehend affects has been identified as a common feature across many empirically supported psychotherapies for BPD (Bateman, Gunderson, & Mulder, 2015). For patients with more BPD traits, relatively lower levels of emotional expression in PFPP may reflect successful work by the therapist and patient to contain affect in order to better work with the meanings and circumstances surrounding panic and anxiety (Fonagy & Luyten, 2009). Relatively lower emotional expression may also indicate defensive flexibility among BPD patients, who typically engage in defenses focusing on prominent, moment-to-moment affects (e.g., splitting; acting out; Kramer, de Roten, Perry, & Despland, 2013; Perry, Presniak, &
Olson, 2013), such that emotional expression per se may sometimes involve typical defensive processes and not primarily a flexible use of defense or processing of conflict.

Method

Participants

Patients. The present study is a secondary analysis of patients randomized to the PFPP condition (N = 80) of a two-site randomized controlled trial comparing PFPP, CBT, and applied relaxation training among patients with primary DSM-IV panic disorder with or without agoraphobia. Patients were recruited at <site A> and <site B>. Participants received study treatment gratis. Participants gave informed written consent. Both sites’ institutional review boards approved the protocol, and the study is registered with ClinicalTrials.gov (identifier: NCT00353470).

Patients were included in the trial if they had one or more weekly spontaneous panic attacks for the month before trial entry, and qualified for DSM-IV diagnosis of primary panic disorder with or without agoraphobia determined as per the ADIS-IV (DiNardo, Brown, & Barlow, 1995). Cross-site agreement on ADIS ratings for panic severity (with a “4” indicating the diagnostic threshold) was excellent (ICC = 1.00; per the norms of Portney & Watkins, 2000). See Milrod et al., 2016 for further details.

Non-study psychotherapy was prohibited. Medications were permitted if stable for at least two months at presentation, and were recorded, held constant, and monitored during the trial. Exclusion criteria were: active substance dependence (less than 6 months’ remission), a history of psychosis or bipolar disorder, acute suicidality, and organic mental syndrome. Additional details on trial design, independent evaluator
training, and therapy adherence can be consulted in the primary outcome paper (Milrod et al., 2016).

**Assessment of personality disorder (PersD).** The Structured Clinical Interview for Axis-II disorders (First et al., 1997) was used to assess the presence of PersD criteria and diagnoses as defined by DSM-IV. Trained, independent masters’ level diagnosticians uninformed to treatment condition administered the interviews. Cross-site interrater reliability was excellent for number of OCPD criteria scored as present \((ICC[2,1] = 1.00)\), and moderate for BPD criteria \((ICC[2,1] = 0.78)\).

**Therapists.** All therapists were experienced therapists (Ph.D. or M.D.) who were specifically trained in PFPP over the span of a 2-day, 12-hour course. Therapists had an average of 15 years of post-graduate experience \((SD = 8.2)\), and an average of 5 years’ experience in some form of time-limited psychodynamic therapy \((SD = 6.3)\). Therapists participated in monthly group supervision and received regular individual supervision from senior clinicians. For the primary outcome paper, adherence to PFPP was established, and additional information on the number and training of therapists can be found there (Milrod et al., 2016).

**Outcome Indices**

**Panic Disorder Severity Scale (PDSS; Shear et al., 1997).** The PDSS is a diagnosis-based, composite, global observer rating of panic disorder severity, with acceptable psychometric properties. The PDSS was administered by trained, master’s-level diagnosticians, uninformed as to treatment condition, based on an interview guide. Interrater reliability on the PDSS was excellent \((ICC[2,1] = 0.95)\). The PDSS was
administered five times during treatment: at baseline (Week 0), Week 1, Week 5, Week 9, and termination (Week 12).

**Panic-Specific Reflective Functioning (PSRF; Rudden et al., 2006).** PSRF is an interview-based measure that assesses the degree to which patients can recognize the psychological contributions to their panic symptoms. Respondents are queried as to their understanding of panic, how it has changed over time, and whether they notice any concordance between panic and emotional states. Each response is rated for psychological mindedness and complexity, and an overall score is based on item scores. An example of a more impaired PSRF answer might be: “It’s the heat, the heat brings them on” in response to “Why do you think you have panic attacks?”; a less impaired response to that question might be: “I notice that I get them when I am feeling a lack of control in my personal relationships. I fear that others will leave me, or that I may want to leave them.” PSRF narratives were reliably scored by trained raters ($ICC[2,1] = 0.80$).

**Process Measures**

**Emotional Expression Rating Scale (EERS; Huque & Keefe, 20).** The EERS is an observer-rated scale developed to assess the degree to which patients are engaged in emotional discourse. It was developed for this investigation by two study authors, with reference to prior attempts to rate emotions in-session (e.g., Klein et al., 1969; McCullough et al., 2003). The measure was designed to be relatively atheoretical and broadly descriptive rather than reflecting a particular perspective on what constitutes (for example) adaptive emotional experience or deep emotional processing. Contrasted with other scales, it was designed for ease of use by individuals with less theoretical or clinical
training. To incorporate information on both verbal and non-verbal indicators of emotional expression, only sessions with usable video recordings could be sampled. Ratings were made every five minutes to allow for a more precise and potentially valid assessment of patient emotion expression relative to whole-session ratings, and to collect temporal information about development of emotional responses throughout a session. Raters made both an omnibus rating for overall emotional expression, and ratings for 4 broad emotion categories.

The primary rating concerns a patient’s overall emotional expression, which is rated on a 0 to 5 scale. Higher scores are dependent on the peak of emotional intensity during the rated segment, but also on the duration of non-neutral emotional expressivity. A score of 0 indicates no notable emotional expression, or a relatively neutral or flat affective state. A score of 2 indicates a low-to-medium but clear presence of emotional expression lasting more than a few seconds, up to a minute, with use of affectively charged language or verbal statements of feeling, and non-verbal indicators such as choking up, tearing up, emphatic gestures, muscular tension, smiling, or physical agitation/restlessness. Scores of 4 and 5 indicate a high, unconstrained peak affective arousal with sustained emotional expression across the majority of the rated segment. Ratings were made for emotional expression within-session, and not when a patient reported how they were feeling in the past, with no clear indication that they are currently feeling that particular emotion (e.g., reporting with a neutral tone that they were sad when a parent died). In Table 2, examples of transcripts from higher- versus lower-emotional expression therapy sessions from the present study are given.
Expressions of four specific broad categories of emotions were coded, three of which are referenced in the following analyses:\(^1\):

- **Grief/Sadness**: Grief/Sadness referred to the expression of grief, sadness, sorrow, or regret. Raters were trained to distinguish between sadness surrounding (for example) an interpersonal figure, and indistinct depressive distress or self-attack, which was rated as Anxiety/Distress. Nonverbal indicators included tearing up, sad vocal tone, choking up, quavering voice, frowning, and crying. Raters were trained to pay attention to clear verbal expressions of sadness, regret, or loss, or particular expressions such as feelings of closeness and tenderness while talking about a close attachment figure who died or whom they fear may no longer be available to them.

- **Anger/Assertion**: Anger/Assertion made a distinction between anger directed toward an external or interpersonal figure, and anger directed toward the self (e.g., self-attack; self-punishment), which was rated as Anxiety/Distress. Anger/Assertion was rated if patients indicated clear anger, criticized someone, asserted their needs or desires to someone, or voiced wishes for recompense for a past misdeed. Nonverbal markers included patients’ showing angry facial expressions, angry vocal tone, becoming more tense, clenching their fists, gesturing aggressively, or gritting their teeth.

---

\(^1\) Positive Affect was rarely rated as present and had low variance, and consequently we did not analyze these scores.
- Anxiety/Distress: Anxiety/Distress was intended to be rated when a patient appeared activated with clearly negative feeling and arousal, such as negative global feelings about the self, anxious tension, or experiencing undifferentiated aversive affect. Non-verbal Anxiety/Distress indicators included visible discomfort, fidgeting and drawing inwards, hyperventilating or full panic attack, and uncontrollable negative affect. Verbal indicators included explicit references to feeling anxious or distressed or statements of self-attack and hopelessness.

Correlations between emotion ratings on a per-5 minute basis indicated independence of the specific emotions (see Appendix D).

A total of 15 undergraduate psychology majors were trained on this measure by study authors JRK and ZH. Raters were trained by rating sessions not used in this investigation until they reliably rated within one point of the trainers. Weekly-to-biweekly anti-drift sessions were held to maintain reliability and discuss rating challenges. Raters were uninformed to patient outcomes and other study patient data. Videotapes were coded in a random order, and the average score between three raters was used for all analyses.

Interrater reliability (Shrout & Fleiss, 1979) as calculated in a linear mixed model (Bates et al., 2017) was good to moderate per-5 minutes for all emotions (ICC [2,3] Overall = 0.74, Grief/Sadness = 0.71, Anger/Assertion = 0.76, Anxiety/Distress = 0.66).

Statistical Analyses
All analyses were conducted in the \( R \) statistical computing language \((R \text{ Core Team, 2017})\). All primary analyses were run using robust regressions as implemented in the \( R \) package “Robustbase” \((\text{Maechler et al., 2016})\). Given the effective sample size \((n = 44)\), robust regression was selected over standard regression for its superior properties of robustness against multivariate outliers and deviation from homoscedasticity \((\text{Huber & Ronchetti, 2009})\). Semi-partial correlation effect sizes \((sr)\) were estimated for parameters of interest.

