Voluntary Breath Holding in Panic and Generalized Anxiety Disorders

Roth, Walton T. MD; Wilhelm, Frank H. PhD; Trabert, Werner MD

From the Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, California, and Department of Veterans Affairs Medical Center, Palo Alto, California.

Address reprint requests to: Walton T. Roth, MD, VA Palo Alto Health Care System (116F-PAD), 3801 Miranda Ave., Palo Alto, CA 94304-1207. E-mail: wtroth@stanford.edu.

Received for publication November 20, 1997; revision received May 14, 1998.

Outline

• Abstract
• INTRODUCTION
• METHODS
  • Subjects
  • Apparatus
  • Procedure
  • Data Analysis
• RESULTS
  • Panic Attacks
  • Respiration Questionnaire
  • Breath Holding Performance
  • Mood Ratings
  • Symptoms During the Breath Holding Procedure
  • Physiological Measures Before the First Breath Holding
  • Physiological Measures During Breath Holdings
  • pCO2 Before and Immediately After Each Breath Holding
  • Physiological Measures After Each Breath Holding
• DISCUSSION
• REFERENCES

Abstract

Objective: Because breath holding causes arterial pCO2 to increase, we used it to test the hypothesis that in panic disorder (PD) a biological suffocation monitor is pathologically sensitive.

Method: Nineteen patients with PD, 17 with generalized anxiety disorder (GAD), and 22 normal controls took deep breaths on signal and held them until a release signal was given 30 seconds later. This was repeated 12 times separated by 60-second normal breathing periods.

Results: PD patients reported having had in the past more symptoms of shortness of breath when anxious, and more frequent frightening suffocation experiences than the other groups. However, increases in self-rated anxiety between periods of normal breathing and periods of breath holding were similar in all three groups. Skin conductance, blood pressure, and T-wave amplitude reactions to breath holdings were also similar, but heart rate acceleration upon taking a deep breath was greater in GAD patients. Before and after individual breath holdings, end-tidal pCO2 was lower in PD patients than in normal controls; GAD patients were intermediate. Inspiratory flow rate did not differ between groups.

Conclusions: Our physiological results provide no direct support for an overly sensitive suffocation alarm system in PD. Lower pCO2 may be due to anxiety causing hyperventilation in patients prone to panic.

Key words: panic disorder, generalized anxiety disorder, respiration, breath holding, suffocation, autonomic nervous system. PD = panic disorder; GAD = generalized anxiety disorder; HR = heart rate; ANOVA = analysis of variance.
INTRODUCTION

Klein [1,2] has hypothesized that in PD a biological suffocation monitor is pathologically sensitive and easily misfires leading to sudden respiratory distress, hyperventilation, panic, and an urge to flee. Recent evidence includes findings that the physiological symptom with the largest effect size for differentiating clinical and nonclinical panic is suffocation [3], and that a 0.5-mg/kg dose of the respiratory stimulant doxapram was able to trigger panic attacks in 75% of PD patients but in only 12.5% of normal controls [4]. In contrast, psychological theories of PD give a primary role to the fearful misinterpretation of bodily sensations, particularly those that are associated with anxiety [5-7]. Fear of physical sensations is a good predictor of fear responses to hyperventilation and CO2 challenge in anxiety disorder patients [8] and in nonclinical subjects [9,10]. The potential importance of psychological factors was demonstrated in the doxapram study cited above, in that a cognitive intervention was able to attenuate (although not to obliterate completely) the effects of the drug on the PD patients.

A simple, logical way to test Klein’s hypothesis is to compare reactions to breath holding in PD patients with other groups, because breath holding will cause arterial pCO2 to increase and stimulate the relevant physiological sensors. Three groups of researchers have used this approach. Zandbergen et al. [11] had patients with PD, patients with other anxiety disorders, and normal controls inhale a deep breath of 50% O2 and 50% N2 and then hold that breath as long as they could. PD patients held their breath for a mean of 34 seconds, which was significantly different from normal controls (74 seconds), but not from patients with other anxiety disorders (49 seconds). The inhalation and breath holding procedure was repeated 3 minutes later after 1 minute of hyperventilation with the same results. Increase in self-reported anxiety with breath holding was small and not significantly different between the groups. Asmundson and Stein [12] had patients with PD, patients with generalized social phobia, and normal controls hold their breath three times after a normal exhalation for as long as possible. Patients with PD held their breath for a shorter time (mean = 16 seconds) than either comparison group (21 seconds for social phobics and 21 seconds for normal controls), but the groups did not differ in heart rate, respiratory sinus arrhythmia, respiratory frequency, or endtidal pCO2 before or after trials. Van der Does [13], had patients with PD, depressed patients, and normal controls hold their breath as long as they could, once after deep inhalation and three times after normal exhalation. As with Zanbergen et al. [11], PD patients held their breath shorter than controls after deep inhalation, but contrary to Asmundson and Stein [12], not after normal exhalation.

Our experiment differed from the ones above in several ways. First, we used patients with GAD as a clinically anxious comparison group, because GAD patients have somatic anxiety symptoms similar to those of PD patients. Second, we had subjects complete a questionnaire that asked about past experiences with, and attitudes toward, shortness of breath. Third, we asked the subjects to hold their breaths for a signaled time period of 30 seconds. Specifying a moderate time that all subjects were expected to attain was an attempt to make uniform the physical aspect of the stressor to which we measured subjective and physiological responses and to reduce motivational effects, which cause considerable variation when people are told to hold their breath as long as they can [13-15]. Fourth, by asking subjects to repeat breath holding 12 times, we were better able to look for habituation or sensitization effects [16,17]. Fifth, more measurements were taken, including subjective emotional responses and a variety of cardiovascular, electrodermal, and respiratory variables.

