

Supplementary Materials

Section A: Scripts used with breathlessness assessments

Section B: Confirmatory factor analysis and Exploratory Principal Components Analysis (PCA) for focal periods (past two weeks (daily life) and last minute of walk test.

Table S1: Summary Confirmatory Factor Analysis (1 and 2 factors) for Dyspnoea -12 (D-12) and Multidimensional Dyspnoea Profile (MDP) for two focal periods; average over past two weeks (daily life) and breathlessness during last minute of walk test.

Table S2: Standardised factor Loadings for the two-factor confirmatory factor analysis (CFA) model

Table S3: Summary of results for Exploratory Principal Components Analysis (PCA) for Dyspnoea -12 and Multidimensional Dyspnoea Profile (MDP) for two focal periods; daily life (average over past two weeks) and during last minute of walk test

Section C: Comparisons between scalable instruments.

Figure S1 Bland Altman plots for single developer recommended scores and subdomain scores for breathlessness in daily life and end exercise.

Table S4: Summary of participant characteristics (Mean \pm standard deviation or freq. (%)): Total sample for clinical trial (n=101), sample with data for both breathlessness focal periods (n=84)

Table S5: Comparisons between severity (GOLD grade) with **breathlessness in daily life** and differences between daily life and breathlessness during last minute of six minute walk test for scalable instruments by item (VAS's, Dyspnoea-12 (all items) and Multidimensional Dyspnoea Profile (all items)

Table S6: Comparisons between severity (GOLD grade) with **breathlessness at end of walk test** and differences between daily life and breathlessness during last minute of six minute walk test for scalable instruments by item (VAS's, Dyspnoea-12 (all items) and Multidimensional Dyspnoea Profile (all items)

Associations and tests for difference in associations between scalable instruments.

Table S7. Recalled sensation of breathlessness in **daily life**: Pearson's Correlation Coefficient r (95% confidence intervals) and differences in associations (z-test based on Fisher's z transformation; Bonferroni adjusted p for significance ≤ 0.002)

Table S8. Recalled sensation of breathlessness **during last minute of walk test**: Pearson's Correlation Coefficient r (95% confidence intervals) and differences in associations (z-test based on Fisher's z transformation; Bonferroni adjusted p for significance ≤ 0.002)

Figure S2: Significant differences between visual analogue scale for breathlessness intensity (VAS-I) associations with Dyspnoea -12 (D-12) and Multidimensional Dyspnoea Profile (MDP) scores ($p < 0.002$)

Section D: Consistency of sensory quality choice across breathlessness assessments

Table S9 Classification of descriptors /items for breathlessness assessments.

Table S10. Breathlessness in daily life: individual consistency for descriptors across breathlessness assessments

Table S11. Breathlessness at end exercise: individual consistency for descriptors across breathlessness assessments

Section A: Scripts used with breathlessness assessments.

All five breathlessness assessments were compiled into a protocol book with written scripts to be used by the research personnel. Each assessment was presented on a separate page (or subsections of instruments on separate pages). Breathless assessments were undertaken prior to functional exercise tests (sensation recalled over the past two weeks) and immediately on completion of the walk test (sensation during the last minute of the walk test). Assessments of intensity (Visual analogue scales) and volunteered descriptors of breathlessness were always completed first for both focal periods in order to limit “leading or biasing” participants after viewing breathlessness descriptors presented in other instruments), with the remaining three assessments (endorsed descriptors, Multidimensional Dyspnoea Profile (MDP) and Dyspnoea-12 (D-12) undertaken by predetermined randomised order). While developers of the MDP recommend that where other instruments are used in conjunction with the MDP, these should not be interspersed with MDP items in order to maintain the integrity of the MDP, the introductory script of the MDP (defining and drawing analogies to intensity an unpleasantness of music) was used to preface all breathlessness assessments in order to clarify the conceptual differences between intensity and unpleasantness.

Introductory script:

The next series of questions is all about the sensation you feel when you are breathless. For these questions you will need to think about your sensation of breathlessness OVER THE PAST TWO WEEKS.

Some of these questions ask you to focus on the intensity of the sensation, some ask you to focus on how unpleasant the sensation is and some focus on the quality or how you would describe the sensation.

The distinction between the aspects of **intensity** and **unpleasantness** breathing sensations might be made clearer if you think of listening to a sound. As the volume of the sound increases, I can ask you how loud it sounds or how unpleasant it is to hear it. For example, music that you hate can be unpleasant even when the volume is low, and will become more unpleasant as the volume increases; music that you like will not be unpleasant, even when the volume increases.

Does this make sense?

There are no right or wrong answers. We want to know what you tell us about your own breathing.

Visual analogue scales for intensity and unpleasantness

First, I’d like you to think about how INTENSE your sensation of breathlessness has been **on average over the past two weeks**.

I’m going to ask you to draw a line in this scale. This end of the scale means that you did not feel any degree of breathlessness over the past two weeks (left side) while this end of the scale is the worse breathlessness you have ever felt (right side). Could you make put a line to indicate how INTENSE your sensation of breathlessness has been over the past two weeks?

Not breathless at all	_____	The worst breathlessness I’ve ever felt
-----------------------------	-------	---

Now I’d like you to think about how UNPLEASANT your sensation of breathlessness has been **on average over the past two weeks**.

I’m going to ask you to draw a line in this scale. This end of the scale means that you did not feel any degree of unpleasantness (left side) while this end of the scale is the most unpleasant experience of breathlessness you have ever felt (right side). Could you make put a line to indicate how UNPLEASANT your sensation of breathlessness has been on average over the past two weeks?

Not
unpleasant
at all

The most
unpleasant
experience of
breathlessness
I've ever felt

Volunteered Descriptors

I'm going to get you to think about your breathing over the past two weeks.

Which words would you use to describe your breathing when it is uncomfortable?

Prompt:

Could you try and explain the sensation you feel when your breathing is uncomfortable

"I'm trying to understand what it feels like for you when your breathing is uncomfortable".

Do these words or phrases describe the worst breathlessness experience you've had or your sensation on an average day?

Can you describe your worst experience of breathlessness? When was this?

Note: Data for the two sub questions has not been reported in this manuscript. The majority of participants indicated that the volunteered words described breathless sensations on average rather than a specific or worst experience over the past two weeks.

Endorsed Descriptors

This is a list of statements which other people have used to describe their breathing when it is uncomfortable. Please read through this list. On this list of statements, please **choose up to three (3)** which best describe the feeling you have had **over the past two weeks** when your breathing is uncomfortable?

Participant was invited to view standardised list of 15 descriptor phrases (Mahler et al 1996). Research personnel confirmed and circled each of the selected descriptor.

Do these words or phrases describe the worst breathlessness experience you've had or your sensation on an average day?

Note: Data for the sub questions has not been reported in this manuscript. The majority of participants indicated that the descriptor statements selected described breathless sensations on average rather than a specific or worst experience over the past two weeks.

Dyspnoea 12 Questionnaire

This questionnaire is designed to help us **learn more about how your breathing is troubling you.**

Please read each item and then tick the box that best matches your breathing **these days (during the past two weeks)**. If you did not experience an item over the past two weeks, tick the "none" box. Please respond to all items.

Participant was invited to view D-12 questionnaire (York et al 2010). Research personnel checked that each item had been completed.

Multidimensional Dyspnoea Profile

(Note: During planning and initial commencement of the clinical trial a penultimate version of the MDP (Banzett et al 2008) was provided by the developers which was upgraded to the ultimate version during the initial stages of the clinical trial)

On this page we ask you to tell us how unpleasant your breathing feels. On a later page we will ask you about the intensity or strength of your breathing sensations.

Affective 1 (A1) scale

Use this scale to rate the **unpleasantness or discomfort** of your breathing sensations, how bad your breathing felt.

Please focus on the period during the past two weeks.

A1 scale here

Sensory quality (SQ) choice (new page)

Below are phrases or terms arranged in groups of similar meaning.

Step 1: Check each group that describes how your breathing felt during the past two weeks

Step 2: Please also mark one group that most accurately describes how your breathing felt during the past two weeks

Sensory quality table displayed here

SQ scales (new page)

Use these scales to rate the intensity of the breathing sensations you felt (like the loudness of sound regardless of whether the sensation is pleasant or unpleasant; for example a sensation could be intense without being unpleasant.)

Please focus on the period during the past two weeks.

Rating of sensory qualities table here

Affective 2 (A2) scales (new page)

When your breathing doesn't feel normal, you may experience emotions or feelings. Using the scales below, please tell us about how your breathing sensations made you feel - rate zero for any emotion you did not feel.