**Missing data.** In the parent clinical trial, a not missing at random (NMAR) pattern of treatment dropout was detected, such that patients with worse PDSS symptom trajectories were more likely to terminate from treatment prematurely \((\text{Milrod et al., 2016})\). When outcomes for treatment noncompleters are imputed in the NMAR context, imputation and other missing data methods can lead to biased estimates and confidence intervals \((\text{Graham, 2009})\). As such, only individuals who provided data up to Week 9 \((\text{the 4}^{\text{th}} \text{assessment})\) were eligible for inclusion \((n = 65; 81.3\% \text{ of the intent-to-treat sample})\). Moreover, due to the need for video to rate nonverbal indicators of emotional expression, only individuals with videotaped sessions \((\text{rather than audio backups})\) were used. The final sample size having videotaped sessions was \(n = 44\).

Some trial completers were missing one out of three videotaped session due to technical issues or therapist/research assistant error \((n = 13, 29.5\% \text{ of the reduced sample})\). In contrast to noncompleters, process ratings for completers missing a video-recording of a session can be presumed to be missing at random in relation to panic symptom outcomes \((\text{Rubin & Little, 2002})\), and psychotherapy process ratings have been imputed in past investigations \((\text{Keefe, Solomonov, et al., 2018; Lorenzo-Luaces et al.,})\).
Random forest imputation (Stekhoven & Bühlmann, 2011) was used to impute missing data for completers. For this dataset, a normalized root mean square error of prediction was estimated at 0.12, indicating that imputation accuracy was estimated to be more than adequate (Stekhoven & Bühlmann, 2011).

**Analytic strategy.** Using robust linear regressions, the average level of overall emotional expression between Sessions 2, 5, and 10 (i.e., three sessions between baseline and Week 5) was used to predict subsequent change in PDSS-measured panic symptomatology between Week 5 and treatment termination (Week 12). We also analyzed PSRF as a secondary outcome. Because site differences were detected in the parent trial (Milrod et al., 2016), we explored whether site moderated focal effects for this and all subsequently described analyses at a significance of at least $p < .10$. Control covariates in each regression included the baseline value for either outcome, and the degree of change in the variable of interest that occurred between baseline and Week 5 (i.e., during the measurement of emotional expression). In addition, two planned analyses of moderators of the relationship between average emotional expression and PDSS symptom improvement were conducted: baseline SCID-II OCPD criteria, and SCID-II BPD criteria.

We also performed two secondary, exploratory analyses. First, for both PDSS and PSRF change, we examined whether any of three specific types of emotional expression (i.e., Grief/Sadness, Anger/Assertion, Anxiety/Distress) were predictive of subsequent change. Second, in the context of recent debates about the degree to which stable therapy process across sessions merely reflects trait-like features of patients (Falkenström, Finkel, Sandell, Rubel, & Holmqvist, 2017), we examined whether having emotional expression
at Session 10 greater than one’s average across Sessions 2, 5, and 10 predicted subsequent symptom change (i.e., a term reflecting Session 10 minus the average of all sessions). This analysis also simultaneously controlled for average emotional expression within the same model.

Results

Descriptive Statistics

The average patient in this study had a baseline PDSS score of 13.8 (range 9 to 20), considered to be in the moderately-ill severity range for patients with comorbid agoraphobia (Furukawa et al., 2009). Other patient demographics and clinical information can be found in Table 1.

Descriptive statistics on mean levels of emotional expression across each session can be found in Appendix D. We also examined the consistency of emotional expression across measured sessions, using a linear mixed model framework (Bates et al., 2017; Appendix D), which revealed that the majority of variance in emotional expression was not trait-like across sessions.

PDSS Symptom Change

Patients who had higher levels of emotional expression across the three sampled sessions (2, 5, and 10) experienced more panic symptom improvement subsequent to Session 10 ($B = -3.10$ [95% CI: $-5.27$ to $-0.93$], $SE = 1.07$, $t[40] = -2.88$, $p = 0.006$, $sr = 0.31$), controlling for baseline panic symptoms and prior change in panic symptoms.

In a secondary analysis, having emotional expression at Session 10 that was higher than one’s average emotional expression across Sessions 2, 5, and 10 was also predictive of having more symptom improvement subsequent to Session 10 ($B = -3.58$
52

[95% CI: -5.78 to -1.37], $SE = 1.09$, $t(39) = -3.28$, $p = 0.002$, $sr = 0.38$). This was true even though the model also controlled for a patient’s average level of emotional expression, which remained itself a statistically significant predictor of subsequent panic symptom improvements ($B = -2.90$ [95% CI: -4.70 to -1.09], $SE = 0.89$, $t(39) = -3.24$, $p = 0.002$, $sr = 0.29$).

We further explored whether any specific type of emotional expression was particularly responsible for this relationship, employing a model simultaneously including as predictors average expression levels of grief/sadness, anger/assertion, and anxiety/distress. Patients who expressed more grief/sadness across the sampled sessions had superior subsequent panic outcomes ($B = -5.25$ [95% CI: -9.38 to -1.14], $SE = 2.18$, $t(38) = -2.41$, $p = 0.021$, $sr = 0.28$), while neither anger/assertion ($B = 2.43$ [95% CI: -1.61 to 6.46], $SE = 1.93$, $t(38) = 1.25$, $p = 0.217$, $sr = -0.16$) nor anxiety/distress ($B = 0.01$ [95% CI: -5.37 to 5.40], $SE = 2.68$, $t(38) = 0.01$, $p = 0.996$, $sr = 0.00$) were significant predictors of symptom change. Figure 1 displays these relationships for overall emotional expression and specific emotions.

**Personality Moderators of Symptom Change**

Next, we examined our moderation hypotheses as to whether BPD and OCPD personality traits would predict a smaller or greater relationship (respectively) between emotional expression and symptom improvements. Consistent with our hypothesis, the number of baseline SCID-II BPD criteria a patient met significantly interacted with overall emotional expression, such that meeting more BPD criteria attenuated the relationship between overall emotional expression and symptom improvement ($B = 2.36$ [95% CI: 1.09 to 3.63], $SE = 0.63$, $t(38) = 3.76$, $p < .001$, $sr = 0.29$). Unpacking this
continuous interaction using the Johnson-Neyman technique, we found that patients
meeting 1 or no BPD criteria evidenced a significant, positive relationship between
emotional expression and symptom improvement ($sr = 0.33, p = 0.003$; see Figure 2). By
contrast, there was a nonsignificant, negative relationship between emotional expression
and improvement for patients meeting 2 or more BPD criteria ($sr = -0.20, p = 0.062$).
Thus, patients qualifying for a relatively low number of DSM-defined BPD criteria (2 or
more) exhibited no beneficial (or potentially even a negative) relationship to degree of
emotional expression compared to those with minimal (1) or no BPD pathology, who had
a significant, positive relationship.

However, as concerned SCID-II OCPD criteria, there was a significant interaction
between OCPD criteria, emotional expression, and site ($p = 0.036$). At Site A, the
hypothesized interaction was obtained, whereby patients meeting more OCPD criteria at
baseline had an increasingly positive relationship between early emotional expression and
subsequent symptom improvement ($B = -4.48 [95\% CI: -7.72 to -1.24], SE = 1.59, t[34] =
-2.81, p = 0.008, sr = 0.34$). At Site B, there was no significant interaction ($B = 0.48
[95\% CI: -2.11 to 3.07], SE = 1.27, t[34] = 0.38, p = 0.707, sr = -0.04$). Thus, OCPD
criteria cannot be considered a clear moderator.

**Panic-Specific Reflective Functioning Change**

Mirroring the model analyzing symptomatic outcomes, patients with a higher
level of emotional expression across the three sessions had greater gains in PSRF
subsequent to the measured sessions ($B = 0.93 [95\% CI: 0.03 to 1.83], SE = 0.45, t[40] =
2.10, p = 0.042, sr = 0.23$), controlling for their baseline PSRF and early PSRF changes.
Neither the number of OCPD ($B = 0.42 [95\% CI: -0.27 to 1.11], t[38] = 1.24, p = 0.223$,
sr = 0.21) nor BPD criteria (B = -0.42 [95% CI: -1.12 to 0.27], t[38] = -1.23, p = 0.226, sr = 0.20) a patient met significantly moderated this relationship.

We examined whether any specific type of emotional expression drove this relationship. In this case, there were no specific significant relations as concerned expression of grief/sadness (B = 1.43 [95% CI: -1.62 to 4.48], SE = 1.51, t[38] = 0.95, p = 0.350, sr = 0.15), anger/assertion (B = 1.20 [95% CI: -0.54 to 2.95], SE = 0.86, t[38] = 1.40, p = 0.170, sr = 0.17), or anxiety/distress (B = -0.15 [95% CI: -2.02 to 1.71], SE = 0.92, t[38] = -0.17, p = 0.869, sr = -0.02).

Discussion

In panic-focused psychodynamic psychotherapy, patients who engage in more emotional expression over the course of the first five weeks of therapy have superior symptomatic outcomes across the remainder of the treatment. For interpretive context, just under half of the average symptomatic change occurs after the first five weeks of treatment (see Table 1). Our results are convergent with findings from psychodynamic psychotherapy process studies investigating other operationalizations of emotional processing or experiencing in-session (Abbass et al., 2017; Fisher et al., 2016; Friederich et al., 2017; Johansson et al., 2014; Kramer et al., 2015; Town et al., 2013; Town, Salvadori et al., 2017). These findings expand the literature on emotions in psychodynamic therapy to the context of short-term manualized panic disorder-focused treatment, and provide further evidence that emotional expression precedes rather than contemporaneously occurs with improvements. This study thus provides relatively stronger evidence that emotional expression is a therapy process that may help give rise to good outcomes, rather than being merely a product of symptomatic alleviation.
Notably, early emotional expression also predicted subsequent gains in PSRF, which were found in the broader trial to predict subsequent panic improvements (Barber et al., under review). Emotional expression may be reflective of mentalization in-session of emotional meanings and conflicts contributing to instances of panic and anxiety. As the patient works to tolerate and understand affectively-laden meanings and conflicts, they are expected to be able to engage with them rather than feel them as undifferentiated anxiety or frightening somatic fantasy (Busch et al., 2012). Emotionally working through conflict may also serve to detoxify certain affects, meanings, or wishes for patients, which may be typically experienced as distressing, dangerous, uncontrollable, or guilt-provoking to the point of having a panic attack.