Our hypotheses were as follows: If the suffocation false alarm hypothesis is true, PD patients should report more fears related to breathing. They should have the largest psychological and physiological anxiety responses to breath holding, especially higher heart rate and skin conductance levels [18] during the breathing holding themselves. After individual breath holdings PD patients should show signs of greater respiratory drive and continued signs of greater autonomic activation, inasmuch as physiological anxiety reactions would be unlikely to vanish immediately upon temporary resumption of breathing, especially because breath holding will be soon repeated. However, if psychological causes predominate, we would expect fewer differences between groups, inasmuch as breath holding lacks some of the threat of the other suffocation experiences: Subjects have a ready explanation for their sensations and know that they can turn them off at any time by resuming breathing. Control is known to attenuate the effects of anxiety provocations in PD patients [19], although reassuring explanations do not always block the effects of lactate [20] or CO2 effects [21] challenges.

METHODS

Subjects

Nineteen patients with PD, 17 with GAD, and 22 normal controls were recruited by advertisement. Patients were candidates for clinical drug trials and were not paid for their participation, whereas controls were paid. Advertisements for controls asked for volunteers not suffering from excessive anxiety or stress. Diagnosis of patients and exclusion of anxiety disorders and other Axis I disorders in controls was made by the Structured Clinical Interview for DSM-III-R [22]. Eligible subjects signed a consent form before participation. Three of the GAD patients had a history of PD, but their PD was in full remission and they had no agoraphobia. Three of the PD patients currently met criteria for GAD in addition to PD. On the structured interview, PD patients had a mean (SD) panic attack severity of 1.9 (0.7) and agoraphobic avoidance severity of 2.3 (0.9) on scales where 4 is "severe" and 0 is "in full or partial remission." Three of the PD patients, two of the GAD patients, and one of the normal controls had a history of a major depressive episode, but none was having such an episode at the time of testing. All patients and normal controls denied taking psychoactive or cardiovascularly active medication in the 2 weeks before testing, although this was not verified by
drug screens. None of the subjects reported current respiratory disease. As noted in Table 1, groups were successfully selected not to differ significantly in age or in the proportion of women.

Table 1. Description of Groups: Mean (SD)

<table>
<thead>
<tr>
<th></th>
<th>PD</th>
<th>GAD</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>40.8 ± 9.2</td>
<td>41.3 ± 9.5</td>
<td>42.0 ± 8.5</td>
</tr>
<tr>
<td>Proportion of women (%)</td>
<td>47.9</td>
<td>47.0</td>
<td>51.6</td>
</tr>
</tbody>
</table>

Before testing subjects filled out several questionnaires, including the Agoraphobic Cognitions Questionnaire [23], the Body Sensations Questionnaire [23], the Beck Depression Inventory [24], and the State-Trait Anxiety Inventory (STAI) Trait Form [25]. As can be seen in Table 1, the expected group differences were present. In addition, subjects filled out a Respiration Questionnaire, devised by the authors, comprising the 16 questions outlined in Table 2. For example, the first question was "How often have you had these experiences?" followed by a list of five items, the first of which was "1. Shortness of breath when anxious" followed by the numbers 1 to 4, above which were the anchors never/seldom/sometimes/often. Subjects were to circle the appropriate number.

Table 2. Respiration Questionnaire: Mean (SD)

<table>
<thead>
<tr>
<th>Question</th>
<th>PD</th>
<th>GAD</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often have you had these experiences?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Shortness of breath when anxious</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. A feeling of being cut off from others</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. A sense of coming towards something</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. A feeling of being trapped</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. A feeling of losing control</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Immediately before testing, PD patients reported on a questionnaire item having had a mean (SD) of 2.6 (2.1) panic attacks in the past 2 weeks; 14 of the 19 PD patients reported attacks in that period. Surprisingly, 3 of the 17 GAD patients reported panic attacks also: one reported five and two reported one. Note that subjects were not asked how many symptoms accompanied these attacks or other details to verify that the attacks met DSM criteria. The STAI-State Form was used in addition to our Mood Questionnaire (described below) to quantify state anxiety immediately before testing (20 items rated 1-4 allow a maximum score of 80) [25].

Apparatus

Eight physiological channels were recorded using the Viatport I (Becker Engineering, Karlsruhe, Germany): 1) An electrocardiogram; 2) two channels of respiration from Respibands (Respitrace Corporation, Ardsley, NY) placed around the upper thorax centered on nipples and around the abdomen between the xiphoid process and the navel, and calibrated by having subjects breathe into a spirometer; 3) a continuous arterial blood pressure waveform from the index finger by means of the Finapres 2300 (Omeda, Madison, WI) system; 4) skin conductance using a constant 0.5 V and isotonic electrode paste; 5) expiratory pCO2 measured continuously by a calibrated infrared capnograph (Datex B, Puritan-Bennett Corporation, San Ramon, CA) into which air was drawn with a flow rate of 150 ml/minute through a 1.2-mm diameter plastic tube ending 2 to 4 mm inside the left nostril (subjects were instructed to breathe only through the nose); and 6) body movement sensed by accelerometers attached to the right thigh and shoulder.

Procedure

Subjects sat upright in a comfortable chair during testing in a room held between 21 and 23° C. Time of day was matched between the three groups. Subjects had had at least 45 minutes to adjust to their surroundings before the breath holding procedure began. It was introduced by written instructions: The subjects were informed that when they heard a single tone, they were first to quickly breathe out and then take a deep breath and hold it until they heard a double tone, which signaled that they were to resume breathing normally. All breathing was to be through the nose, not the mouth. The interval between single and double tones (the breath holding period) was 30 seconds, and the interval from a double tone to the next single tone (the normal breathing period) was 60 seconds. Twelve breath holding periods (trials) separated by 11 normal breathing periods occurred over 18 minutes. In addition, the 30 seconds before the first breath holding and the 30 seconds after the last breath holding were included as normal breathing periods for purposes of data analysis. There was no mention in the instructions that anxiety or panic attacks might occur during the procedure nor provisions for its premature termination.