Please focus on feelings during the past two weeks.

Rating of emotional responses table here

On completion of assessments of breathlessness (daily life).

Ensure that the participant is instructed PRIOR to the walk test that a second language of breathlessness interview will occur immediately at the end of the walk test.

We are going to repeat all the questions I asked you about your sensation of breathlessness at the end of this walking test. But this time, I'm going to ask you to think about your sensation of breathlessness **DURING THE LAST MINUTE OF THE WALKING TEST**. So when walking test is finished, I'll ask you to sit in this chair and then I'll ask you the series of questions about your sensation of breathlessness.

The focal period for all assessments completed at cessation of the six minute walk test was “...during the **LAST MINUTE** of the walk test.”

Section B Confirmatory and exploratory and factor analysis for focal periods (past two weeks (daily life) and last minute of walk test.

Confirmatory and exploratory factor analysis for the two focal periods reported within this paper;

Breathlessness on average over past two weeks (daily life) (sample included within clinical trial (n=101);

Breathlessness during last minute of the walk test (participants with paired data for sensations of breathlessness (n = 84).

Suitability of the data for factor analysis was first examined. The correlation matrices contained numerous values greater than 0.3 and Bartlett's test of sphericity (Bartlett 1954) reached statistical significance in all cases, supporting the factorability of the correlation matrix. Furthermore, Kaiser-Meyer-Olkin values ranged from 0.79 through 0.92, exceeding the recommended value of 0.60 (Kaiser 1974).

A two-factor Confirmatory Factor analysis (CFA) model was first applied to the data set for each focal period and evaluated using model-fit criteria recommended by Bentler (2007), namely the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the standardised root mean square residual (SRMR). The sample size was relatively small for CFA, and the following results should be treated with caution. While model fit statistics should be treated as guidelines and not as golden rules (Marsh et al., 2004), it is recommended (Hu and Bentler 1999) that RMSEA be no greater than 0.06, CFI be greater than 0.9, and SRMR less than 0.08. The resulting fit indices show mixed results (**Table 1**) suggesting none of the models fit ideally, however all factor loadings being ≥ 0.47 and significant ($p < .05$) (**Table 2**). By way of comparison, a simple one-factor CFA model was also run, with results indicating the two factor model was a better fit (**Table 1**).

Due to the mixed results of the CFA models, exploratory Principal Component Analysis with varimax rotations was undertaken to test whether our modification ("past two weeks" and "Last minute of the walk test") of the focal periods used during development studies of the D-12 ("These days") and MDP ("right now", activity or time based) had affected the underlying factor structures. Data for both focal periods resulted in the same 2-factor structures indicated by Meek et al (2012) and Yorke et al (2010) with similar levels of variance explanation and Cronbach Alpha scores (**Table 3**).

Table S1: Summary Confirmatory Factor Analysis (1 and 2 factors) for Dyspnoea -12 (D-12) and Multidimensional Dyspnoea Profile (MDP) for two focal periods; average over past two weeks (daily life) and breathlessness during last minute of walk test.

Multidimensional Dyspnoea Profile (MDP)	Daily life		End of walk test		Comments: (based on 2 factor solution)
	1 Factor	2 Factors	1 Factor	2 Factors	
Standardised root mean square residual (SRMR < 0.08 good fit)	0.08	0.06	0.18	0.10	Daily life < 0.08 (good fit) End walk test > 0.08 (ill fit)
Root mean square error of approximation (RMSEA \leq 0.06 good fit)	0.14	0.10	0.23	0.12	Both > 0.06 (ill fit)
Bentler comparative fit index (CFI) (> 0.9 good fit)	0.83	0.92	0.71	0.92	Both > 0.9 (good fit)
Dyspnoea-12					
Standardised root mean square residual (SRMR < 0.08 good fit)	0.07	0.06	0.10	0.08	Daily life < 0.08 (good fit) End of walk test > 0.08 (ill fit)
Root mean square error of approximation (RMSEA \leq 0.06 good fit)	0.12	0.08	0.22	0.18	Both > 0.06 (ill fit)
Bentler comparative fit index CFI(> 0.9 good fit)	0.89	0.95	0.75	0.83	Daily life > 0.9 (good fit) End of walk test < 0.9 (ill fit)

Table S2: Standardised factor Loadings for the two-factor confirmatory factor analysis (CFA) model

	Factor and Items	(Rotated) Factor Loadings	
		Daily life ("last two weeks)	End of walk test
Dyspnoea-12	Factor 1: "Physical"		
	My breathe does not go in all the way	0.65	0.65
	My breathing requires more work	0.67	0.67
	I feel short of breath	0.85	0.85
	I have difficulty catching my breath	0.87	0.97
	I cannot get enough air	0.89	0.89
	My breathing is uncomfortable	0.86	0.86
	My breathing is exhausting	0.80	0.80
	Factor 2: "Affective"		
	My breathing makes me feel depressed	0.81	0.81
	My breathing makes me feel miserable	0.88	0.88
	My breathing is distressing	0.84	0.84
	My breathing makes me feel agitated	0.81	0.81
	My breathing is irritating	0.71	0.71
Multidimensional Dyspnoea Profile	Factor 1: "Immediate Perception"		
	Unpleasantness	0.66	0.72
	My breathing requires muscle work or effort	0.71	0.72
	I am not getting enough air or I am smothering or I feel hunger for air	0.78	0.81
	My chest or lungs feel tight or constricted	0.64	0.71
	My breathing requires mental effort or concentration	0.68	0.73
	I am breathing a lot	0.71	0.79
	Factor 1: "Emotional Response"		
	Depressed	0.78	0.98
	Anxious	0.84	0.57
	Frustrated	0.72	0.54
	Angry	0.56	0.80
	Afraid	0.78	0.99

Note: CFA used maximum likelihood estimation. All factor loadings were significant ($p < .05$).

Table S3 : Summary of results for Exploratory Principal Components Analysis (PCA) for Dyspnoea -12 and Multidimensional Dyspnoea Profile (MDP) for two focal periods; daily life (average over past two weeks) and during last minute of walk test.

Focus period (“Focus on your breathing felt/feels.....”)					
	“on average over last 2 weeks” (Daily life)	“during the last minute of walk test”	D-12 (Yorke et al 2010) (“These days”)	MDP (Meek et al 2012) (“right now”	MDP (Parshall et al 2012) (“When decided to visit the ED”)
Multidimensional Dyspnoea Profile (MDP)					
Exploratory PCA => 2 factors:					
Variance explained by factors	61%	68%		T1 =66% T2 =74% T3 =72%	T0a=63% T0b=71% T0c=64%
Standardised Cronbach's alpha:					
Factor 1 (Immediate perception domain)	0.85	0.88		T1=0.93 T2=0.96 T3=0.96 T4 =0.94 T5 =0.97	T0a=0.89 T0b=0.94 T0c=0.90
Factor 2 (Emotional response domain)	0.86	0.88		T1=0.84 T2=0.86 T3=0.86 T4 =0.92 T5 = 0.94	T0a=0.81 T0b=0.85 T0c=0.81
Dyspnoea-12 (D-12)					
Exploratory PCA => 2 factors:					
Variance explained by factors	63%	72%	64%		
Standardised Cronbach's alpha:					
Factor 1 (Physical domain)	0.92	0.92	NR		
Factor 2 (Affective domain)	0.90	0.90	NR		

Yorke et al (2010) –participants recruited through outpatients clinics completed assessments during a clinic visit NR –Not reported in published paper

Meek et al (2012) –T1 =Time one -initial emergency department presentation for breathlessness; T2 =60 minutes after T1; T3 =at least one hour after T2, T4 -4-6 weeks after emergency department visit, T5 -4 to 6 weeks after Emergency department visits following completion of spirometry, oximetry and questionnaires (recall of T1).

Parshall et al (2012) Time 0a: initial emergency department presentation for breathlessness; Time 0b: = 60 minutes after T0a;Time 0c: 4 to 6 weeks after Emergency department visits (recall of T0a).