Clinically, our results suggest that emotionally flat or withdrawn discussions of the contributors and contexts of anxiety are unlikely to help a patient substantively experience and work with these meanings in psychodynamic therapy. Concordantly, past work has found that psychodynamic interventions focused on expression of difficult-to-express affects appear to be correlated with improvements in anxiety (Pitman et al., 2014; Pitman et al., 2017), and that anxiety patients rate as especially helpful attempts to explore unexpressed/avoided feelings in-session (Glock, Hilsenroth, & Curtis, 2018). The presence versus absence of emotional expression may help a therapist distinguish between a patient’s compliant or pseudo-insightful (intellectualized) acquiescence to the therapist’s attempts to explore the emotional underpinnings of their anxiety, versus productive, affective engagement in the therapeutic process.

In our exploratory analyses concerning which specific emotions contribute to the observed relationships, only expression of grief/sadness significantly predicted
subsequent symptom outcomes. PFPP’s clinical theory emphasizes the degree to which actual or prospective attachment losses can precipitate panic disorder itself, often because patients have difficulty acknowledging the emotional importance of the loss (Busch et al., 2012; Klass et al., 2009; Milrod, Leon, & Shear, 2004). Expression of grief/sadness may sometimes reflect patients engaging with and becoming more tolerant of the true emotional impact of the actual or feared loss.

Anger may not have emerged as a significant predictor despite its role in psychodynamic models of panic (Busch et al., 1999; Busch et al., 2012), as anger may sometimes have been directed toward figures regarding whom the patient already feels comfortable experiencing anger. In PFPP, anger is conceptualized as being repressed particularly when the patient is worried that his/her anger will provoke separation/retaliation and hence loss of an ambivalently-held attachment figure (Busch et al., 1999; Busch et al., 2012; Rudden et al., 2003). As such, access to and expression of grief/sadness about loss that might be imagined to emerge as a result of angry feelings might indicate a deeper processing of the underlying conflict. Expression of disavowed or unconscious anger, for instance regarding attachment figures, may be found to have stronger relationships to outcomes compared to general anger. Interestingly, in contrast with our findings, in an investigation of intensive short-term dynamic psychotherapy for treatment-resistant major depressive disorder, experiencing of anger but not grief was related to better depression outcomes (Town, Falkenstrom, Abbass, & Stride, 2017; Town, Salvadori et al., 2017). It may be that difficulties tolerating particular, core emotions are more common to specific symptoms.
Finally, anxiety/distress in-session was nonpredictive of outcome. In PFPP, presence of anxiety in-session is not considered negative per se, as for example the patient may feel anxious because the therapeutic work is addressing important conflicted material, and that anxiety can be dealt with in real-time (Busch et al., 2012). From the perspective of PFPP, a therapy that transpires without affects sometimes considered “negative” or “inhibitory” (McCullough et al., 2003) may reflect that particular important dynamics are not being brought into the treatment by the patient or pursued by the therapist. Our metric of Overall emotional expression incorporated all deviations from a neutral emotional state—which could include so-called negative/inhibitory affects—in part to reflect our interest in emotional engagement in treatment writ large relative to therapies in which comparatively minimal affect was mobilized in-session. Other investigations with different goals may choose to use a modified version of the EERS explicitly excluding Anxiety/Distress codes from consideration in the Overall score, or other metrics that attempt to explicitly distinguish negative/inhibitory affects, such as the Achievement of Therapeutic Objectives Scale (McCullough et al., 2003).

We also examined whether personality disorder traits theoretically related to constrained (OCPD) or heightened (BPD) emotional expression moderated the relationship between emotional expression and symptom change. OCPD traits only moderated the predictive value of emotional expression at one of the two treatment sites, such that perhaps unmeasured differences in process or patient population further affected this relationship—for instance, patients at Site B in the trial were much more likely to be on psychotropic medications and had a significantly higher number of medication classes taken (Milrod et al., 2016). On the other hand, the presence of
elevated BPD traits diminished the relationship between emotional expression and symptom improvements. With patients meeting relatively more BPD criteria (>1), the PFPP therapist may wish to focus on helping a patient slow down and contain their strong affects to better work with the meanings giving rise to powerful emotional experiences. Of note, meeting relatively more subthreshold BPD criteria per se was not a negative prognostic indicator for PFPP patients in this trial and in fact BPD symptoms improved more in PFPP than in CBT in this study (Keefe, Milrod, Gallop, Barber, & Chambless, 2018), suggesting that these patients can benefit from PFPP, given therapeutic processes that are tailored to their needs. However, few patients in this study met SCID II criteria for a full comorbid BPD diagnosis, potentially because acute suicidality was an exclusion criterion, limiting our assessments to PD patients with low levels of borderline pathology.

**Limitations and future directions**

Several patients had one session (of three) unavailable to be rated due to videotape missingness (n = 13, 29.5% of the reduced sample). However, we employed a standard, validated method for imputing missing ratings (Stekhoven & Bühlmann, 2011), and a complete data analysis showed similar results. In addition, our effective sample size was small (n = 44).

In our study, average levels of emotional expression on the level of the session were relatively low (around a 1 on the 5-point scale), with the range of rated values consisting of emotionally flat sessions (mean 0) to heightened but not (consistently) highly activated sessions (mean 2). Our obtained relationships for emotional expression generally do not include representation of consistently high-to-extremely emotional sessions, and it is possible that such consistently high expression would not relate to
positive outcomes. There was mixed evidence whether maximum emotional expression rated in a session (rather than average across segments) was predictive of outcomes (see Footnote 2), with a trend level relationship for grief/sadness (cf. Kramer et al., 2015). Emotional expression may also have a different relationship to outcome later in treatment.

Our process measure to assess emotional expression, based on the apparent intensity and duration of expression, consisted of a simpler operationalization of emotional experiencing/processing than other measures used heretofore by psychodynamic and process-experiential researchers. For example, there is compelling evidence that in emotion-focused experiential psychotherapy (Greenberg, 2015), a prototypical sequence of emotional processing moves from feelings of undifferentiated global distress toward feeling self-assertive anger or adaptive, relieving experiences of grief over past hurts and losses (Pascual-Leone, 2017). It could also be that emotional expression at particular moments in therapy—such as discussing specific relationship episodes—may be a stronger marker of good process.

Emotional expression as measured by the EERS also did not attempt to distinguish between more versus less adaptive expressions. We were instead primarily interested in how manifest emotional expression was a marker of good clinical process in this treatment. Defining adaptive versus maladaptive/inhibitory affect in-session is a potentially important but complex task, and different psychodynamic approaches use varying lenses for understanding affect as a treatment mechanism (for a broad survey of heterogeneity in defining adaptiveness of affect in different therapy schools, see the different measures used in the Peluso & Freund et al., 2018 meta-analysis). What is
“adaptive” affect may also strongly differ from patient to patient: for many patients entering PFPP, expressing anger to a romantic partner might be a novel experience *contra* their typical defenses against aggression, whereas for a patient with a more borderline personality organization this same action could more frequently reflect splitting as a defense. To our knowledge, our study is one of the first to examine how patient context (e.g., BPD traits) influences how emotional expression relates to treatment success; such personalized examinations may more specifically elucidate the vagaries of emotional process in-session, even when using scales that aim to identify only adaptive or deep emotional expression.

However, one advantage of our method is that, by definition, for more nuanced emotional experiencing or processing to take place, a basic level of emotional expression must nearly always also be present. Future work on this dataset might use our assay of emotional expression across segments of sessions to orient and focus more detailed research.

This study does not address the question of how a therapist may best affectively engage patients in-session. Process work in psychodynamic therapy indicates that focused confrontation interventions (Town, Hardy, McCullough, & Stride, 2012) or interventions attempting to orient patients to their affects (Ulvenes et al., 2014) tend to be associated with greater emotional experiencing on the part of the patient. Preliminary work in this sample suggests that panic-focused interpretations—found to relate to subsequent outcomes in PFPP (Keefe, Solomonov et al., 2018)—in one segment of treatment may predict higher emotional expression in the next segment of treatment when
patients were discussing an attachment relationship (Keefe, Huque et al., 2018). This research is ongoing.

We also did not examine the extent to which patients engaged in emotional expressions in the CBT and applied relaxation therapies in the trial. In CBT for depression, some studies report that emotional processing in-session may predict better post-treatment outcomes (Aafjes-van Doorn & Barber, 2017), but the role of emotional processing and experiencing is less commonly explored in this family of treatments. In CBT for panic, we might expect emotional expression to be unrelated or positively related to treatment outcomes (e.g., restructuring “hot cognitions”; intensely engaging in interoceptive exposure). For applied relaxation, emotional expression may have a negative relationship to outcomes, given the goal of in-session progressive relaxation. Future work could apply the EERS or other experiencing scales to less affect-focused psychotherapies for panic.