Subjects rated their subjective emotional state immediately before and after (but not during) the breath holding procedure on our Mood Questionnaire, which consisted of four items: "anxious," "excited," "alert," and "in control." Using scales from 0 (not at all) to 10 (extremely strong) for each of the four items, subjects made ratings for four different time points. Before the procedure, they rated their current mood (this rating will be referred to subsequently as before). After the procedure they rated their current mood again (after) and, retrospectively, their average emotional state while breathing normally during the procedure (breathe) and their average while breath holding (hold).

The last item on this questionnaire asked whether subjects had "attacks during the test when you suddenly felt more frightened, anxious, or extremely uncomfortable," and how many such attacks they had had. If they answered yes, subjects were asked to fill out Symptom Questionnaire A for this episode (or the worst episode) so that we could determine whether it had resulted in four or more symptoms. Subjects for whom that was the case were referred to below as panicickers. Symptom Questionnaire A was a list of the 13 symptoms comprising the diagnostic criteria for a panic attack in the Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition-Revised (DSM-III-R) [26], after each one of which subjects circled "yes" or "no."
Finally, all subjects filled out Symptom Questionnaire B asking them "how you felt during any part of the last test." It listed the 13 DSM-III-R symptoms with "heart racing, pounding, or skipping" combined with "chest pain or pressure" to give a total of 12 items. Subjects rated each item on a 5-point scale from 0 = not at all to 4 = to a high degree.

Data Analysis

In off-line analyses, measurements derived from the channels were synchronized in time by the beginning of breath holdings for each subject and trial. The computer detected the onset of breath holding as a leveling off to zero of inspiratory change in the respiratory channels after a single signal tone. This allowed analysis to be divided into sets of three time periods: before (the last 30 seconds of normal breathing), during (mean of 28 seconds from the end of inspiration to the beginning of expiration), and after each breath holding (the first 30 seconds of normal breathing). To avoid transition effects, only data from the first 20 seconds of the before period and the last 20 seconds of the after period were used.

HR was analyzed with a program that computed estimates of instantaneous heart rate for every 0.5-second interval [27]. Respiratory sinus arrhythmia was calculated by spectral analysis [28]; 20-second epochs of heart rate data were linearly detrended, and then analyzed by the Fast Fourier transform using Welch's method. Respiratory sinus arrhythmia was estimated as the spectral power between 0.14 and 0.50 Hz. T-wave amplitude was determined by first locating the maximum value from 50 to 300 msec after an R wave. The difference between this maximum and the mean of the last 50 msec before the occurrence of the present R wave was taken as T wave amplitude [29]. Systolic and diastolic blood pressure were local maxima and minima in the continuous arterial blood pressure waveform. Movement channels were rectified, integrated, and added together. Skin conductance was analyzed as mean levels.

The first step in analyzing the respiration channels was to convert the raw signals of the two bands to a calibrated instantaneous lung volume established by a spirometer calibration procedure conducted before the tests began [30]. We calculated a wide variety of respiratory parameters including respiratory rate, tidal volume, minute volume, inspiratory flow rate, inspiratory time, expiratory time, duty cycle (inspiratory time/total time), pause after inspiration, pause after expiration. Onset of inspiration, expiration, and pauses were defined by slope criteria (flow rates exceeding 100 ml/min defined onsets). Our analytic approach followed that introduced by Gautier [31], which assumes that respiration is under the control of a central "driving" mechanism turned on and off by a cyclic "timing" mechanism [32]. We estimate that after calibration the discrepancy between concurrent Respitrace and spirometric volume measures is within +/- 10%.

End-tidal pCO₂ was determined as the level at which pCO₂ stopped rising at the end of an expiration (final maximum). Expirations with a low percentage of alveolar air can be recognized by the CO₂ waveform not reaching a plateau [33], and these were excluded. We required plateaus to be present in 50% of breaths in a measuring epoch or it was discarded. For measurement of pCO₂ in the first expiration after breath holding, the second expiration value was substituted in the few breath holdings in which it was greater. Presumably, in these cases the first expiration did not go completely through the nose. Such cases were not restricted to one group. Our statistical analysis for pCO₂ was based on data for 11 PD patients, 14 GAD patients, and 14 normal controls. Besides failures to reach a plateau, in 3 PD patients and 2 normal controls the nasal tube did not remain in place, and in one PD patient and 3 normal controls data recording was faulty.

Heart rates during breath holding were determined from ensemble averages synchronized with breath holding onsets and offsets. Offsets were set to +28 seconds (median breath holding time) from onsets, which resulted in the discarding of some data when breath holding was longer than 28 seconds, and the filling in of interpolated data when breath holding was shorter than 28 seconds. Trials in which breath holding was less than 20 seconds (1.8% of all trials) were discarded.

Statistical analysis was generally repeated-measures ANOVA with probability levels adjusted by the Greenhouse-Geisser correction. Time effects during the procedure were examined after dividing the breath holdings into three consecutive sets of four trials, and taking the mean of each measure for each set. Using means helped to reduce error variance in the physiological measures. The Tukey-Kramer statistic was used for post hoc comparisons of means of the three diagnostic groups, and the unpaired t test for means of panickers and nonpanickers. Only main group effects and interactions with group are reported except for the analyses of the Mood Questionnaire. Certain demographic and questionnaire items needed to be analyzed with nonparametric tests. Our comparisons of patient panickers and patient nonpanickers included patients from both the PD and GAD groups. These comparisons were made selectively when particularly relevant.