References

1. Bartlett M. S. A note on the multiplying factors for various chi square approximations. Journal of the Royal Statistical Society: Series B, 1954, 16, 296-298.
2. Kaiser H. An index of factorial simplicity. Psychometrika, 1974, 39, 31-36.
3. Bentler P.M. On tests and indices for evaluating structural models. Personality and Individual Differences, 2007, 42, 825-829.
4. Hu L. and Bentler P.M. Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modelling, 6, 1-55.
5. Kline R.B. Principles and Practice of Structural Equation Modeling (2nd Edition ed.) 2005. New York: The Guilford Press.
6. Marsh, H. W., Hau, K. T., Wen, Z. In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's 1999 findings. Structural Equation Modeling, 2004, 11, 320-341.
7. Hayduk L., Cummings G.G., Boadu K., Pazderka-Robinson H., and Boulianne S. Testing! Testing! One, Two Three - Testing the theory in structural equation models! Personality and Individual Differences, 2007, 42 (2), 841-50.
8. Yorke J, Moosavi SH, Shulldham C, Jones PW. Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12. Thorax 2010; 65(1): 21-26.
9. Meek PM, Banzett R, Parshall MB, Gracely RH, Schwartzstein RM, Lansing R. Reliability and Validity of the Multidimensional Dyspnoea Profile. Chest 2012;141(6):1546-1563
10. Parshall MB, Meek PM, Sklar D, Alcock J, Bittner P. Test-retest reliability of multidimensional dyspnoea profile recall ratings in the emergency department: a prospective, longitudinal study. BMC Emergency Medicine 2012; 12(6): doi: 10.1186/1471-227X-12-6

Section C: Comparisons between scalable instruments.

Figure S1 Bland Altman plots for single developer recommended scores and subdomain scores for breathlessness in daily life and end exercise (scores presented as percentage of maximum score for each instrument).

Each individual point indicates an individual participant's score (presented as percentage of maximum score for each developer recommended or subdomain score). **x-axis** (Average score (%)) displays the mean of the two scores under consideration (D-12 and MDP recommended or subdomain scores). **y-axis** (Difference (%)) displays the difference between the D-12 and MDP scores (D12 minus MDP) with the mean (estimated bias) and ± 1.96 standard deviation (95% limits of agreement) lines indicated for each assessment. "Outliers" are indicated by individual points that fall above or below the ± 1.96 standard deviation (95% limits of agreement) lines.

If the majority of points are scattered above and below the zero line, this suggests that there is no consistent bias in favour of one instrument compared to the other. If the majority of points fall above (or below) the zero line, this suggests that one instrument consistently scores higher (or lower) than the other.

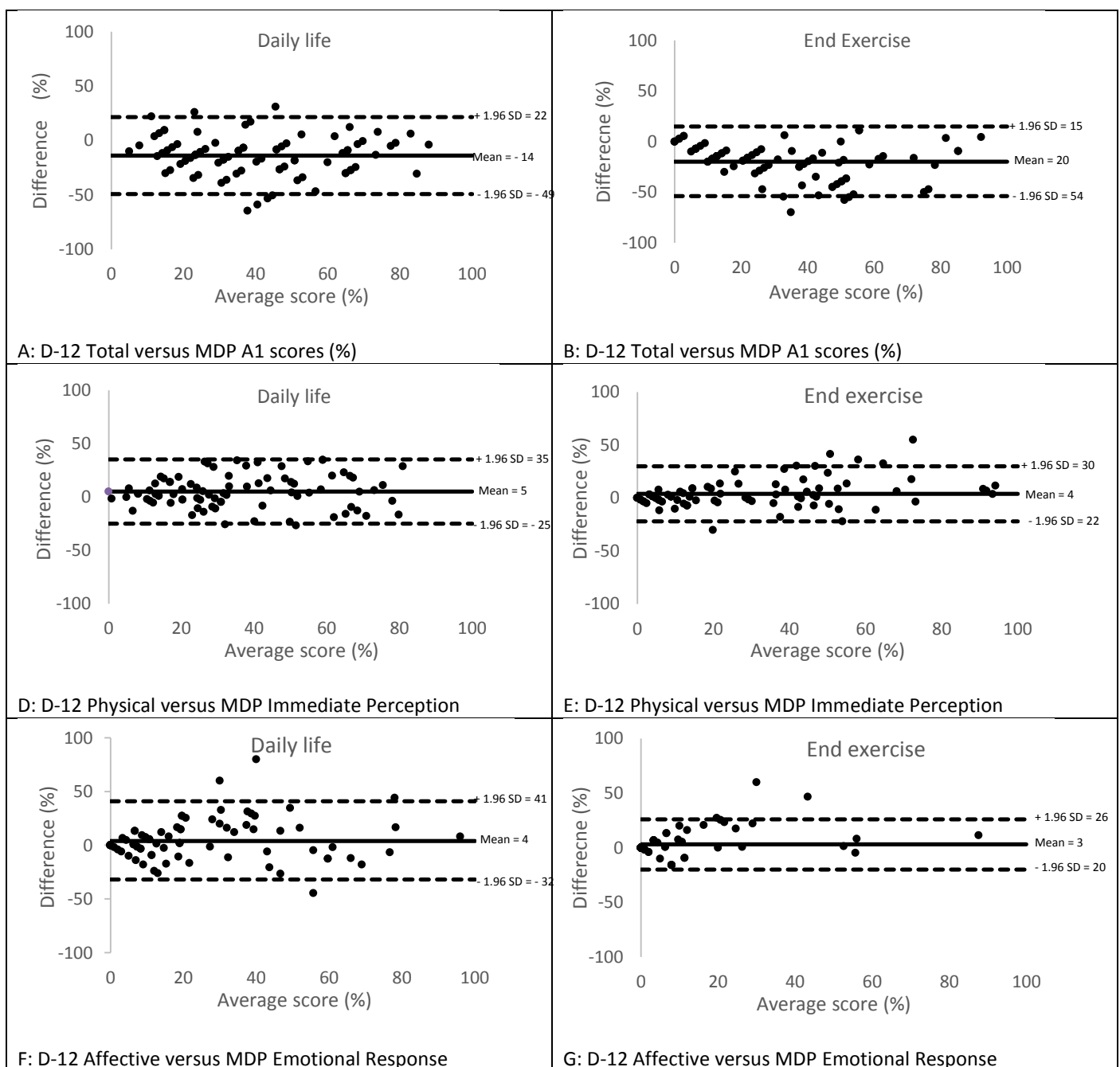


Table S4: Summary of participant characteristics (Mean \pm standard deviation or freq. (%)): Study sample with data for both breathlessness focal periods (n=84) and by Global initiative for Obstructive Lung Disease (GOLD) Grade for severity of airflow limitation. Note: Bonferroni adjusted p value for statistical significance $p \leq 0.001$