Finally, in psychotherapy research, there has been increasing attention on distinguishing a patient’s tendency to have a given process score across all sessions (i.e., the “between-patients” component) from within-patient changes in process scores, accounting for their average levels of that process (Falkenström et al., 2017). Our secondary analysis partially addresses this critique, suggesting that having higher than one’s average level of emotional expression (Sessions 2, 5, 10) at Session 10 predicted superior subsequent symptom improvement, even when simultaneously modeling the predictive value of one’s average (i.e., “between-patients”) emotional expression. This provides limited evidence that, even among patients who have a relatively higher level of emotional expression across their early therapy, emotional expression over one’s typical
level at a later session is positively prognostic of symptom improvement. Moreover, the majority of variance in emotional expression was not trait-like and consistent across patient sessions, indicating that it is less likely (though not impossible) that the “between-patients” component of emotional expression is primarily driving the observed relations. On the other hand, consistently high levels of a particular process, such as emotional expression, may also reflect specific work within a unique therapeutic dyad, such that the same patient working with a different or more/less skillful therapist would not show such a pattern.

**Conclusion**

Emotional expression in short-term psychodynamic treatment of panic disorder early in the course of therapy predicts greater symptomatic improvements later in treatment, possibly through encouraging insight into the emotional meanings and conflicts surrounding episodes of panic and anxiety. These results require replication.

**References**


Barber, J. P., & Muenz, L. R. (1996). The role of avoidance and obsessiveness in matching patients to cognitive and interpersonal psychotherapy: empirical
findings from the treatment for depression collaborative research program. J Consult Clin Psychol, 64(5), 951-958.


Zimmerman, M., Chelminski, I., Young, D., Dalrymple, K., & Martinez, J. (2012). Does the presence of one feature of borderline personality disorder have clinical
Table 2.1
Descriptive data for baseline characteristics and for symptom change

<table>
<thead>
<tr>
<th>Baseline Measure</th>
<th>Mean (SD) or # (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline PDSS</td>
<td>13.8 (2.9)</td>
</tr>
<tr>
<td>Baseline to Week 5 PDSS Change</td>
<td>-3.9 (4.0)</td>
</tr>
<tr>
<td>Week 5 to Termination PDSS Change</td>
<td>-2.9 (2.9)</td>
</tr>
<tr>
<td>Baseline PSRF</td>
<td>3.3 (1.1)</td>
</tr>
<tr>
<td>Baseline to Week 5 PSRF Change</td>
<td>0.9 (1.4)</td>
</tr>
<tr>
<td>Week 5 to Termination PSRF Change</td>
<td>0.0 (1.2)</td>
</tr>
<tr>
<td>SDS</td>
<td>16.9 (6.1)</td>
</tr>
<tr>
<td>HAM-D</td>
<td>10.4 (3.8)</td>
</tr>
<tr>
<td>Agoraphobia Diagnosis</td>
<td>35 (79.5%)</td>
</tr>
<tr>
<td>Age</td>
<td>37.8 (14.0)</td>
</tr>
<tr>
<td>Gender (Female)</td>
<td>31 (70.5%)</td>
</tr>
<tr>
<td>Concurrent Psychopharmacology</td>
<td>12 (27.3%)</td>
</tr>
<tr>
<td>Age of Panic Onset (years)</td>
<td>26.6 (11.2)</td>
</tr>
<tr>
<td>SCID-II OcPD Criteria</td>
<td>2.1 (1.7)</td>
</tr>
<tr>
<td>SCID-II BPD Criteria</td>
<td>1.0 (1.7)</td>
</tr>
</tbody>
</table>

Note. HAM-D = Hamilton Rating Scale for Depression; IIP = Inventory of Interpersonal Problems; PDSS = Panic Disorder Severity Scale; SCID-II = Structured Clinical Interview for the Diagnosis of Axis-II Disorders; SDS = Sheehan Disability Scale
### Table 2.2

**Example transcripts from sessions rated high/low in emotional expression**

| High Emotional Expression Transcript  
<table>
<thead>
<tr>
<th>(Session Average 1.75)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T: Is your boyfriend being on this nocturnal schedule and maybe not finishing his program a worry for you? I feel that you’re frustrated and anxious in an immediate sense of what’s going on right now, but there’s some concern about getting married, proceeding with that.</td>
</tr>
<tr>
<td>P: No, I mean, the pros outweigh the cons &lt;quiet, restrained tone of voice&gt;.</td>
</tr>
<tr>
<td>T: I think maybe part of what might go on inside is that thinking about these concerns is at times a scary, anxiety-provoking thought… because you want to proceed with marrying him, and the pros outweigh the cons, but he’s not taking care of himself.</td>
</tr>
<tr>
<td>P: Yes… yes, I guess so. &lt;tearing up&gt; Maybe that’s why… I don’t know, I mean, he’s really smart, I just hope that he’ll soon figure what he wants to focus on… I know that in a few years if he didn’t he’d… &lt;begins to fully cry, cries for a few moments&gt;…&lt;therapist offers tissue&gt;… Thank you…</td>
</tr>
<tr>
<td>T: You’re worried about that, and what might happen to you both if that were to happen.</td>
</tr>
<tr>
<td>P: &lt;teary, weepy&gt; I don’t know. He’s such a Type-A perfectionist, I don’t know how he keeps spiraling. I’m so worried about him all the time, it really gets to me. I’m really sad…</td>
</tr>
</tbody>
</table>

| Low Emotional Expression Transcript  
<table>
<thead>
<tr>
<th>(Session Average 0.28)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P: It’s awful. I hate saying goodbyes. I hate driving my son to the station to say goodbye. I hate driving my sister to say goodbye. &lt;neutral tone and facial expression&gt;</td>
</tr>
<tr>
<td>T: How does it make you feel when you do that?</td>
</tr>
<tr>
<td>P: I guess, and this is just pure speculation, but I guess…</td>
</tr>
<tr>
<td>T: What is it that you feel? / P: Um. / T: Sorry I cut you off. / P: No, that’s fine. It makes me feel anxious, it makes me feel sad, um, it’s just very hard, I don’t like it at all &lt;laughter; smiling incongruously to statement&gt;.</td>
</tr>
<tr>
<td>T: Just like the way you feel here.</td>
</tr>
<tr>
<td>P: When I have tears, they could literally be for you and what you’re going through, they could be for us, or it could be, I don’t know. Because, I mean, there would have to be a thought associated with the emotion. A thought that, I’m assuming, triggered the emotion. &lt;continued neutral tone/expression&gt; And I don’t know what that thought is. I’d like to know what that thought is. But, um, the leaving I don’t like. I don’t like goodbyes. Never liked goodbyes. Well, I don’t know if that’s true, but I don’t like goodbyes. &lt;coded small elevation in Anxiety/Distress from increase in physical agitation and pace of speech, but overall still relatively neutral/flat&gt;</td>
</tr>
</tbody>
</table>
Note. Both transcripts are adapted from sessions from the trial, particularly Session 10, for a patient with one of the top 3 ratings for emotional expression at this session versus a patient with one of the bottom 3 ratings for emotional expression. Transcripts have been edited for clarity and to eliminate potentially identifying information.
Figure 2.1. Estimated effect sizes for the relationship between the average level of emotion expression across sessions 2, 5, and 10, and subsequent improvement in panic symptoms as measured by the PDSS. Positive semipartial correlations indicate that higher levels of expression are associated with more subsequent symptom improvement. Bars are 95% confidence intervals.
Figure 2.2. Estimated change in panic symptoms as measured by the PDSS between Weeks 5 to Termination (Week 12), as a function of the average overall emotional expression between sessions 2, 5, and 10, and their interaction with the number of SCID-II borderline personality disorder criteria a patient met at baseline. Increasingly negative values represent greater predicted symptom improvement. All regression variables not displayed in the figure were set to the sample means.
What Mechanism Works for Whom in Panic

Heterogeneity between patients within a DSM-described disorder is widely recognized clinically and, increasingly, empirically. Personalized medicine approaches to psychiatric treatment seek to capitalize on this heterogeneity either by matching patients to the treatment most likely to target their primary and comorbid DSM disorders, or, for psychotherapies, by modifying a given treatment to best address the needs of a given patient (Cohen & DeRubeis, in press).

While some dominant models of psychiatry operate explicitly or implicitly from a latent trait conception—that different symptoms emerge from an underlying, common disease—an alternative conception is that different causal biopsychosocial factors interacting with one another (e.g., in a network) may promote and maintain different patterns of presenting symptomatology (Hofmann, Curtiss, & McNally, 2016; Kendler, Zachar, & Craver, 2011). Treatment selection models attempting to precisely allocate patients on the basis of their personal characteristics to the treatment most likely to benefit them may capitalize on the fact that different treatments activate different mechanisms of change, which might be more or less likely to help a particular patient (DeRubeis, Cohen, et al., 2014; Wallace, Frank, & Kraemer, 2013). If linked factors may cause symptoms to correlate and “hang together” in a disorder-like manner (Kendler et al., 2011), there is no particular reason why two individuals must share the same pathway to manifesting an overt, diagnosable symptom. For example, a longitudinal network analysis of symptoms and experiences among depressed and anxious patients examined how a patient’s reports of these factors at one time point predicted higher or lower levels
of the other measured variables at a subsequent timepoint (e.g., from feelings of anger at T_0 to depressed mood at T_1) (Fisher, Reeves, Lawyer, Medaglia, & Rubel, 2017). While some cross-patient commonalities could be extracted across networks of relationships between the measured psychological variables across time, patients nevertheless exhibited substantive heterogeneity in their networks. For patient A, sleep deprivation could lead most to lowered mood, while for patient B, experiences of anger were most related to later lowered mood. This suggests that patients with the same set of symptoms could differ in the paths that led to their development.