RESULTS

Panic Attacks

During the test as defined above, 32% (6 of 19) of PD patients, 18% (3 of 17) of GAD patients, and 0% (0 of 22) of normal controls were panickers (chi squared (2) = 7.84, p < .02). The difference between PD and GAD patients was not significant (chi squared (1) = 0.93). None of the three GAD patients with a history of PD were panickers. Of the patients reporting attacks, the mean (SD) number of attacks reported was 3.1 (2.0) with a range of 1 to 8. No subject asked for the procedure to be terminated early.
Respiration Questionnaire

Results for the Respiration Questionnaire are presented in Table 2. PD patients endorsed higher scores than normal controls on 8 of the 16 items and more than GAD patients on 3 items.

Breath Holding Performance

Breath holding time was calculated from the end of initial inspiration following a single signal tone to the beginning of expiration following a double signal tone (or earlier, for subjects unable to hold their breath that long). The mean (SD) for PD was 26.5 (3.0) seconds, for GAD, 27.0 (2.1) seconds, and for normal controls 27.6 (1.7) seconds. ANOVA with time (three consecutive sets of four trials) as a repeated measure showed no main Group effect but a Group by Time interaction (F(4,102) = 4.03, p < .02), because for Trials 1 to 4 the PD patients held their breaths about 2 second shorter than other groups, whereas at the two other time points the mean times of the three groups differed less than 1 second. Although follow-up ANOVAs did not show significant diagnostic group differences, we analyzed the first four trials separately. There was no main Group effect, but again there was a Group by Time interaction (F(6,153) = 3.40, p < .01). Follow-up ANOVAs showed a Group difference for Trial 1 (F(2,51) = 4.50, p < .02), but not for any of the other three trials. Post hoc Tukey-Kramer tests showed that on Trial 1, PD patients differed (p < .05) from GAD patients and normal controls, but GAD patients and normal controls did not differ from each other. Mean breath holding time on Trial 1 was 23.6 seconds for PD, 27.3 seconds for GAD, and 27.5 seconds for normal controls.

Mood Ratings

The diagnostic groups differed in mood ratings just before the procedure on the STAI-State Form (F(2,54) = 26.4, p < .0001). Post hoc Tukey-Kramer tests showed that PD and GAD patients both were higher (p < .05) than normal controls, but not different from each other. The mean for PD patients was 43.8, for GAD patients was 46.9, and for normal controls was 27.9. Figure 1 shows Mood Questionnaire scores for four time points: just before the procedure (before), average while breathing normally during the procedure (breathe), average while breath holding (hold), and just after the procedure (after). One set of ANOVAs compared pretest ratings, another pretest rating and posttest ratings (the first and last time points in the figure) and a third, breathe and hold data (the middle two time points in the figure). For before ratings Group effects were found only for “anxious” (F(2,55) = 17.7, p < .0001) and “excited” (F(2,55) = 3.40, p < .05). Post hoc tests showed that PD and GAD patients both rated themselves more anxious than normal controls but not different from each other. The statistics of the following two analyses appear in Table 3. For before/after (closely related to the previous analysis but capable of detecting changes over the test) both “anxious” and “excited” showed Group effects, but no before/after effects or interactions. For breathe/hold, “anxious” and “excited” showed the same Group effects as for before/after, but in addition a large breathe/hold effect due to increases from normal breathing to breath holding. The lack of interactions means breath holding did not provoke more mood change in PD or GAD patients than in normal controls.

Figure 1. Means and standard errors of Mood Questionnaire scores for four time points: just before the breath holding procedure (before), average during the procedure while breathing normally (breathe), average during the procedure while breath holding (hold), and just after the procedure ended (after). Breathe and hold were rated retrospectively at that time. Groups are panic disorder patients (PD), generalized anxiety disorder patients (GAD), and normal controls (Controls).

Table 3. ANOVAS for Mood Questionnaire Ratings: PD Patients, GAD Patients, and Normal Controls

Panicking and nonpanicking patients did not differ significantly on the STAI-State Form, but differed in the Mood Questionnaire in that panickers were higher than nonpanickers on before ratings of "anxious" (t(34) = 2.68, p < .02) and "excited" (t(34) = 2.94, p < .006) but did not differ on "alert" or "in control." An absence of Group by before/after interactions (for "anxious," F(1,38) = 0.51) means that panickers did not change more or less than nonpanickers. An absence of Group by breathe/hold interactions (for "anxious," F(1,38) = 2.51) means that panickers and nonpanickers did not rate their amount of change in mood during breath holding differently.

Symptoms During the Breath Holding Procedure

The diagnostic groups differed on Symptom Questionnaire B in total scores (F(2,54) = 6.46, p < .005), number of symptoms (F(2,54) = 6.88, p < .005), and average severity of symptoms (F(2,54) = 5.77, p < .005). Post hoc Tukey-Kramer tests showed that for these three scores PD patients were significantly (p < .05) higher than normal controls and that GAD patients had intermediate...
scores that were not significantly different from those of PD patients or normal controls. Individual symptoms that showed a group effect by the Kruskal-Wallis test with tie correction were "heart racing, pounding, skipping or chest pain or pressure," "dizziness, unsteadiness, or faintness," "unreality or detachment," "flushing," and "fear of going crazy," but not "shortness of breath." Panicking and nonpanicking patients also differed significantly on most items of Symptom Questionnaire B. Absolute means (SDs) for shortness of breath (on a five-point scale) were 2.78 (1.64) for panickers and 0.92 (1.16) for nonpanickers (p < .01 by the Mann-Whitney U test).