	Study sample	GOLD Grade (for study sample n=84)		
		2	3	4
Sample size	84	40	31	13
Age	69.7 \pm 8.9	70.8 \pm 8.7	71.4 \pm 7.8	62.5 \pm 8.8
Height	166 \pm 9	164.1 \pm 8.2	165.8 \pm 8.6	171.9 \pm 9.1
Weight	75.1 \pm 18.6	77.6 \pm 20.8	75.4 \pm 16.8	66.8 \pm 14.1
Male Sex	47 (56)	20 (50)	18 (58)	9 (69)
English spoken at home	81 (96)	39 (98)	29 (94)	13 (100)
Mini Mental State Examination	29.2 \pm 1.7	29.1 \pm 2	29.1 \pm 1.5	29.5 \pm 0.7
mMRC n (%)	1.9 \pm 1.2	1.8 \pm 1.3	1.8 \pm 1.1	2.8 \pm 1.1
0	5 (6)	3 (7)	2 (6)	0 (0)
1	37 (44)	21 (53)	14 (45)	2 (15)
2	13 (15)	4 (10)	6 (19)	3 (23)
3	17 (20)	6 (15)	7 (23)	4 (31)
4	12 (14)	6 (15)	2 (6)	4 (31)
HADs-Anxiety	7.0 \pm 4.4	7.0 \pm 4.6	7.2 \pm 4.5	6.8 \pm 4
HADs-Depression	6.3 \pm 4.2	6.1 \pm 4.4	5.5 \pm 3.6	9.2 \pm 4
CRQ-Dyspnoea	4.7 \pm 1.4	5.0 \pm 1.3	4.5 \pm 1.5	4.1 \pm 1.7
CRQ-Fatigue	3.9 \pm 1.3	3.9 \pm 1.3	4 \pm 1.2	4.1 \pm 1.2
CRQ-Emotion	4.6 \pm 1.2	4.6 \pm 1.1	4.7 \pm 1.2	4.6 \pm 1.5
CRQ-Mastery	4.8 \pm 1.4	5.0 \pm 1.4	4.7 \pm 1.3	4.3 \pm 1.6
pH	7.4 \pm 0	7.4 \pm 0	7.4 \pm 0	7.4 \pm 0
Arterial carbon dioxide kPa	5.4 \pm 0.8	*5.0 \pm 0.4	*5.4 \pm 0.5	*6.3 \pm 1.1
Arterial oxygen kPa	10.0 \pm 1.4	10.3 \pm 1.3	9.9 \pm 1.5	9.2 \pm 0.8
SaO ₂ (percent)	94.9 \pm 2.1	95.4 \pm 1.9	94.7 \pm 2.4	94.1 \pm 1.6
PRE-rest	0.9 \pm 1.1	0.9 \pm 1.1	0.8 \pm 1.1	1.1 \pm 1.1
PRE -post 6MWT	3.5 \pm 1.8	3.2 \pm 1.7	3.5 \pm 1.8	4.3 \pm 1.6
Maximum distance 6MWT (m)	385 \pm 135	404 \pm 127	385 \pm 121	325 \pm 177
FEV ₁ percent predicted	47.9 \pm 16.6	*61.7 \pm 9.2	*41.0 \pm 5.7	*21.6 \pm 6.0
FEV ₁ /FVC	42.6 \pm 14.1	*51.6 \pm 11.2	*38.6 \pm 9.6	*24.4 \pm 8.0
PEF % pred	62.4 \pm 22	*79.3 \pm 15.5	*53.0 \pm 11.8	*32.9 \pm 9.5
IC %pred (p)	97.9 \pm 24.3	*121 \pm 27	*94.3 \pm 20.2	*76.6 \pm 19
VC % pred (p)	92.3 \pm 19.5	97.1 \pm 18.3	90.4 \pm 19.8	82.5 \pm 20.1
ITGV % pred (p)	142.7 \pm 41	*121 \pm 27	*147 \pm 36	*200 \pm 34
RV % pred (p)	167 \pm 56.4	*139 \pm 36	*170 \pm 43	*251 \pm 63
TLC % pred (p)	117 \pm 22	*110 \pm 20	*119 \pm 23	*138 \pm 11
RV/TLC % pred (p)	134 \pm 26	*121 \pm 17	*138 \pm 18	*168 \pm 36
DLCO % pred	51.5 \pm 18.2	*58.5 \pm 18.9	*48.1 \pm 13.8	*35 \pm 13
Key				
*	$P \leq 0.001$			
mMRC	Modified Medical Research Council Scale	(p)	Plethysmography (n=71 valid measures)	
HAD	Hospital Anxiety and Depression scale	IC % pred	Inspiratory capacity percent predicted	
CRQ	Chronic Respiratory questionnaire	VC % pred	Vital capacity percent predicted	
6MWT	Six minute walk test	ITGV % pred	Intrathoracic gas volume percent predicted	
PRE	Perceived rate of exertion	RV % pred	Residual volume percent predicted	
FEV ₁ % pred	Forced expiratory volume in one second percent predicted	TLC % pred	Total lung capacity percent predicted	
FEV ₁ /FVC	Ratio between FEV ₁ and Forced Vital Capacity	RV/TLC%pred	Residual volume percent predicted total lung capacity	
PEF % pred	Peak expiratory flow percent predicted	DLCO %pred	Diffusion capacity for carbon monoxide percent predicted	

Table S5: Comparisons between severity (GOLD grade) with **breathlessness in daily life** and differences between daily life and breathlessness during last minute of six minute walk test for scalable instruments by item (VAS's, Dyspnoea-12 (all items) and Multidimensional Dyspnoea Profile (all items))

Instrument		GOLD Grade (n = 84) (daily life)				Focal periods (n = 84)		Paired t-test
		Grade 2 (n = 40)	Grade 3 (n = 31)	Grade 4 (n = 13)	ANOVA	Daily life	Last minute of walk test	
		Mean \pm SD	Mean \pm SD	Mean \pm SD	p=	Mean \pm SD	Mean \pm SD	P=
Breathlessness intensity (VAS)		44.8 \pm 21.5	51.0 \pm 20.5	61.2 \pm 26.7	0.06	49.6 \pm 22.5	52.4 \pm 26.9	0.35
Breathlessness Unpleasantness (VAS)		40.0 \pm 24.9	56.2 \pm 26.6	53.6 \pm 29.0	0.03*	48.1 \pm 27.0	41.1 \pm 30.2	0.07
Dyspnoea-12								
My breathe does not go in all the way		1.1 \pm 0.9	0.9 \pm 1	0.9 \pm 0.8	0.78	1.0 \pm 0.9	0.8 \pm 1.0	0.13
My breathing requires more work		1.3 \pm 0.8	1.4 \pm 1	1.5 \pm 1	0.59	1.3 \pm 0.9	1.1 \pm 0.9	0.08
I feel short of breath		1.4 \pm 0.8	1.6 \pm 0.9	2 \pm 0.6	0.08	1.6 \pm 0.8	1.3 \pm 1.0	0.02
I have difficulty catching my breath		1.3 \pm 1	1.2 \pm 1.1	1.2 \pm 1	0.99	1.2 \pm 1.0	0.8 \pm 1.1	0.001*
I cannot get enough air		0.8 \pm 1	1.2 \pm 0.9	1 \pm 0.8	0.33	1.0 \pm 0.9	0.9 \pm 1.1	0.52
My breathing is uncomfortable		1 \pm 1	1.1 \pm 0.9	1.4 \pm 0.8	0.38	1.1 \pm 0.9	0.9 \pm 1.0	0.52
My breathing is exhausting		0.9 \pm 1.1	1.1 \pm 1	0.8 \pm 1.1	0.63	1.0 \pm 1.1	0.8 \pm 1.1	0.18
Physical sub score		7.7 \pm 4.5	8.5 \pm 5.5	8.9 \pm 4.3	0.65	8.1 \pm 4.9	6.7 \pm 5.9	0.02
My breathing makes me feel depressed		0.5 \pm 0.9	0.7 \pm 0.9	0.6 \pm 0.9	0.70	0.6 \pm 0.9	0.2 \pm 0.5	<0.001*
My breathing makes me feel miserable		0.6 \pm 1	0.7 \pm 0.8	0.6 \pm 0.8	0.83	0.7 \pm 0.9	0.2 \pm 0.6	<0.001*
My breathing is distressing		0.7 \pm 0.9	1 \pm 1	0.8 \pm 0.8	0.26	0.8 \pm 1.0	0.4 \pm 0.7	<0.001*
My breathing makes me feel agitated		0.7 \pm 1	0.9 \pm 1	0.7 \pm 0.9	0.74	0.8 \pm 1.0	0.2 \pm 0.6	<0.001*
My breathing is irritating		1 \pm 1	1.1 \pm 0.9	1.3 \pm 0.6	0.52	1.1 \pm 0.9	0.4 \pm 0.8	<0.001*
Affective sub score		3.5 \pm 4	4.4 \pm 3.9	4.1 \pm 3.1	0.57	3.9 \pm 3.8	1.4 \pm 2.8	<0.001*
Total score (range 0 to 36)		11.1 \pm 7.8	12.9 \pm 8.7	13 \pm 7.1	0.59	12.0 \pm 8.0	8.1 \pm 8.1	<0.001*
Multidimensional Dyspnoea Profile	Forced Choice N = (%)###	Grade 2 (n = 40)	Grade 3 (n = 31)	Grade 4 (n = 13)	ANOVA	Daily life	Last minute of walk test	Paired t-test
Unpleasantness (A1)	NA	4.5 \pm 2.4	4.7 \pm 2.3	5.4 \pm 2.2	0.50	4.7 \pm 2.3	4.2 \pm 2.8	0.06
My breathing requires muscle work or effort	10 (14%)	2.5 \pm 2.7	2.8 \pm 2.8	5 \pm 3.6	0.03*	3.0 \pm 3.0	2.7 \pm 3.1	0.40
I am not getting enough air or I am smothering or I feel hunger for air	19 (26%)	3.2 \pm 2.8	3 \pm 3	4.1 \pm 3.4	0.51	3.2 \pm 2.9	2.0 \pm 2.9	<0.001*
My chest or lungs feel tight or constricted	27 (37%)	3.4 \pm 2.9	3.4 \pm 3.3	4 \pm 3	0.80	3.5 \pm 3.0	2.0 \pm 2.8	<0.001*
My breathing requires mental effort or concentration	6 (8%)	2.2 \pm 2.9	2.6 \pm 2.7	3.3 \pm 3.1	0.49	2.5 \pm 2.9	2.5 \pm 3.1	0.92
I am breathing a lot	11 (15%)	3.2 \pm 3.2	2.9 \pm 2.8	4.5 \pm 3	0.29	3.3 \pm 3.0	3.3 \pm 3.2	0.97
Immediate Perception sub score		18.9 \pm 13.8	19.5 \pm 12.1	26.2 \pm 13.6	0.21	20.2 \pm 13.3	16.7 \pm 14.1	0.02
Depressed		1.5 \pm 2.9	1 \pm 2.1	3.4 \pm 3	0.02	1.6 \pm 2.7	0.4 \pm 1.6	<0.001*
Anxious		2.6 \pm 3.2	2.1 \pm 2.6	3 \pm 3.4	0.61	2.5 \pm 3.0	1.0 \pm 2.3	<0.001*
Frustrated		2.9 \pm 2.9	2.9 \pm 3	5.1 \pm 3.2	0.06	3.2 \pm 3.1	1.2 \pm 2.5	<0.001*
Angry		1.4 \pm 2.7	1.9 \pm 3.1	2.8 \pm 2.7	0.28	1.8 \pm 2.9	0.4 \pm 1.3	<0.001*
Afraid		1.8 \pm 3.2	1.5 \pm 2.6	2 \pm 2.3	0.85	1.7 \pm 2.8	0.2 \pm 0.9	<0.001*
Emotional Response sub score		10.1 \pm 12.2	9.4 \pm 10.2	16.3 \pm 12.6	0.18	10.8 \pm 11.7	3.2 \pm 7.0	<0.001*