This view has strong implications for mediation analyses that aim identify a potential mechanism by which a treatment leads to the amelioration of symptoms of a disorder. A mechanism can be improvement in a process that led to symptom development (e.g., defense mechanisms; Perry & Bond, 2012) or a gain in a capacity that counteracts a psychopathological process (e.g., cognitive therapy skill use; Strunk, DeRubeis, Chiu, & Alvarez, 2007). If such changes are found to promote symptom relief in some patients and not others, and insofar as this can be predicted on the basis of measurable, individual differences, the term moderated mediation is applicable (Preacher, Rucker, & Hayes, 2007).² Differential responsiveness between patients to change in

² Note that the term “moderated mediation” has been used to characterize instances in which the symptomatic impact of a psychological change is dependent on the type of treatment a patient received (e.g. (Zilcha-Mano et al., 2016)). The theoretical implications of this type of finding are arguably unclear. If a particular psychological change X is helpful in treatment A but not treatment B, this would necessitate that at least one of the following is also true: (a) there is another change encouraged by treatment A but not B that, when combined with change X, allows for X to have a salutary effect; (b) there is another change encouraged in treatment B but not A that attenuates the impact of change X; or (c) the treatment a patient receives affects the meaning of scores (in one treatment or both) on the measure that assesses change X, such that shifts in the scores that index X reflect different constructs in treatments A and B.
psychological mechanisms may indicate more or less precise targeting of their particular deficits or dysfunctions that maintain their symptoms.

**Mechanisms of change in therapies for panic disorder**

We examined the notion of moderated mediation in the context of a 2:2:1 randomized, two-site controlled trial (Milrod et al., 2016) of cognitive-behavioral therapy (CBT; Craske, Barlow, & Meadows, 2000) and panic-focused psychodynamic psychotherapy (PFPP; Busch, Milrod, Singer, & Aronson, 2012) for the treatment of panic disorder with and without agoraphobia. We predicted that individual patient differences would influence the symptomatic impact of change in two putative mediators of panic improvement, catastrophic cognitions concerning bodily sensations (Clark, 1986), and panic-specific reflective functioning (PSRF) concerning the interpersonal-emotional contexts and triggers of panic and anxiety (Rudden, Milrod, Target, Ackerman, & Graf, 2006).

Cognitive-behavioral therapy (CBT) and psychodynamic therapy (PDT) for the treatment of panic are each associated with their own theory of the nature of panic attacks and of the mechanisms that treatment will marshal to ameliorate panic disorder. There is increasing evidence that changes in panic-related cognitions can precede and predict improvements in panic symptoms (Lorenzo-Luaces, Keefe, & DeRubeis, 2016). In CBT models of panic disorder, catastrophic cognitions concerning the ramifications of specific

---

3 The ART condition was not included in the present investigation due to the smaller group of patients randomized to ART, in addition to significantly higher dropout that was related to being more symptomatically and psychosocially impaired at baseline (only in ART), and having worse symptomatic trajectories during treatment. These facts imply that the few treatment-completing ART patients were particularly unrepresentative compared to CBT and PFPP patients.
frightening body sensations, found to be elevated among patients with panic disorder compared to other anxiety disorders (Clark et al., 1997), are considered causal to the emergence of panic attacks (Clark, 1986). Panic patients are conceptualized as frequently interpreting normal body sensations or day-to-day anxiety as indicating somatic problems and distress, which initiates a vicious cycle of increases in arousal and a belief that a catastrophe is imminent, leading to a panic attack. Cognitive restructuring of these cognitions is thought to alleviate panic as it disrupts the cycle of interpretation and arousal. (Teachman, Marker, & Clerkin, 2010), reported that improvements on the Brief Body Sensations Interpretation Questionnaire (BBSIQ), a measure of catastrophic interpretation, predicted subsequent improvements in core panic and agoraphobic symptoms in CBT. Barber and colleagues (under review) replicated this finding, and found that this pattern held also in panic-focused psychodynamic psychotherapy (PFPP; Busch, Milrod, Singer, & Aronson, 2012).

By contrast, psychodynamic theories of panic emphasize how the inability to recognize, experience, and tolerate affects and fantasies that lead to anxiety promotes panic. From a psychodynamic perspective, these causes often derive from unconscious conflict (e.g., repressed anger against a loved one you are fearful of losing if anger is acknowledged). Panic patients report high alexithymia and experiential avoidance (Galderisi et al., 2008; Izci et al., 2014; Parker, Taylor, Bagby, & Acklin, 1993; Tull & Roemer, 2007) and, despite the fact that in the DSM, a panic attack is defined as appearing to “come out of the blue,” emotional-interpersonal anxiety triggers can often be identified (Busch, Shear, Cooper, Shapiro, & Leon, 1995; Chambless, 2010; Klass et al., 2009; Scocco, Barbieri, & Frank, 2007). Moreover, panic patients exhibit
emotionally-avoidant strategies in response to both positively- and negatively-valenced stimuli (Tull & Roemer, 2007), and experimentally inducing use of these strategies promotes distress and physiological arousal (Tull, Jakupcak, & Roemer, 2010). This lack of awareness of triggers, whether facilitated by ignorance or avoidance, can produce the upsetting sense that arousal and consequent panic comes “out of the blue” (Busch et al., 1995). Panic-specific reflective functioning (PSRF), which refers to the tendency of an individual to be aware of and to understand emotional meanings surrounding panic, is typically impaired in patients with PD (Rudden, Milrod, Target, Ackerman, & Graf, 2006). Early improvements in PSRF were found to predict subsequent symptom change in both PFPP and CBT, controlling for early symptom change (Barber et al., under review). These results are consistent with research findings on gains in insight in the context of psychodynamic psychotherapy in treating other disorders, which typically demonstrate that patients who have improved insight over the course of psychotherapy show further or more stable improvements across post-treatment follow-up (Gibbons et al., 2009; Johansson et al., 2010; Kallestad et al., 2010; Ulberg, Amlo, Dahl, & Høglend, 2017).

**Moderation Hypotheses**

For both the BBSIQ and PSRF, we hypothesized that improvements in the BBSIQ and PSRF were expected to be more predictive of subsequent symptom improvement when a patient’s baseline values for that measure were in the more impaired range. We based this on the conjecture that, for patients with these clinical impairments, these core problems likely have a greater role in sustaining the symptom disorder. In other words, we proposed that the presence of impaired values on these measures would indicate the
presence of a problem that must be addressed and compensated for, rather than indicating an area of rigid dysfunction that implies therapist should help bolster and mobilize other, already stronger capacities (Cheavens, Strunk, Lazarus, & Goldstein, 2012). We therefore tested baseline values of each change variable as moderators of the impact of change in both BBSIQ and PSRF.

We also examined the role of co-morbid personality disorder (PersD) in the change processes for these therapies. A little less than 50% of panic patients meet criteria for a SCID-II diagnosable comorbid personality disorder (Friborg, Martinussen, Kaiser, Øvergård, & Rosenvinge, 2013; Keefe, Milrod, Gallop, Barber, & Chambless, in press), and anxiety patients with PersD experience worse trajectories of symptom improvements over time generally (Ansell et al., 2011; Skodol, Geier, Grant, & Hasin, 2014) and in CBT for panic specifically (Porter & Chambless, 2015). Studied mechanisms of change in personality disorder are generally based on personality change (or, in the case of DBT, acquisition of skills) rather than the change processes studied in this investigation (Keefe & DeRubeis, 2018). Psychotherapy with PersD-comorbid patients may sometimes require adaptations in technique (Keefe, Webb, & DeRubeis, 2016). Taken together, we hypothesized that the personality factors giving rise to PersD may also help contribute to panic experiences, and that thus the presence of PersD could attenuate and render less primary the relationship between BBSIQ and PSRF improvements and symptomatic response.

Method

Participants
The present study is a secondary analysis of 138 of 201 patients randomized to three psychotherapy treatments in a two-site randomized controlled trial comparing CBT, applied relaxation training, and PFPP among patients with primary DSM-IV panic disorder with or without agoraphobia. Patients were recruited at Weill Cornell Medical College and the University of Pennsylvania. Participants received study treatment *gratis*. Participants gave informed written consent. Both sites’ institutional review boards approved the protocol, and the study is registered with ClinicalTrials.gov (identifier: NCT00353470).

Patients were included in the trial if they had the spontaneous occurrence of one or more panic attacks for the month before trial entry, and qualified for a primary DSM-IV panic disorder with or without agoraphobia diagnosis determined as per the ADIS-IV version (DiNardo, Brown, & Barlow, 1995). Cross-site agreement on ADIS ratings for panic severity (with a “4” indicating the diagnostic threshold) was excellent (ICC = 1.00). See Milrod et al., 2016 for further details.

Additional ongoing psychotherapy was prohibited. Medications were permitted if stable for at least two months at presentation, and were recorded, held constant, and monitored during the trial. Exclusion criteria were active substance dependence (of less than 6 month’s remission), a history of psychosis or bipolar disorder, acute suicidality, or organic mental syndrome. Additional details on trial design, diagnostician training, and therapy adherence can be consulted in the primary outcome paper (Milrod et al., 2016).

**Therapies**

PFPP is based on the central assumption that panic symptoms have a partly unconscious psychological meaning. It explores feelings and subjective content of panic
episodes, so the patient can begin to address these meanings rather than experiencing conflicts physically as somatic anxiety leading to panic (Busch et al., 2012; Milrod et al., 1997). The therapy helps patients understand and alter core conflicts (e.g., regarding attachment and dependency) to avert future panic vulnerability. CBT for PD followed a modified version of the Panic Control Therapy protocol (Craske et al., 2000), entailing education about panic, correction of maladaptive thoughts about anxiety and body sensations, and both in-session and homework interoceptive exposures to bodily sensations designed to mimic those experienced during panic (Craske et al., 2000). Both psychotherapies comprised 24 sessions delivered twice weekly (12 weeks). Additional information on treatments, including training, supervision, and adherence monitoring, can be found in Milrod et al. (2016).