Physiological Measures Before the First Breath Holding

The 20-second period before the first breath holding is an indicator of baseline and anticipatory effects. None of the cardiovascular, electrodermal, or respiratory variables differed between the diagnostic groups at this time point. The F ratios for heart rate, systolic and diastolic blood pressure, skin conductance, and respiratory measures (including end-tidal pCO₂) were less than 1.0. A comparison of the panicking and nonpanicking patients also failed to show differences. For end-tidal pCO₂, t(21) equaled 0.03. The correlation between end-tidal pCO₂ values before the first breath holding and the breath holding times for Trial 1 was significantly positive for all groups combined (r = 0.33, p < .03), but not for individual groups. Perhaps intolerance to feelings of suffocation arising from stress led to both initially lower pCO₂ values and shorter breath holding times. These end-tidal pCO₂ values did not correlate with changes in anxiety between before and after or breathe and hold for the groups combined or any group.

Physiological Measures During Breath Holdings

Cardiovascular measures responded to breath holding in a complex manner, illustrated for heart rate in Figure 2. Inspiration and onset of a breath holding is accompanied by cardiac acceleration (measured in individual subjects as the "inspiratory maximum" heart rate) followed by deceleration (measured as the "postinspiratory minimum") and a return to an intermediate HR level ("final breath holding" heart rate measured as the mean in a 10-second interval ending at the end of breath holding). The maximum may represent gating of cardiac vagal tone whereas the minimum is a component of the "diving reflex" [35]. Expiration at the end of a breath holding is followed by a second acceleration ("post-breath holding maximum") due to the first inspiration after breath holding. This is followed by a "recovery" to intermediate levels. All of these measures were calculated as differences from 20-second baselines ending 10 seconds before the onset of each breath holding. Only inspiratory maximum differed significantly (F(4,102) = 5.31, p < .005) between the three groups. Post hoc tests confirmed the statistical significance of the unexpected result that GAD patients had larger accelerations than PD patients or normal controls, with a tendency for this effect to become stronger in later trials. The changes in inspiratory time and tidal volume could not explain the difference in heart rate acceleration, inasmuch as they showed no Group effects or interactions. Skin conductance response amplitude to onset of breath holding showed no significant effects. None of the other cardiovascular variables, nor skin conductance level, nor bodily movement differed between groups or showed group interactions at any of the measurement points during breath holding.

Figure 2. Averages of HR synchronized to onset and offset of breath holding over all 12 trials for the three diagnostic groups (see legend for Figure 1). Unlike in Figure 1, "before" and "after" refer to measurement points before and after individual trials, not before and after the entire procedure.

pCO₂ Before and Immediately After Each Breath Holding

pCO₂ for the first expiration after each breath holding and during the 20 seconds before each breath holding were averaged separately over three consecutive sets of four trials so that time effects could be evaluated. (Note that data from the time point before the first trial also appeared above in the analysis of measures before the first breath holding. Note also that in the following analyses, "before" and "after" refer to measurement points before and after individual breath holding trials, not before and after the entire procedure.) Figure 3 shows that of the three diagnostic groups, PD patients tended to have a lower pCO₂ before and after, whereas the differences between before and after within each diagnostic group were approximately equal. A three-way ANOVA showed significant effects for Group (F(2,36) = 3.65, p < .04), Before-After (F(1,36) = 227.0, p < .0001), and Time (F(2,72) = 23.1, p < .0001), but not for any of the interactions. Post hoc tests confirmed that pCO₂ in PD was lower than in normal controls (p < .05), but in GAD was not significantly different from pCO₂ in other groups. Times 1 and 2, and 1 and 3 differed significantly (p values < .0002), but not Times 2 and 3.

Figure 3. Means and standard errors of end-tidal pCO₂ before and at end of breath holdings for three consecutive sets of four trials for the three diagnostic groups (see legend for Figure 1). Unlike in Figure 1, "before breath holding" refers to measurements before
individual trials; correspondingly "at end of breath holding" refers to measurements after individual trials.

### Physiological Measures After Each Breath Holding

The 20-second post-breath holding periods were grouped into three consecutive sets of four trials so that interactions of Time and Group effects could be evaluated. Among the respiratory measures, which included minute volume, tidal volume, and inspiratory flow rate, only pCO$_2$ showed a main Group effect ($F(2,36) = 3.81, p < .05$). Post hoc tests showed again that pCO$_2$ in PD was lower than in normal controls ($p < .05$), but in GAD was not significantly different from those in other groups. None of the respiratory measures showed a Group by Time interaction. Among measures of cardiovascular activation (heart rate, systolic and diastolic blood pressure, T wave amplitude, and respiratory sinus arrhythmia), electrodermal activation, and movement there were no main Group effects or Group by Time interactions.

Exclusion of the Three GAD Patients With a Past History of PD. Because these patients may have had the psychophysiological characteristics of PD patients, additional analyses were made excluding them. On the Respiration Questionnaire, the same pattern of significance was present except for the item "Not getting enough air swimming." GAD patients had significantly higher scores than either PD patients or normal controls for both "How often" and "Ever severely frightened." The pattern of significance for the Mood Questionnaire was unaltered except that Group by Before/After for in control was no longer significant. The patterns of significance for heart rate change during breath holdings and for other physiological measures were not changed.