Note: Bonferroni adjusted p value for statistical significance is p < 0.002

* p \leq 0.002

MDP Sensory Quality forced choice –percentages calculated out of the number of participants that selected a forced choice (n= 73)

Multidimensional Dyspnoea Profile Sensory Quality forced choice question responses: 73 participants (from 84) provided a forced choice (11 did not select); 61 participant's forced choice was in agreement with SQ highest intensity rating (84%), 11 were in disagreement (11/61 =15%) and six participants selected a forced choice but did not rate intensity in any SQ category. Of the 61 participants in agreement with SQ forced choice and highest rating, 29 rated two or more SQ categories equally as high (48%).

Table S6: Comparisons between severity (GOLD grade) with **breathlessness at end of walk test** and differences between daily life and breathlessness during last minute of six minute walk test for scalable instruments by item (VAS's, Dyspnoea-12 (all items) and Multidimensional Dyspnoea Profile (all items).

Descriptor categories#			GOLD Grade (End Exercise) (n = 84)			χ^2
			Grade 2 (n = 40)	Grade 3 (n = 31)	Grade 4 (n = 13)	
			N (%)	N (%)	N (%)	P =
Volunteered	Air hunger/ Suffocating		16 (40)	9 (29)	5 (38)	0.62
	Work /Effort		13 (33)	13 (42)	5 (38)	0.71
	Tight/ Constricted		2 (5)	1 (3)	2(15)	0.28
	Frightening /Awful /Worried		3 (8)	1 (3)	1 (8)	0.72
	Uncomfortable /Annoying		6 (15)	3 (10)	2 (15)	0.78
	Helpless/ Regret/Depressed		1 (3)	2 (6)	1 (8)	0.64
Endorsed	Air hunger		14 (35)	14 (45)	4 (31)	0.57
	Work		15 (38)	13 (42)	8 (62)	0.31
	Tight		10 (25)	4 (13)	3 (23)	0.44
	Unnamed		33 (83)	24 (77)	9 (69)	0.59
Scalable instruments		Forced choice N= (%)##	Mean \pm SD	Mean \pm SD	Mean \pm SD	ANOVA p=
VAS	Breathlessness intensity (VAS)		49.9 \pm 27.2	51.7 \pm 26.9	61.9 \pm 26.0	0.38
	Breathlessness Unpleasantness (VAS)		39.1 \pm 31.3	36.6 \pm 27.8	57.9 \pm 28.6	0.09
Dyspnoea-12	My breathe does not go in all the way		0.7 \pm 0.9	0.9 \pm 1.0	0.9 \pm 1.1	0.47
	My breathing requires more work		0.9 \pm 0.9	1.2 \pm 0.9	1.6 \pm 1.0	0.06
	I feel short of breath		1.1 \pm 1.0	1.5 \pm 0.7	1.8 \pm 1.1	0.04
	I have difficulty catching my breath		0.6 \pm 1.0	0.9 \pm 1.0	1.2 \pm 1.3	0.16
	I cannot get enough air		0.7 \pm 1.1	1.0 \pm 1.0	1.4 \pm 1.3	0.09
	My breathing is uncomfortable		0.8 \pm 1.0	1.0 \pm 0.9	1.4 \pm 1.3	0.13
	My breathing is exhausting		0.7 \pm 1.0	0.8 \pm 1.0	1.2 \pm 1.2	0.23
	Physical sub score		5.3 \pm 5.5	7.2 \pm 5.2	9.6 \pm 7.5	0.06
	My breathing makes me feel depressed		0.1 \pm 0.3	0.1 \pm 0.4	0.5 \pm 1.1	0.01
	My breathing makes me feel miserable		0.2 \pm 0.5	0.1 \pm 0.4	0.5 \pm 1.1	0.12
	My breathing is distressing		0.2 \pm 0.4	0.4 \pm 0.8	0.8 \pm 1.1	0.03
	My breathing makes me feel agitated		0.2 \pm 0.4	0.2 \pm 0.6	0.6 \pm 1.0	0.05
	My breathing is irritating		0.4 \pm 0.8	0.3 \pm 0.7	0.8 \pm 1.0	0.11
	Affective sub score		1.1 \pm 2.1	1.1 \pm 2.3	3.3 \pm 4.7	0.03
Total score (range 0 to 36)			6.4 \pm 7.2	8.3 \pm 6.8	12.9 \pm 11.9	0.04
Multidimensional Dyspnoea Profile	Unpleasantness (A1)	N/A	3.9 \pm 2.9	4.2 \pm 2.5	5.2 \pm 2.9	0.38
	My breathing requires muscle work or effort	13 (19%)	2.1 \pm 2.9	2.9 \pm 2.7	4.2 \pm 4.0	0.11
	I am not getting enough air or I am smothering or I feel hunger for air	12 (18%)	1.9 \pm 2.9	1.5 \pm 2.2	3.4 \pm 3.9	0.14
	My chest or lungs feel tight or constricted	9 (13%)	2.0 \pm 2.8	1.6 \pm 2.2	3.1 \pm 3.8	0.27
	My breathing requires mental effort or concentration	14 (21%)	1.9 \pm 2.8	2.1 \pm 2.5	5.5 \pm 3.4	<0.001*
	I am breathing a lot	19 (28%)	2.9 \pm 3.2	3.1 \pm 2.8	4.7 \pm 3.8	0.20
	Immediate Perception sub score		14.7 \pm 14.1	15.3 \pm 10.1	25.9 \pm 19.2	0.03
	Depressed	NA	0.3 \pm 0.8	0.2 \pm 1.3	1.3 \pm 3.3	0.09
	Anxious		1.2 \pm 2.5	0.6 \pm 1.8	1.3 \pm 2.8	0.56
	Frustrated		0.9 \pm 2.0	0.6 \pm 1.6	3.5 \pm 4.3	<0.001*
	Angry		0.3 \pm 1.1	0.3 \pm 1.1	0.8 \pm 2.2	0.40
	Afraid		0.2 \pm 0.7	0.1 \pm 0.7	0.7 \pm 1.7	0.16
	Emotional Response sub score		2.8 \pm 5.2	1.9 \pm 5.5	7.7 \pm 12.4	0.04

Note: Bonferroni adjusted p value for statistical significance is p <0.002

* p \leq 0.002

Descriptors categories reported in Williams et al (2010) with data representing the number of participants volunteering/endorsing descriptor within each category. Note, in cases where the majority of cells were small, a Fisher's Exact Test was also carried out. In all cases, a similar result (ie significant/non-significant) was obtained

MDP Sensory quality forced choice –percentages calculated out of the number that selected a forced choice (n=67)

Multidimensional Dyspnoea Profile Sensory Quality Forced Choice question responses: 67 participants (from 84) provided a forced choice (17 did not select). Of the 67, 57 participant's forced choice was in agreement with SQ highest intensity rating (85%), nine were in disagreement (9/67 =13%) and one participant selected a forced choice but did not rate intensity in any SQ category. Of the 57 participant's in agreement with SQ forced choice and highest rating, 25 rated two or more SQ categories equally as high (44%).

Associations and tests for difference in associations between scalable instruments.