**Outcome Index**

*Panic Disorder Severity Scale (PDSS; (Shear et al., 1997)).* The PDSS is a diagnosis-based, composite, global rating of panic disorder severity, with acceptable psychometric properties. The PDSS was administered by trained, master’s-level diagnosticians who were uninformed as to treatment condition. Interrater reliability on the PDSS was excellent ($ICC[2,1] = 0.95$). The PDSS was administered five times during treatment: at baseline (Week 0), Week 1, Week 5, Week 9, and termination (Week 12). The mediational analyses in this manuscript concern PDSS change between Week 5 to termination (Week 12), controlling for early PDSS change from Week 0 to Week 5.

**Psychological Change and Moderator Measures**

*Brief Bodily Sensations Interpretation Questionnaire (BBSIQ; Clark et al., 1997).* The BBSIQ is a 7-item measure of catastrophic misinterpretation of panic-related
bodily sensations. Respondents read seven scenarios about bodily sensations, and rank three provided interpretations in the order by which they would most likely come to the respondent’s mind in that scenario. A score for each item is determined as a function of the reported order and the level of catastrophizing of the interpretation, and the item scores are summed. The BBSIQ exhibits psychometric consistency conforming to a one-factor solution, in addition to strong construct validity and test-retest reliability over three months. It showed adequate internal consistency in this trial (Cronbach’s $\alpha = 0.86$).

**Panic-Specific Reflective Functioning (PSRF; Rudden et al., 2006).** PSRF is an interview-based measure that assesses the degree to which a patient can recognize the psychological contributions to panic symptoms. Respondents are queried as to their theory of panic, how it has changed over time, and whether they notice any concordance between panic and particular emotional or interpersonal states. Each response is rated for its psychological mindedness and complexity, and an overall score is assigned based on the item scores. An example of a more impaired PSRF answer might be saying “It’s the heat, the heat brings them on” in response to “reasons for your panic attacks.” A less impaired response to that same question might be “I notice that I get them when I am feeling a lack of control in my personal relationships. I fear that others will leave me, or that I may want to leave them.” PSRF narratives were reliably scored by three trained raters at the Weill-Cornell Medical College ($ICC = 0.80$).

**Structured Clinical Interview for Axis-II Disorders (SCID-II; First, Gibbon, Spitzer, Williams, & Benjamin, 1997).** The Structured Clinical Interview for Axis-II disorders (First et al., 1997) was used to assess the presence of PersD criteria and
diagnoses as defined by DSM-IV. SCID-II criteria counts rather than diagnostic status were employed as a moderator variable for reasons of statistical power and validity (MacCallum, Zhang, Preacher, & Rucker, 2002), following research suggesting that DSM-IV PersD reflects a continuum of illness rather than categorical taxonomies (Harford, Chen, & Grant, 2014; Harford et al., 2013; Haslam, Holland, & Kuppens, 2012).

Trained, independent masters’ level diagnosticians, uninformed to treatment condition administered the interviews. Cross-site inter-rater reliability for number of total PersD criteria scored as present was excellent ($ICC[2,1] = 0.92$).

**Statistical Analyses**

**Missing data.** Due to the nonrandom nature of dropout in this trial, wherein patients with worse symptom improvement trajectories tended to drop out significantly more, and because early psychological changes on the BBSIQ and PSRF were the crucial predictors in this study, we could not use early dropout patients in our investigation ($n = 24$ in CBT/PFPP; 14.8% dropout rate), resulting in a final sample size of $n = 138$ across the two arms. Attempts to multiply impute their outcome and predictor values would likely be statistically biased due to the nonrandom dropout (White, Royston, & Wood, 2011). As a result, 138 CBT and PFPP patients who performed assessments through at least Week 5 (Assessment 3) were the focus of this study.

**Modeling.** Statistical analyses were conducted in SAS using the PROC MIXED command. Repeated measures mixed models were employed on the three PDSS evaluations between Week 5 and Termination (Week 12). Random slopes of PDSS symptom change across the time unit of week of treatment during the evaluation period
were nested in random person-specific intercepts, with an unstructured covariance structure. Covariance structures were allowed to vary between the two sites of the trial. In every model, the baseline PDSS score and early PDSS change (between baseline and Week 5) were entered as covariates.

Within a mixed model, the primary analyses predicted slopes of PDSS change between Week 5 and termination as a function of early change in psychological change mechanisms (i.e., on the BBSIQ or PSRF) between baseline and Week 5. Time was modeled as linear, following the pattern of change detected in the parent trial. Moderators were tested as an interaction between the moderator value, early psychological change, and time ($p < .05$ threshold). Significance of this interaction indicates that the degree to which early psychological change predicts subsequent symptom change is dependent on baseline levels of the moderator variable. Covariates included baseline PDSS score and early PDSS change occurring between baseline and Week 5 (when the mediator was measured the second time). Separate models were run for each of the combinations of two psychological mechanism variables (BBSIQ, PSRF) and three baseline moderator variables (BBSIQ, PSRF, total SCID-II criteria).

We separately tested whether any of the above three potential moderators found to be statistically significant further interacted with site or treatment at $p < .10$. However, no such interactions emerged, indicating that moderation effects were not reliably different across sites or treatments.

**Results**

**Descriptive statistics**
Information on patient demographics and clinical characteristics can be found in Table 1. Patients with better PSRF were somewhat less, but not significantly, likely to have a catastrophic, body-focused interpretation style as measured by the BBSIQ ($r[134] = -0.16, p = 0.064$). Early improvements on the PSRF (Baseline to Week 5) did not correlate with early improvements on the BBSIQ ($r[70] = 0.19, p = 0.111$). The relationship between baseline values in mediator and moderator measures at baseline and early mediator change is reported in Table 2.

**Moderators of the impact of PSRF improvements**

As reported in the Barber et al. (in press) paper on mediation in the trial, patients who experienced more PSRF improvements during the first five weeks of treatment exhibited greater panic symptom improvements over the subsequent seven weeks ($B = -0.47$ [95% CI: -0.82, -0.13], $t[166] = -2.69, p = 0.008$).

Two of the three tested variables moderated the impact of early PSRF improvements on PDSS change. Our primary hypothesis concerning baseline PSRF severity was not supported, as baseline PSRF was not related to the degree to which improvements in PSRF predicted subsequent symptom change ($B = 0.10$ [95 CI: -0.20, 0.39], $t[159] = 0.64, p = 0.520$). Contrary to our expectation, however, patients who engaged in more severe catastrophic interpretations at baseline on the BBSIQ experienced more subsequent symptomatic benefits from early improvements in PSRF ($B = -1.05$ [95% CI: -2.06, -0.03], $t[151] = -2.04, p = 0.043$) (see Figure 1).

Nevertheless, as hypothesized, personality pathology was a significant moderator, as patients with more personality pathology as measured by the SCID-II were less likely to experience symptom improvement after gains in PSRF ($B = 0.06$ [0.00, 0.11], $t[158] =$...
2.00, \( p = 0.047 \)), whereas patients with less personality pathology burden experienced relatively more improvement after gains in PSRF. The critical value of this interaction, at which point early PSRF change was no longer significantly related to subsequent PDSS improvement, was 12 SCID-II criteria, as identified with the Johnson-Neyman technique (Johnson & Fay, 1950; see Figure 2).

**Moderators of the impact of BBSIQ improvements**

Again, as was found in the Barber et al. (in press) paper on overall mediation in this trial, early improvements in the BBSIQ (between baseline and Week 5) predicted greater subsequent symptom change on the PDSS (\( B = 0.13 \ [0.0, 0.25], t[142] = 2.05, p = 0.042 \), controlling for baseline PDSS symptom severity and early symptom improvement.

In a test of our moderation hypothesis, we found that patients with higher BBSIQ scores at baseline experienced more symptom relief following early improvements on the BBSIQ (\( B = 0.27 \ [95\% \ CI: 0.05, 0.49], t[140] = 2.43, p = 0.016 \)). Similar to the pattern shown in Figure 1 for PSRF improvements, patients with relatively lower baseline BBSIQ exhibited no extra symptomatic improvement subsequent to early decreases on the BBSIQ, while those with increasingly higher BBSIQ scores at baseline were predicted to benefit more and more from BBSIQ improvements. As expected, baseline levels of PSRF did not moderate the relationship between BBSIQ change and symptom improvement (\( B = -0.06 \ [95\% \ CI: -0.16, 0.03], t[131] =-1.28, p = 0.203 \)). The test of the hypothesis that a patient’s level of comorbid personality pathology would affect negatively the impact of improvements in catastrophic cognitions on symptomatic benefit was not supported (\( B = 0.01 \ [95\% \ CI: -0.01, 0.03], t[134] = 1.17, p = 0.244 \)).
Discussion

A panic patient’s degree of symptomatic improvement after experiencing a particular psychological change in psychotherapy differs significantly as a function of their psychological characteristics. For certain patients, focus on helping stimulate particular psychological changes might improve the efficacy of their treatment. From a theoretical perspective, these results align with models of psychopathology that posit that different psychological variables may contribute to what appear to be identical symptoms for different patients (Fisher et al., 2017). As such, ameliorating the same putatively pathological process or helping a patient gain an apparently adaptive capacity may be more helpful for some patients than those same changes will be for others. Moreover, our findings align with the viewpoint that patients’ characteristics may meaningfully contribute to which kind of treatment they should receive to optimize their chances of success (Cohen & DeRubeis, 2018).