### DISCUSSION

As stated in the introduction, if Klein's hypothesis is true, we would expect the largest psychological and physiological anxiety responses to breath holding in the PD patients, including higher heart rate and skin conductance responses during breath holdings and signs of greater respiratory drive and autonomic activation between them. After the first trial, our strategy of having subjects hold their breaths for 30 seconds was successful in producing a constant stimulus situation for all groups, but few of the expected results were observed. Contrary to the hypothesis, increases in self-reported anxiety from normal breathing to breath holding were the same for all diagnostic groups. Instead, group differences consisted of higher anxiety levels before the procedure in PD and GAD patients than in normal controls, with reactions to breath holding appearing as equal phasic increases added to unequal baselines, a pattern that has been observed in PD patients and normal controls with diverse provocations (eg, [36]).

The hypothesized greater physiological anxiety reactions in PD patients during and after individual breath holdings did not appear either. Because measuring a broad spectrum of potential physiological anxiety responses (HR, T wave amplitude, respiratory sinus arrhythmia, systolic and diastolic blood pressure, skin conductance level, and a wide variety of respiratory measures) made our testing procedure quite sensitive, the lack of any group differences goes strongly against Klein's hypothesis. However, in the case of HR measurements during breath holding, the "diving reflex" [35] may have altered our results somewhat. A component of this reflex is heart rate deceleration, which, if more intense in panicking patients than in nonpanickers, could have decreased rather than increased their heart rates during breath holding. Fear in some studies seemed to produce greater heart rate deceleration during breath holding [38]. However, these decelerative effects should be short-lived and not affect our final breath holding interval measurement, which also showed no HR differences. Finally, contrary to the suffocation alarm hypothesis, inspiratory flow rate in intervals after individual breath holdings, a measure of central respiratory drive [31], did not differ between groups.

A curious, unexpected result was that an HR acceleration occurring just before breath holdings was greater for GAD patients than for other groups. This acceleration was probably the result of a combination of the physical effort expended for quick expiration and deep inspiration, and the transient decrease in intrathoracic pressure associated with inspiration. One might speculate that their worrying led GAD patients to prepare themselves more vigorously for breath holding than did other groups, but this was not supported by inspiratory time and tidal volume measurements. Because GAD and PD patients were well matched on state anxiety just before differences, between the patient groups can be more plausibly attributed to their type of anxiety than to its severity. However, it is possible that certain of our GAD patients had the specific biological characteristic of PD, if such a characteristic exists. Interview diagnosis is unlikely to distinguish perfectly between anxiety syndromes [39]. This may explain why GAD patients tended to fall between PD patients and normal controls on several measures and, in a few cases, reported panic attacks before or during the procedure.

The pattern of results for one important physiological measure, end-tidal pCO$_2$, is compatible with the suffocation false alarm hypothesis. The lower end-tidal pCO$_2$ in PD patients compared with normal controls (although the difference between PD and GAD patients does not reach significance), could have been caused by hyperventilation in response to greater air hunger produced by an overly sensitive suffocation alarm. Because of the lower initial pCO$_2$ levels in PD patients, breath holding increased pCO$_2$ only to normal, not to hypercapnic levels. This may have prevented psychological or physiological anxiety responses to breath...
holdings from being greater in PD patients than in the other groups. In other words, an overly sensitive suffocation alarm could initially have produced increased respiratory drive, but as pCO₂ decreased, this drive might have decreased to a normal level where phasic reactions to transient increases in pCO₂ were also normal. Apparently, greater hyperventilation in PD patients seems to have been initially stimulated by breath holding, inasmuch as pCO₂ did not differ between groups before the first trial. All groups showed small progressive decreases in pCO₂ during the procedure, perhaps as a physiological adjustment to repeated breath holding or perhaps as a nonspecific stress reaction. Oddly, none of these pCO₂ differences were reflected in significant differences in minute volume or inspiratory flow. Because pCO₂ levels can change with minor imbalances between CO₂ production and ventilation averaged over many minutes, we assume that the changes in ventilation were below the threshold of detection of our measurement methods.

However, interpreting the lower end-tidal pCO₂ in PD as a confirmation of the suffocation false alarm hypothesis is open to several objections. First, a more parsimonious interpretation of the lower pCO₂ levels is that they are evidence for a hyperventilation theory of PD (eg, [40]). Observations of end-tidal pCO₂ (ż) and respiration over longer time periods might help clarify whether hyperventilation is primary or secondary in modulating the anxiety level of PD patients. Second, although absolute pCO₂ levels after breath holding were lower in PD patients, possibly making the provocation weaker for them, relative changes are important for stimulating ventilation (see [41], p. 480). An average increase of 6 nm pCO₂ is known to be a strong respiratory stimulus, more than doubling minute ventilation in normals even if pO₂ is not allowed to decrease (eg, [42]). Thus, it is puzzling that the hypothesized hypersensitive suffocation alarm of PD patients failed to produce greater than normal psychological and physiological anxiety responses to individual breath holdings. Third, the hypocapnia that we observed may be a nonspecific reaction to stressors: pCO₂ levels in PD may be more a reflection of the patients' fear rather than its cause. Lower pCO₂ has been observed in normal subjects doing stressful desk work [43] and anxious non-PD patients engaged in fearful imagery [44]. Consistent with nonspecificity, it is observed following opposite respiratory stressors: breath holding and instructed hyperventilation [45]. On the other hand, one experiment suggests that hypocapnia is more characteristic of PD than GAD under mild challenge or baseline conditions even when self-reported anxiety is equal [46], which is similar to our results.