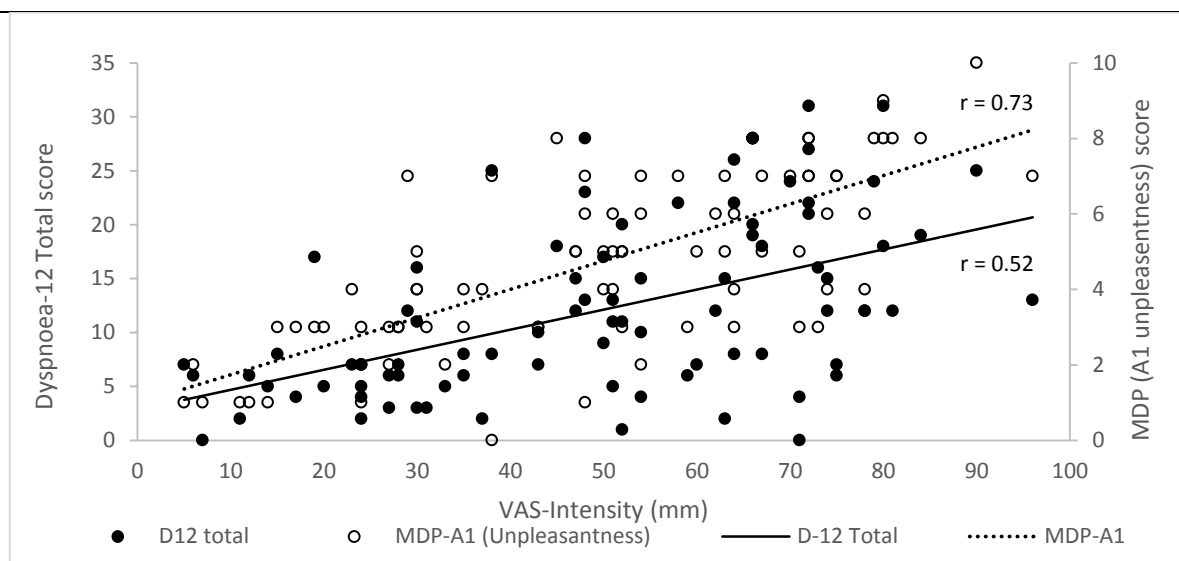
Table S7. Sensation of breathlessness in **daily life**: Pearson's Correlation Coefficient r (95% confidence intervals) and differences in associations (z-test based on Fisher's z transformation; Bonferroni adjusted p for significance ≤ 0.002).

	VAS-I	VAS-U	D-12 Total	A1 MDP	D-12 vs A1 MDP	D-12 Subdomains		D-12 Physical vs MDP IP	MDP subdomains		D-12 Affective vs MDP ER
	r I95,u95	r I95,u95	r I95,u95	r I95,u95	Z score P=	Physical r I95,u95	Affective r I95,u95	Z score P=	IP r I95,u95	ER r I95,u95	Z score P=
Association or difference											
D-12 Physical				0.67 0.53,0.77					0.77 0.67,0.84		
D-12 Affective				0.59 0.43,0.71						0.72 0.59,0.80	
D-12 Total				0.69 0.55,0.78							
VAS - unpleasantness	0.75 0.64,0.83		0.58 0.41,0.70	0.70 0.57,0.80	-2.00 0.05	0.53 0.35,0.67	0.54 0.36,0.67	-0.51 0.61	0.56 0.39,0.69	0.52 0.34,0.66	0.30 0.77
VAS - intensity	-	-	0.52 0.34,0.66	0.73 0.60,0.81	-3.16 0.002*	0.54 0.36,0.67	0.41 0.22,0.58	-0.72 0.47	0.58 0.41,0.70	0.37 0.17,0.54	0.57 0.57
mMRC	0.53 0.36,0.67	0.47 0.28,0.62	0.40 0.20,0.56	0.57 0.40,0.70	-2.29 0.02	0.48 0.30,0.63	0.22 0.00,0.41	0.46 0.64	0.45 0.26,0.61	0.36 0.15,0.53	-1.75 0.08
HADS-Anxiety	0.18 -0.04,0.38	0.26 0.04,0.45	0.46 0.26,0.61	0.26 0.04,0.45	2.44 0.01	0.33 0.12,0.51	0.54 0.36,0.67	-0.14 0.89	0.34 0.13,0.52	0.52 0.34,0.66	0.18 0.86
HADS-Depression	0.38 0.18,0.55	0.35 0.14,0.52	0.34 0.13,0.51	0.35 0.14,0.53	-0.19 0.85	0.22 0.01,0.42	0.42 0.22,0.58	-2.05 0.04	0.37 0.16,0.54	0.57 0.41,0.70	-2.15 0.03*
CRQ-Dyspnoea	-0.45 -0.60,-0.25	-0.51 -0.65,-0.33	-0.42 -0.58,-0.22	-0.54 -0.67,-0.36	1.54 0.12	-0.42 -0.58,-0.22	-0.36 -0.53,-0.14	-0.52 0.60	-0.38 -0.55,-0.18	-0.44 -0.60,-0.24	1.10 0.27
CRQ-Fatigue	-0.36 -0.54,-0.15	-0.41 -0.57,-0.21	-0.37 -0.54,-0.16	-0.47 -0.62,-0.27	1.22 0.22	-0.31 -0.49,-0.09	-0.39 -0.56,-0.19	1.94 0.05	-0.44 -0.60,-0.24	-0.50 -0.65,-0.32	1.47 0.14
CRQ-Emotion	-0.36 -0.53,-0.15	-0.47 -0.62,-0.27	-0.46 -0.62,-0.27	-0.47 -0.62,-0.28	0.13 0.90	-0.35 -0.53,-0.14	-0.53 -0.67,-0.35	1.18 0.24	-0.43 -0.59,-0.23	-0.62 -0.74,-0.46	1.29 0.20
CRQ-Mastery	-0.52 -0.66,-0.34	-0.61 -0.73,-0.44	-0.60 -0.72,-0.43	-0.63 -0.74,-0.47	0.43 0.67	-0.51 -0.65,-0.32	-0.62 -0.73,-0.46	1.22 0.22	-0.58 -0.71,-0.42	-0.63 -0.75,-0.48	0.23 0.82
PRE (Borg) - rest	0.35 0.14,0.52	0.19 -0.02,0.39	0.26 0.05,0.45	0.42 0.22,0.58	-1.91 0.06	0.30 0.09,0.48	0.17 -0.05,0.37	-0.05 0.96	0.30 0.09,0.48	0.18 -0.04,0.38	-0.13 0.90
PRE (Borg) – post 6MWT	0.38 0.18,0.55	0.33 0.12,0.51	0.33 0.12,0.51	0.48 0.29,0.63	-1.86 0.06	0.37 0.16,0.54	0.23 0.01,0.42	-0.43 0.67	0.40 0.20,0.56	0.37 0.17,0.54	-0.43 0.67
6MWT (max metres)	-0.29 -0.48,-0.08	-0.25 -0.44,-0.03	-0.13 -0.34,0.08	-0.31 -0.49,-0.10	2.09 0.04	-0.16 -0.36,0.06	-0.08 -0.29,0.14	0.73 0.47	-0.22 -0.41,0.00	-0.15 -0.36,0.06	0.95 0.34
FEV1 %pred	-0.27 -0.46,-0.06	-0.26 -0.45,-0.05	-0.10 -0.30,0.12	-0.12 -0.32,0.10	0.25 0.80	-0.06 -0.27,0.16	-0.12 -0.33,0.10	0.98 0.33	-0.15 -0.35,0.07	-0.16 -0.36,0.06	0.47 0.63
FEV1/FVC	-0.17 -0.37,0.04	-0.11 -0.32,0.11	-0.13 -0.33,0.09	-0.11 -0.32,0.10	0.25 0.80	-0.09 -0.30,0.13	-0.15 -0.36,0.06	0.98 0.33	-0.21 -0.41,-0.00	-0.13 -0.33,0.09	0.47 0.63
PEF %pred	-0.18 -0.38,0.04	-0.17 -0.37,0.05	-0.09 -0.30,0.13	-0.11 -0.31,0.11	-0.19 0.85	-0.03 -0.24,0.19	-0.15 -0.35,0.07	1.62 0.11	-0.12 -0.33,0.10	-0.15 -0.35,0.07	-0.34 0.73
IC %pred	-0.38 -0.56,-0.15	-0.32 -0.51,-0.09	-0.12 -0.35,0.12	-0.27 -0.47,-0.04	0.18 0.86	-0.13 -0.35,0.11	-0.09 -0.32,0.15	1.18 0.24	-0.11 -0.34,0.12	-0.10 -0.32,0.14	-0.01 0.99
VC % pred	-0.25 -0.46,-0.01	-0.28 -0.48,-0.05	-0.20 -0.42,0.03	-0.20 -0.42,0.04	1.73 0.08	-0.20 -0.41,0.04	-0.17 -0.38,0.07	-0.15 0.88	-0.13 -0.35,0.11	-0.18 -0.39,0.06	0.07 0.94
ITGV % pred	0.10 -0.14,0.33	0.06 -0.17,0.29	-0.02 -0.25,0.22	0.07 -0.17,0.30	-0.001 0.99	-0.04 -0.28,0.19	0.02 -0.22,0.25	-0.97 0.33	0.16 -0.08,0.38	0.10 -0.14,0.33	0.14 0.89
RV % pred	0.12 -0.12,0.34	0.11 -0.12,0.34	0.10 -0.14,0.32	0.10 -0.14,0.32	-1.01 0.31	0.07 -0.17,0.30	0.12 -0.12,0.34	-2.69 0.01	0.23 -0.00,0.44	0.22 -0.01,0.43	-1.05 0.29
TLC %_pred	-0.04 -0.27,0.19	-0.06 -0.29,0.18	-0.05 -0.28,0.19	-0.04 -0.27,0.20	0.01 0.99	-0.08 -0.31,0.16	-0.00 -0.23,0.23	-2.21 0.03*	0.11 -0.12,0.34	0.06 -0.17,0.29	-1.28 0.20
RV/ TLC% pred	0.17 -0.07,0.39	0.19 -0.05,0.40	0.20 -0.04,0.41	0.16 -0.08,0.38	-1.03 0.30	0.17 -0.07,0.39	0.20 -0.04,0.41	-2.56 0.01	0.24 0.01,0.45	0.26 0.03,0.47	0.42
DLCO% pred	-0.18 -0.39,0.06	-0.20 -0.41,0.03	-0.09 -0.31,0.14	-0.16 -0.37,0.08	0.50 0.62	-0.08 -0.30,0.15	-0.09 -0.31,0.14	-1.03 0.30	-0.22 -0.43,-0.01	-0.04 -0.27,0.19	0.55
*	$p \leq 0.002$				6MWT		Six minute walk test				
I95	Lower 95 % confidence limit				FEV ₁ % pred		Forced expiratory volume in one second percent predicted				
u95	Upper 95 % confidence limit				FEV1/FVC		Forced expiratory volume in one second/ Forced Vital Capacity ratio				
VAS-I	VAS intensity				PEF % pred		Peak expiratory flow percent predicted				
VAS-U	VAS unpleasantness				IC % pred		Inspiratory capacity percent predicted (plethysmography)				
D-12	Dyspnoea 12				VC % pred		Vital capacity percent predicted (plethysmography)				
MDP	Multidimensional Dyspnoea Profile				ITGV % pred		Intrathoracic gas volume percent predicted (plethysmography)				
MDP (IP)	Immediate Perception				RV % pred		Residual volume percent predicted				
MDP (ER)	Emotional Response				TLC % pred		Total lung capacity percent predicted				
HAD	Hospital Anxiety and Depression scale				RV/TLC%pred		Residual volume percent predicted total lung capacity				
CRQ	Chronic Respiratory questionnaire				DLCO %pred		Diffusion capacity for carbon monoxide percent predicted				
mMRC	Modified Medical Research Council Scale						Significant linear association ($r > 0.2$)				