Possessing a more catastrophic style of interpreting bodily sensations indicated that patients experienced greater subsequent symptom change from early improvements in both catastrophic interpretations as measured by BBSIQ, or insight into the possible psychological meanings surrounding panic as assessed by the PSRF. For these patients, the cognitive model of panic may be especially applicable (i.e., that such interpretations spark and fuel the cycle into panic attacks; Clark, 1986), such that restructuring those interpretations more reliably helps patients thwart the panic cycle. By contrast, patients without a strong tendency to catastrophize may not benefit as much from attempts to reduce this tendency further.
We did not expect to detect the apparent moderating influence of catastrophic cognitive style on the impact of gains in PSRF. In PFPP, tendency to somatize is understood as a defense mechanism. From a psychodynamic perspective, focus on body sensations can arise due to a discomfort with the underlying emotions and fantasies contributing to physiological arousal—somatization avoids direct acknowledgement of these unacceptable, sometimes partially unconscious contents (Busch et al., 2012; Busch et al., 1995). Shifting from understanding anxiety and arousal as implying somatic danger to deriving from central psychological meanings may help patients to recognize and work with those meanings rather than to ascribe arousal to frightening somatic causes, especially given a heightened sense of catastrophe.

By contrast, patients with lower versus higher insight about possible psychological meanings of their panic (i.e., PSRF) evidenced equal apparent benefit from either type of psychological change. Ergo, patients beginning psychotherapy with little or no insight as to possible emotional or interpersonal links to their panic appear to nonetheless benefit from work fostering early improvements in PSRF.

Taken together, these findings imply that patients with a highly catastrophic, body-focused interpretation style may need either to ameliorate this cognitive style or gain insight into their emotional-interpersonal triggers in order to experience greater relief from panic disorder. Both types of changes may commonly reflect patients recognizing that bodily sensations do not indicate danger and catastrophe, but rather that they are harmless (BBSIQ improvement) or may signify other psychological meanings (PSRF improvement). This may also reflect that, while BBSIQ indexes severity of a cognitive style that specifically distinguishes and potentially causes panic disorder (Clark
et al., 1997), low levels of PSRF may reflect a lack of understanding of symptom links that can be gained to help with panic, rather than a deficit particular to panic patients that must be improved. It is possible that individual differences concerning the severity of more general emotional awareness deficits (e.g., alexithymia; experiential avoidance) documented in panic patients may be a better moderator for panic mediational models. Unfortunately, it was not possible to meaningfully investigate the case of simultaneous improvement in both constructs, as the type of changes a patient experienced was dependent on the therapy they received (CBT vs. PFPP), and simultaneous improvements were thus uncommon.

Patients with more personality pathology burden benefitted less from improvements in PSRF, while the converse was true for patients with few such traits. Indeed, patients without any personality pathology whatsoever exhibited the greatest benefit from PSRF improvements. Nevertheless, in this subsample, there was no indication that patients meeting more PersD criteria had significantly worse symptomatic outcomes in PFPP as compared to CBT. Furthermore, in this trial PFPP was superior to CBT in improving comorbid severe PersD symptoms on the SCID-II (Keefe et al., 2018). It may be that PFPP’s capacity to treat severely PersD-comorbid patients is due to other psychological changes targeted by psychodynamic therapies, such as improvements in defensive functioning or attachment security (Johansson et al., 2010; Keefe & DeRubeis, 2018; Perry & Bond, 2012; Rossouw & Fonagy, 2012). On the other hand, BBSIQ improvements were predictive of symptom change regardless of patients’ personality pathology, which is consistent with an account that PersD is not a clear contraindication to cognitive restructuring for panic symptom relief.
This investigation was not designed to elucidate specific techniques or therapeutic foci within the two treatments that best promote the psychological changes we examined. Most psychological treatments consist of multiple interventions targeting multiple types of changes. CBT for panic disorder includes teaching relaxation (a coping skill), cognitive restructuring concerning catastrophic thoughts about the body, and both interoceptive exposure to feared bodily sensations and *in vivo* exposures to phobic, panic-engendering spaces (facilitating extinction and potentially fueling cognitive change). Panic-focused psychodynamic psychotherapy also entails several foci, including the exploration of dysregulated attachment, the interpretation of transference and of defenses, and efforts to enhance the patient’s awareness and tolerance of emotions and fantasy life. Different psychological interventions within a given therapy may best encourage the types of psychological changes most likely to help a particular patient (Keefe et al., 2016; Sasso, Strunk, Braun, DeRubeis, & Brotman, 2015). These kinds of questions require further investigation.

**Limitations and future directions**

Due to the nonrandom nature of dropout in this trial such that patients with worse trajectories of symptom change were more likely to drop out, only patients with observed data up to at least Week 5 were analyzed in this investigation (14.8% of CBT and PFPP patients randomized to treatment).

For this analysis, we had the option to explore either a wide range of potential moderators of proposed change mechanisms, or to limit our analyses to a select few on the basis of theoretical predictions. In the interest of limiting Type I error, we elected to test a smaller set of moderators. It is possible that other individual traits exist that would
help envisage how each type of psychological change might predict subsequent symptomatic improvements. In addition, only two psychological changes relevant to panic disorder were examined in the primary mediator analyses for this trial (catastrophic interpretations and PSRF), whereas other putative mediators may exist, such as self-efficacy (Fentz et al., 2013) and improvements in attachment security (Milrod, 2015), neither of which were assessed in this trial. In addition, as this is the first investigation of its kind in panic treatment, these relationships should be explored in other treatment trials with relevant data.

Despite our goals to identify which psychological changes might be most important for individual patients, our study takes a nomothetic, between-patients approach to examining “what mechanism works for whom.” An alternative, personalized approach would be to identify these patterns for individual patients, and to examine the differential effects of attempting to intervene upon the core problems that sustain an individual’s panic (Fisher & Boswell, 2016). For example, if a patient’s tendency to catastrophize bodily sensations was observed to reliably predict their experiences of panic and anxiety across time, cognitive interventions targeting that tendency may be found to be more efficacious in improving panic than other cognitive-behavioral interventions (e.g., relaxation training).

References


Table 3.1

Descriptive data for baseline characteristics and for symptom change

<table>
<thead>
<tr>
<th>Baseline Measure</th>
<th>Mean (SD) or # (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline PDSS</td>
<td>13.8 (3.6)</td>
</tr>
<tr>
<td>Week 5 PDSS</td>
<td>9.6 (4.2)</td>
</tr>
<tr>
<td>Termination PDSS</td>
<td>6.5 (4.3)</td>
</tr>
<tr>
<td>Week 5 to Termination PDSS Change</td>
<td>-3.4 (3.5)</td>
</tr>
<tr>
<td>Baseline PSRF</td>
<td>3.4 (1.4)</td>
</tr>
<tr>
<td>Baseline to Week 5 PSRF Change (CBT / PFPP)</td>
<td>0.0 (1.0) / 0.5 (1.3)</td>
</tr>
<tr>
<td>Baseline BBSIQ</td>
<td>1.9 (0.4)</td>
</tr>
<tr>
<td>Baseline to Week 5 BBSIQ Change (CBT / PFPP)</td>
<td>-0.2 (0.4) / 0.0 (0.3)</td>
</tr>
<tr>
<td>SCID-II Criteria Count</td>
<td>7.5 (6.4)</td>
</tr>
<tr>
<td>Agoraphobia Diagnosis (% Yes)</td>
<td>112 (81.2%)</td>
</tr>
<tr>
<td>Age</td>
<td>40.00 (13.5)</td>
</tr>
<tr>
<td>Gender (Female)</td>
<td>86 (58.1%)</td>
</tr>
<tr>
<td>Ethnicity (Hispanic)</td>
<td>23 (16.7%)</td>
</tr>
<tr>
<td>Race (Caucasian)</td>
<td>101 (73.2%)</td>
</tr>
<tr>
<td>Concurrent Psychopharmacology</td>
<td>42 (30.4%)</td>
</tr>
</tbody>
</table>

*Note.* BBSIQ = Brief Bodily Sensation Interpretation Questionnaire; PDSS = Panic Disorder Severity Scale; PSRF = Panic-Specific Reflective Functioning; SCID-II = Structured Clinical Interview for the Diagnosis of Axis-II Disorders
Table 3.2
*Correlations of baseline moderator values to improvement in BBSIQ and PSRF from Baseline (Week 0) to Week 5 of treatment.*

<table>
<thead>
<tr>
<th>Baseline BBSIQ</th>
<th>Baseline PSRF</th>
<th>Baseline SCID-II</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBSIQ Improvement</td>
<td>0.49***</td>
<td>-0.23*</td>
</tr>
<tr>
<td>PSRF Improvement</td>
<td>0.06</td>
<td>-0.57***</td>
</tr>
</tbody>
</table>

*Notes. Higher scores on the BBSIQ indicate a more pronounced tendency to catastrophically interpret body sensations. Higher PSRF scores indicate greater symptom insight. * = p<.05; *** = p<.001*
Figure 3.1. Predicted degree of Week 5 to Termination (Week 12) symptom change attributable to early (Baseline to Week 5) improvements in PSRF, as a function of baseline severity of catastrophic interpretive style as measured by the BBSIQ (interaction $p = 0.043$).
Figure 3.2. Predicted degree of Week 5 to Termination (Week 12) symptom change attributable to early (Baseline to Week 5) improvements in PSRF, as a function of baseline level of personality disorder pathology as measured by number of criteria met on the SCID-II interview (interaction $p = 0.047$).
APPENDICES
Appendix A

Additional notes on statistical analyses (Study 1)

Random forest imputation. Random forest imputation is a single-dataset imputation method that can produce nominally equivalent or more accurate imputations compared to standard multiple imputation by chained equations, and can natively handle simultaneous continuous/categorical variable imputation and datasets with interactions and nonlinear relationships between variables (Liao et al., 2014; Stekhoven & Bühlmann, 2011; Waljee et al., 2013). Random forest algorithms produce nested prediction or decision trees based on splits in continuous or categorical predictor variables (King & Resick, 2014). Each individual tree selects a bootstrapped subset of the main data set, and at each potential split point allows a random subset of possible predictors to be used, until there are no more possible splits to be made. This process is repeated across many different forests of trees, wherein predictions for the missing variables are iteratively updated and improved with successive models, until a stopping criterion is reached. A given imputed data value is based on the average prediction made across several decision trees built from multiple subsets of the data and potential predicting variables.