Finally, it is important to note that not all possible versions of the suffocation false alarm hypothesis were tested by our breath holding experiment. First, if the fault in the biological monitor is intermittent, which could explain the infrequent and unpredictable occurrence of panic attacks in many PD patients, its malfunctioning might not be captured in an experimental run lasting only 18 minutes. More prolonged, probably ambulatory, measurements would be required. Second, if provocations producing uncomfortable levels of dyspnea in normal subjects are required to trigger the alarm in PD patients, our breath holding procedure was inadequate, because the levels of shortness of breath reported by most of our subjects, at least the ones not reporting panic, were low. On the other hand, these low levels of shortness of breath can be taken as evidence against the existence of a pathologically sensitive suffocation alarm. For a biological theory, the critical question is the adequacy of a change in pCO₂ levels to stimulate a biological monitor. Purely psychological theories, such as those postulating catastrophic cognitive interpretations of interoceptions, are sufficient to account for group differences in response to provocations producing definite dyspneic discomfort in normal subjects. Our nonpanicking subjects may have failed to endorse the term shortness of breath very strongly because they did not consider it to apply to a voluntarily induced air hunger that was reliably alleviated by few deep breaths: Such air hunger is uncomfortable and likely to stimulate physiological monitors of suffocation, but is not necessarily frightening.

The Respiration Questionnaire provided evidence both for and against the suffocation alarm hypothesis. For the hypothesis, PD patients experience shortness of breath when anxious more frequently and are more afraid of it than GAD patients or normal controls. PD patients also have had more experiences of past suffocation than the other two groups. On several other questions, PD patients answered in accordance with the suffocation alarm hypothesis compared with normal controls, although differences between PD patients and GAD patients did not reach significance. Against the hypothesis, PD patients were currently not more afraid of suffocation or choking than the other groups, and thought they could hold their breath as long. One explanation of this pattern of results is that after their first panic attacks, PD patients had learned to fear sensations for which they accepted the term "shortness of breath," but that they considered the words "suffocation" and "choking" as more specifically applicable to nonpanic situations such as being in a hot, humid place or having a piece of food caught in the trachea. Of course, retrospective self-report data are liable to various distortions that give them less weight than data obtained from direct physiological testing, especially when a physiocentric theory is being examined.

This research was supported by the Department of Veterans Affairs and National Institute of Mental Health Grant MH56094.

We thank Cassandra L. Lehman, who recruited, diagnosed, and tested the subjects, and Kathrin Borner, who helped analyze the results.

REFERENCES

19. Sanderson WC, Rapee RM, Barlow DH: The influence of an illusion of control on panic attacks induced via inhalation of 5.5% carbon dioxide enriched air. Arch Gen Psychiatry 46:157-162, 1989 [Medline Link] [Context Link]
43. Schleifer L, Ley R: End-tidal pCO2 as an index of psychophysiological activity during VDT data-entry work and relaxation. Ergonomics 37:245-254, 1994 [Medline Link] [Context Link]

<table>
<thead>
<tr>
<th></th>
<th>Panic Disorder</th>
<th>Generalized Anxiety Disorder</th>
<th>Normal Controls</th>
<th>F Ratio or χ²</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. in sample</td>
<td>19</td>
<td>17</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Women (%)</td>
<td>74</td>
<td>65</td>
<td>68</td>
<td>χ²(2)=0.35</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>43.3 (9.8)</td>
<td>39.4 (10.8)</td>
<td>44.8 (16.4)</td>
<td>F(2,35)=0.88</td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
<td>6.9 (5.8)⁺</td>
<td>5.2 (2.2)⁺</td>
<td>1.8 (2.3)⁺</td>
<td>F(2,43)=7.46+++</td>
</tr>
<tr>
<td>Body Sensations Questionaire</td>
<td>2.6 (1.0)⁺</td>
<td>2.3 (0.8)⁺</td>
<td>1.7 (0.7)⁺</td>
<td>F(2,4)=5.87++</td>
</tr>
<tr>
<td>Agoraphobic Cognitions Questionaire</td>
<td>2.2 (0.6)⁺</td>
<td>1.7 (0.6)⁺</td>
<td>1.4 (0.4)⁺</td>
<td>F(2,4)=11.0+++</td>
</tr>
<tr>
<td>Mobility Inventory (unaccompanied)</td>
<td>2.6 (0.9)⁺</td>
<td>1.4 (0.5)⁺</td>
<td>1.4 (0.4)⁺</td>
<td>F(2,4)=17.4+++</td>
</tr>
<tr>
<td>State-Trait Anxiety Inventory-Trail</td>
<td>46.2 (9.0)⁺</td>
<td>44.6 (8.1)⁺</td>
<td>32.6 (8.4)⁺</td>
<td>F(2,4)=12.6+++</td>
</tr>
</tbody>
</table>