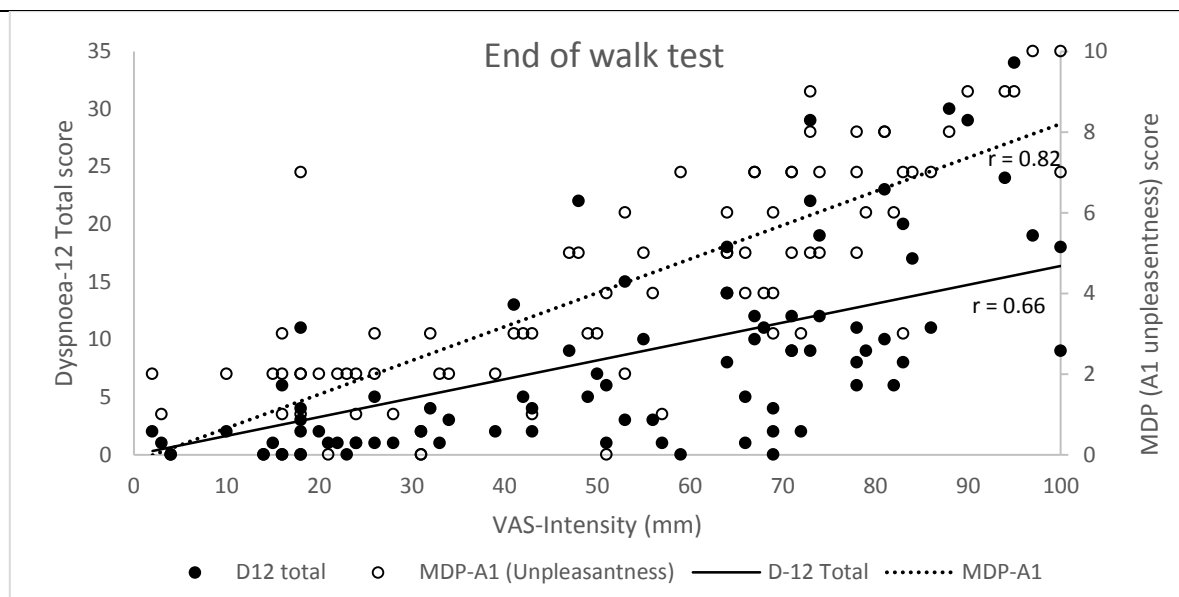
Table S8. Recalled sensation of breathlessness **during last minute of walk test:** Pearson's Correlation Coefficient r (95% confidence intervals) and differences in associations (z-test based on Fisher's z transformation).