All baseline data, technique ratings from a rated session not used in this manuscript (Session 5), and termination and pre-to-post treatment change scores on the PDSS, Sheehan Disability Scale, Inventory of Interpersonal Problems, and Hamilton Rating Scale for Depression were allowed to predict missing data in the set, in addition to all rated process measures and their constituent items.

Additional references


Appendix B

Additional checks on robustness of obtained findings (Study 1)

Controlling for therapist effects

The packages “lme4” and “lmerTest” in the R statistical language were used (Bates et al., 2016; Kuznetsova, Brockhoff, & Christensen, 2016). Therapists were specified as a 2^{nd}-level random effect in which patients were nested. When therapist effects were examined without fixed effects, there was no statistically significant random effect of therapist on PDSS symptom change (ICC_\text{T Week 1 to 5} = 0.05, \text{p} = 0.600; ICC_\text{T Week 5 to 9} = 0.00, \text{p} = 1.000; ICC_\text{T Total Pre-Post Change} = 0.08, \text{p} = 0.300). The lack of estimable effect or statistical significance is likely due to the low sample size and poor patient/therapist ratio leading to issues in estimation. On the other hand, the estimate of reliable therapist-level variance for total pre-post symptom change (8\%) was comparable to the range of 6-9\% often reported in the therapist effects literature (Lutz & Barkham, 2015).

For PDSS symptom change between Weeks 1 and 5 and Weeks 5 to 9, we conducted an exploratory analysis to examine whether our technique-outcome findings differed notably when simultaneously modeling therapist effects, using the same model specifications as the primary analyses (e.g., outcome indices, covariates, technique terms) but as applied to a mixed model framework. Unremarkably, there were no \text{p}-value shifts across the significance level of \text{p} <.05, and effect size estimates were comparable.

Functioning and interpersonal outcomes

In our primary analyses, we found that panic-focused interpretations at Session 10 predicted panic symptom improvement on the PDSS occurring subsequent to Session 10.
One possibility is that panic-focused interpretations are conducive to symptomatic improvements, but that they do not help other aspects of functioning. As a secondary, *post-hoc* analysis, we explored whether these findings extended to improvements on the Sheehan Disability Scale (Leon, Shear, Portera, & Klerman, 1992), a measure of psychosocial dysfunction due to mental illness, and the Inventory of Interpersonal Problems. Unlike the PDSS, which was assessed several times during the trial, the SDS and IIP were administered only pre- and post-therapy.

Our overall modeling strategy was very similar, with the following exceptions: (1) the predicted variable was the final SDS or IIP score; and (2) baseline levels of the SDS or IIP were included as a covariate in lieu of baseline PDSS. We retained the degree of PDSS symptom change between baseline and Week 5 as a covariate, such that early symptomatic progress occurring during the process measurement period could not contribute to the apparent relationships (or lack thereof) between process variables and non-symptomatic outcomes measured at treatment termination.

Panic-focused interpretations at Session 10 predicted significantly lower IIP (*B* = -0.27 [95% CI: -0.45 to -0.09], *SE* = 0.09, *t*[58] = -2.99, *p* = 0.004) scores at treatment termination, and, at the level of a statistical trend, lower SDS scores (*B* = -2.42 [95% CI: -5.21 to 0.38], *SE* = 1.40, *t*[59] = -1.73, *p* = 0.089). By contrast, panic-focused interpretations at Session 2 did not relate to improvements on either the IIP (*B* = -0.20 [95% CI: -0.47 to 0.08], *SE* = 0.14, *t*[58] = -1.44, *p* = 0.156) or the SDS (*B* = -2.64 [95% CI: -6.29 to 1.01], *SE* = 1.82, *t*[59] = -1.45, *p* = 0.152).

These findings are consistent with a perspective that panic-focused interpretations at mid-therapy may also encourage improvements in interpersonal functioning, in
addition to panic symptom response. However, it should be noted that these analyses are less analytically rigorous than our primary analyses, as temporal precedence between technique use and symptom change cannot be assured because the SDS and IIP were measured pre- to post-treatment.

**Additional References**


Appendix C

Supplementary tables (Study 1)

Table S1
*Examples of Panic-focused Clarification and Interpretation*

<table>
<thead>
<tr>
<th>Technique</th>
<th>Example</th>
</tr>
</thead>
</table>
| **Panic-focused Clarification** | Patient: It’s just that I know I might be having a heart attack whenever I get that anxious. Everything in my chest is so tight, I’m so wound up, I’m so clenched, I just feel like my heart is going to give out and die.  
Therapist: You think this even though you’ve been to specialists several times, and all tests show you as being in not only OK, but apparently remarkably good health for your or any age.  
Patient: Yes, yes, but what if they’re wrong? I’m the one who pays the price.  
Therapist: Jane, it’s interesting. Somehow, it seems as though it’s easier for you to believe you have a hidden, undiagnosable heart defect impenetrable to modern medical science, than think that there may be any psychological reason you may be panicking. Part of you would rather feel like you’re on the razor’s edge than think about whatever it is you’re avoiding. |
| **Panic-focused Interpretation** | Therapist: When you finally saw [your boyfriend], did you feel any of that disappointment or anger you described to me last session, about his lack of communication?  
Patient: Maybe just a little. But, I know that he’s been so busy recently, I can’t consider myself the most important thing all the time in his life. That would be so selfish of me.  
Therapist: You’re putting in all this grand effort to welcome him, at the same time he’s been treating you poorly. I wonder if it is easier for you to try to feel excited, rather than experience your disappointment, which might lead you to have to reevaluate the relationship. I think that considering the relationship feels very dangerous to you, you might break up. I wonder if trying to ignore that anger and concern about separation leaves you feeling anxious and panicky. |
Appendix D

Supplementary tables (Study 2)

Table S2.
Correlations between emotional expression ratings within the same 5-minute segment.

<table>
<thead>
<tr>
<th></th>
<th>Overall Expression</th>
<th>Grief/Sadness</th>
<th>Anger/Assertion</th>
<th>Anxiety/Distress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Expression</td>
<td>1</td>
<td>0.48***</td>
<td>0.38***</td>
<td>0.65***</td>
</tr>
<tr>
<td>Grief/Sadness</td>
<td>1</td>
<td>0.10†</td>
<td>0.24***</td>
<td></td>
</tr>
<tr>
<td>Anger/Assertion</td>
<td></td>
<td>1</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Anxiety/Distress</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

† = p < .10; *** = p < .001

Table S3.
Average levels of emotional expression per 5-minute segment across sampled sessions.

<table>
<thead>
<tr>
<th></th>
<th>Session 2 (M / SD)</th>
<th>Session 5 (M / SD)</th>
<th>Session 10 (M / SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Expressivity</td>
<td>1.10 (0.43)</td>
<td>0.99 (0.36)</td>
<td>1.04 (0.39)</td>
</tr>
<tr>
<td>Grief/Sadness</td>
<td>0.33 (0.29)</td>
<td>0.19 (0.17)</td>
<td>0.17 (0.19)</td>
</tr>
<tr>
<td>Anger/Assertion</td>
<td>0.30 (0.21)</td>
<td>0.32 (0.26)</td>
<td>0.36 (0.29)</td>
</tr>
<tr>
<td>Anxiety/Distress</td>
<td>0.68 (0.38)</td>
<td>0.66 (0.32)</td>
<td>0.74 (0.37)</td>
</tr>
</tbody>
</table>

Note. As estimated in a mixed model using the R package lme4 (Bates et al., 2017), grief/sadness expression was estimated to significantly decrease within-person as patients progressed from sessions 2 to 5 to 10, coded as 0, 1, and 2 time indicators (B = -0.08 [95% CI: -0.13 to -0.04], SE = 0.02, t[74.8] = -3.67, r = 0.39, p < .001). There were no other reliable patterns of change in emotional expression experienced by the average patient.
Table S4.
Sources of variance in emotional expression per 5-minute segment.

<table>
<thead>
<tr>
<th></th>
<th>Patient-Level Variance</th>
<th>Session-Level Variance (in Patient)</th>
<th>Unique/Residual Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Expressivity</td>
<td>29.8%</td>
<td>17.9%</td>
<td>52.2%</td>
</tr>
<tr>
<td>Grief/Sadness</td>
<td>15.2%</td>
<td>14.7%</td>
<td>70.1%</td>
</tr>
<tr>
<td>Anger/Assertion</td>
<td>10.0%</td>
<td>20.6%</td>
<td>69.3%</td>
</tr>
<tr>
<td>Anxiety/Distress</td>
<td>23.6%</td>
<td>27.8%</td>
<td>48.6%</td>
</tr>
</tbody>
</table>

Note. Patient-level variance indicates the degree to which a given patient was reliably more versus less expressive across all of their observed ratings and sessions. Session-level variance indicates the degree to which patients sometimes have entire sessions that are more versus less expressive. Unique/residual variance indicates variability in expression that is not predicted by the fact that a rating comes from a given patient, or which session of a given patient is being rated. Variance components were extracted from a mixed model (Bates et al., 2017) including a random effect of session nested within a random effect of patient.