⁺⁻ Means with different superscripted letters differ at the .05 level on post hoc tests (Tukey-Kramer).
* p < .05; ** p < .01; *** p < .005.
<table>
<thead>
<tr>
<th>Shortness of breath when anxious</th>
<th>Panic Disorder</th>
<th>Generalized Anxiety Disorder</th>
<th>Normal Controls</th>
<th>F Ratio or $\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often (1–4)*</td>
<td>3.3 (0.9)*</td>
<td>2.2 (1.1)*</td>
<td>1.6 (0.8)*</td>
<td>$F = 16.9^{***}$</td>
</tr>
<tr>
<td>Ever severely frightened (% yes)?</td>
<td>78*</td>
<td>53*</td>
<td>5*</td>
<td>$\chi^2 = 22.9^{***}$</td>
</tr>
<tr>
<td>Fear now (1–4)*</td>
<td>2.9 (1.3)*</td>
<td>1.6 (0.9)*</td>
<td>1.3 (0.6)*</td>
<td>$F = 14.6^{***}$</td>
</tr>
<tr>
<td>Shortness of breath at other times</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often?</td>
<td>1.8 (0.8)</td>
<td>1.6 (0.7)</td>
<td>1.5 (0.6)</td>
<td>$F = 0.31$</td>
</tr>
<tr>
<td>Ever severely frightened?</td>
<td>4.4*</td>
<td>24*</td>
<td>9*</td>
<td>$\chi^2 = 6.69^{*}$</td>
</tr>
<tr>
<td>Fear now?</td>
<td>2.1 (1.0)*</td>
<td>1.5 (0.8)*</td>
<td>1.3 (0.5)*</td>
<td>$F = 4.69^{*}$</td>
</tr>
<tr>
<td>Past choking</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often?</td>
<td>1.9 (1.3)*</td>
<td>1.5 (0.7)*</td>
<td>1.0 (0.2)*</td>
<td>$F = 5.67^{*}$</td>
</tr>
<tr>
<td>Ever severely frightened?</td>
<td>4.4</td>
<td>35</td>
<td>23</td>
<td>$\chi^2 = 2.14$</td>
</tr>
<tr>
<td>Fear now?</td>
<td>2.0 (1.3)</td>
<td>1.6 (0.9)</td>
<td>1.4 (0.7)</td>
<td>$F = 2.06$</td>
</tr>
<tr>
<td>Past suffocating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often?</td>
<td>2.2 (1.3)*</td>
<td>1.4 (0.8)*</td>
<td>1.0 (0.2)*</td>
<td>$F = 9.92^{***}$</td>
</tr>
<tr>
<td>Ever severely frightened?</td>
<td>50*</td>
<td>24*</td>
<td>9*</td>
<td>$\chi^2 = 8.64^{*}$</td>
</tr>
<tr>
<td>Fear now?</td>
<td>2.1 (1.3)</td>
<td>1.6 (0.9)</td>
<td>1.4 (0.9)</td>
<td>$F = 2.05$</td>
</tr>
<tr>
<td>Not getting enough air swimming</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often?</td>
<td>1.5 (0.9)*</td>
<td>2.1 (1.2)*</td>
<td>1.4 (0.7)*</td>
<td>$F = 3.22^{*}$</td>
</tr>
<tr>
<td>Ever severely frightened?</td>
<td>13</td>
<td>41</td>
<td>13</td>
<td>$\chi^2 = 5.41$</td>
</tr>
<tr>
<td>Fear now?</td>
<td>1.4 (1.0)</td>
<td>1.8 (1.3)</td>
<td>1.4 (0.8)</td>
<td>$F = 0.87$</td>
</tr>
<tr>
<td>Ability to hold breath (1–5)</td>
<td>1.7 (0.7)</td>
<td>1.7 (0.6)</td>
<td>2.0 (0.5)</td>
<td>$F = 2.57$</td>
</tr>
</tbody>
</table>

*How often* was rated as never = 1, seldom = 2, sometimes = 3, and often = 4; *Ever severely frightened* as "yes" or "no"; *Fear now* as not at all = 1; slightly = 2; moderately = 3, and extremely = 4; and *Ability to hold breath* as shorter = 1, average = 2, or longer = 3.

*bc* Means with different superscripted letters differ at the .05 level on post hoc tests (Tukey-Kramer or $\chi^2$).

*Degrees of freedom for $F$ are 2 and 56 except for the next to last row where they are 2 and 53. Degrees of freedom for $\chi^2$ is 2.

* $p < .05$; ** $p < .01$; *** $p < .005$. 

---

**ANXIOUS**

**EXCITED**

**ALERT**

**IN CONTROL**
<table>
<thead>
<tr>
<th>Before (B)/After</th>
<th>Group (G)</th>
<th>B/A</th>
<th>B/H</th>
<th>G by B/A</th>
<th>G by B/H</th>
<th>Post Hoc Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A)</td>
<td>Hazardous</td>
<td>F(2,55) = 15.0***</td>
<td>F(2,55) = 0.45</td>
<td>F(2,55) = 2.78</td>
<td>PD^a &gt; GAD^a &gt; C^a</td>
<td></td>
</tr>
<tr>
<td>Excited</td>
<td>F(2,55) = 5.38**</td>
<td>F(2,55) = 0.12</td>
<td>F(2,55) = 2.37</td>
<td>GAD^b &gt; PD^b &gt; C^b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alert</td>
<td>F(2,55) = 1.14</td>
<td>F(2,55) = 22.0***</td>
<td>F(2,55) = 0.56</td>
<td>PD: B^c &gt; A^c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In control</td>
<td>F(2,55) = 2.09</td>
<td>F(2,55) = 2.77</td>
<td>F(2,55) = 3.23*</td>
<td>GAD: B^c &gt; A^c</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Breath (B)/Hold (H)**

| Anxious | F(2,55) = 14.0*** | F(1,55) = 24.3*** | F(2,55) = 0.19 | GAD^a > PD^a > C^a |
| Excited | F(2,55) = 7.07** | F(1,55) = 14.8*** | F(2,55) = 0.13 | GAD^b > PD^b > C^b |
| Alert   | F(2,55) = 2.03 | F(1,55) = 1.24 | F(2,55) = 0.78 | PD^b < GAD^ab < C^c |
| In control | F(2,55) = 1.88* | F(1,55) = 1.34 | F(2,55) = 0.74 | PD^c < GAD^ab < C^c |

^a,b Means with different superscripted letters differ at the .05 level on post hoc tests.

*p < .05; ** p < .01; *** p < .005.

---

**AVERAGE HEART RATE RESPONSE**

![Average Heart Rate Response Graph](image)

**PD**  
(N=11)  
**GAD**  
(N=14)  
**CONTROLS**  
(N=14)

Before Breath Holding

At End of Breath Holding

![Trial Graph](image)