	VAS-I	VAS-U	D-12 Total	A1 MDP	D-12 versus A1 MDP	D-12		D-12 Physical vs MDP IP	MDP		D-12 Affective vs MDP ER
						Physical	Affective		IP	ER	
Association or difference			r I95,u95	r I95,u95	Z score P=	r I95,u95	r I95,u95	Z score P=	r I95,u95	r I95,u95	Z score P=
D-12 Physical				0.80 0.71,0.87					0.88 0.82,0.92		
D-12 Affective				0.57 0.40,0.70						0.78 0.67,0.85	
D-12 Total				0.77 0.67,0.85							
VAS - unpleasantness			0.66 0.51,0.76	0.79 0.69,0.86	-2.76 0.006	0.71 0.58,0.80	0.42 0.22,0.58	0.58 0.56	0.69 0.55,0.78	0.54 0.36,0.67	-1.86 0.06
VAS -intensity		0.84 -0.77,0.90	0.66 0.52,0.77	0.82 0.73,0.88	-3.35 0.001*	0.71 0.58,0.80	0.44 0.25,0.60	0.86 0.40	0.68 0.54,0.78	0.46 0.27,0.61	-0.20 0.84
mMRC (continuous)	0.38 0.18,0.55	0.30 0.09,0.48	0.27 0.06,0.46	0.38 0.18,0.55	-1.64 0.10	0.29 0.08,0.47	0.17 -0.04,0.37	-0.68 0.50	0.32 0.12,0.50	0.11 -0.11,0.32	0.88 0.38
HADS-Anxiety	0.15 -0.07,0.35	0.18 -0.04,0.38	0.30 0.08,0.48	0.26 0.04,0.45	0.57 0.57	0.26 0.05,0.45	0.31 0.10,0.50	-1.67 0.10	0.35 0.14,0.52	0.23 0.02,0.43	1.09 0.28
HADS- Depression	0.18 -0.04,0.38	0.17 -0.05,0.38	0.33 0.11,0.51	0.27 0.06,0.46	0.72 0.47	0.28 0.07,0.47	0.36 0.15,0.53	-1.37 0.17	0.35 0.14,0.53	0.29 0.07,0.47	1.00 0.32
CRQ-Dyspnoea	-0.48 -0.63,-0.29	-0.45 -0.60,-0.25	-0.51 -0.66,-0.33	-0.59 -0.72,-0.43	1.29 0.20	-0.50 -0.65,-0.31	-0.44 -0.60,-0.24	1.18 0.24	-0.55 -0.69,-0.38	-0.28 -0.46,-0.06	-2.35 0.02
CRQ-Fatigue	-0.31 -0.49,-0.09	-0.34 -0.52,-0.13	-0.39 -0.56,-0.18	-0.46 -0.61,-0.26	1.02 0.31	-0.37 -0.54,-0.16	-0.35 -0.53,-0.14	1.50 0.13	-0.44 -0.60,-0.24	-0.33 -0.51,-0.12	-0.26 0.80
CRQ-Emotion	-0.27 -0.46,-0.05	-0.33 -0.51,-0.11	-0.35 -0.53,-0.14	-0.39 -0.56,-0.19	0.61 0.54	-0.31 -0.49,-0.09	-0.38 -0.55,-0.17	2.17 0.03	-0.42 -0.58,-0.21	-0.39 -0.56,-0.19	0.19 0.85
CRQ-Mastery	-0.25 -0.44,-0.03	-0.25 -0.44,-0.03	-0.30 -0.49,-0.09	-0.37 -0.54,-0.16	0.93 0.35	-0.28 -0.47,-0.06	-0.29 -0.47,-0.07	1.40 0.16	-0.35 -0.53,-0.14	-0.29 -0.48,-0.08	0.09 0.93
PRE (Borg) - rest	0.47 0.28,0.62	0.51 0.33,0.64	0.43 0.23,0.59	0.48 0.29,0.63	-0.83 0.40	0.42 0.23,0.58	0.35 0.15,0.53	0.23 0.82	0.41 0.21,0.57	0.35 0.15,0.53	0.01 0.99
PRE (Borg) – post 6MWT	0.79 0.69,0.86	0.69 0.55,0.79	0.69 0.55,0.79	0.78 0.67,0.85	-1.84 0.07	0.70 0.56,0.79	0.54 0.37,0.68	-0.73 0.46	0.72 0.60,0.81	0.59 0.42,0.71	-0.70 0.48
6MWT (max)	-0.30 -0.48,-0.09	-0.23 -0.42,-0.01	-0.21 -0.41,-0.00	-0.33 -0.51,-0.12	1.62 0.10	-0.21 -0.40,0.01	-0.18 -0.38,0.04	0.07 0.94	-0.21 -0.41,0.00	-0.07 -0.28,0.15	-1.46 0.14
FEV1 %pred	-0.15 -0.35,0.07	-0.16 -0.36,0.05	-0.28 -0.47,-0.07	-0.12 -0.33,0.10	-2.22 0.03	-0.28 -0.47,-0.07	-0.23 -0.42,-0.02	-0.22 -0.41,0.00	-0.20 -0.40,0.01	-0.12 -0.33,0.10	-1.52 0.13
FEV1/FVC	-0.07 -0.28,0.15	-0.12 -0.33,0.09	-0.23 -0.42,-0.01	-0.09 -0.30,0.13	-2.22 0.03	-0.24 -0.43,-0.03	-0.15 -0.35,0.07	-1.55 0.12	-0.14 -0.35,0.07	-0.02 -0.23,0.19	-1.52 0.13
PEF %_pred	-0.07 -0.28,0.14	-0.12 -0.33,0.10	-0.21 -0.40,0.01	-0.08 -0.29,0.14	-1.87 0.06	-0.21 -0.41,0.00	-0.15 -0.35,0.07	-1.85 0.06	-0.15 -0.35,0.07	-0.05 -0.26,0.16	-1.73 0.08
IC % pred	-0.10 -0.32,0.14	-0.04 -0.28,0.19	-0.13 -0.35,0.11	-0.09 -0.32,0.14	-1.69 0.09	-0.12 -0.35,0.12	-0.11 -0.34,0.13	-1.29 0.20	-0.15 -0.37,0.09	-0.02 -0.25,0.22	-1.31 0.19
VC % pred	-0.08 -0.31,0.16	-0.02 -0.26,0.21	-0.18 -0.39,0.06	-0.09 -0.32,0.15	-0.44 0.66	-0.16 -0.38,0.08	-0.18 -0.40,0.05	0.45 0.65	-0.18 -0.40,0.06	-0.11 -0.34,0.13	-1.27 0.20
ITGV % pred	0.01 -0.23,0.24	0.11 -0.13,0.33	0.28 0.05,0.48	0.08 -0.16,0.31	-1.12 0.26	0.27 0.03,0.47	0.25 0.01,0.46	0.47 0.64	0.16 -0.08,0.38	0.05 -0.19,0.28	-1.00 0.32
RV % pred	0.04 -0.19,0.27	0.15 -0.09,0.37	0.37 0.15,0.56	0.13 -0.10,0.36	2.76 0.006	0.35 0.12,0.54	0.36 0.13,0.54	2.12 0.03	0.26 0.03,0.46	0.14 -0.10,0.36	2.73 0.006
TLC % pred	-0.02 -0.25,0.21	0.09 -0.15,0.31	0.22 -0.02,0.43	0.05 -0.19,0.28	3.32 0.001*	0.22 -0.02,0.43	0.19 -0.05,0.40	1.74 0.08	0.10 -0.14,0.32	0.04 -0.20,0.27	3.02 0.003
RV /TLC % pred	0.10 -0.14,0.33	0.16 -0.08,0.38	0.39 0.17,0.57	0.15 -0.09,0.37	2.33 0.02	0.36 0.13,0.54	0.38 0.16,0.56	2.24 0.03	0.29 0.06,0.49	0.17 -0.06,0.39	2.02 0.04
DLCO %pred	-0.09 -0.31,0.14	0.002 -0.23,0.23	-0.10 -0.32,0.14	-0.04 -0.27,0.19	3.39 0.001*	-0.10 -0.32,0.13	-0.07 -0.29,0.17	1.30 0.19	-0.09 -0.31,0.15	0.05 -0.18,0.28	2.95 0.003
*	Bonferroni adjusted $p \leq 0.002$				6MWT		Six minute walk test				
I95	Lower 95 % confidence limit				FEV ₁ % pred		Forced expiratory volume in one second percent predicted				
u95	Upper 95 % confidence limit				FEV1/FVC		Forced expiratory volume in one second/ Forced Vital Capacity ratio				
VAS-I	VAS intensity				PEF % pred		Peak expiratory flow percent predicted				
VAS-U	VAS unpleasantness				IC % pred		Inspiratory capacity percent predicted (plethysmography)				
D-12	Dyspnoea 12				VC % pred		Vital capacity percent predicted (plethysmography)				
MDP	Multidimensional Dyspnoea Profile				ITGV % pred		Intrathoracic gas volume percent predicted (plethysmography)				
MDP (IP)	Immediate Perception				RV % pred		Residual volume percent predicted				
MDP (ER)	Emotional Response				TLC % pred		Total lung capacity percent predicted				
HAD	Hospital Anxiety and Depression scale				RV/TLC%pred		Residual volume percent predicted total lung capacity				
CRQ	Chronic Respiratory questionnaire				DLCO %pred		Diffusion capacity for carbon monoxide percent predicted				
mMRC	Modified Medical Research Council Scale						Significant association ($r > 0.2$)				

Figure S2: Significant differences between visual analogue scale for breathlessness intensity (VAS-I) associations with Dyspnoea -12 (D-12) and Multidimensional Dyspnoea Profile (MDP) scores ($p < 0.002$)



A: Breathlessness in daily life: VAS-Intensity associations with Dyspnoea -12 (D-12) and Multidimensional Dyspnoea Profile (MDP) scores



B: Breathlessness at end exercise: VAS-Intensity associations with Dyspnoea -12 (D-12) and Multidimensional Dyspnoea Profile (MDP) scores

Section D: Consistency of descriptor choice across breathlessness assessments

Assessing individual consistency for sensory quality choices across breathlessness assessments raised a number of issues:

1. **Best match descriptor only included in one assessment.** The MDP sensory quality forced choice question permitted participants to nominate their “best match” for sensory quality (“Most accurately describes”). We did not request participants to volunteer or select their best descriptors for volunteered descriptors or nominate the “best match” when selecting “up to three” statements from Mahler et al’s list of 15 statements.
- 1) Where **individual MDP SQ forced choice items were not represented within the other three assessments** e.g no items for “tightness/constricted” within the D-12, no items for unpleasantness or emotions in descriptor list.
- 2) **Volunteered descriptors were always asked first**, but with subsequent breathlessness assessments completed in randomised order (Descriptor list, D-12 and MDP) which is likely to have provided participants with options that may have assisted /prompted or provided greater choices. Participants generally offered fewer volunteered descriptors then selected or indicated applied when presented with prescribed descriptors.
- 3) **Number of items** within sensory categories differing between instruments –i.e SQ for air hunger in MDP (1 item comprised of three phrases of similar meaning) versus 5 statements grouped for air hunger in Mahler et al’s descriptor list (based on our previous work Chest 2010; 138: 315-322) and 1 item in the D-12 (which is common to both Mahler’s list and MDP SQ –“I cannot get enough air”)
- 4) **Best choice versus choice “ever made”:** The MDP SQ “best choice” item presents a mutually exclusive sensory quality choice. When MDP SQ forced choice is used to compare across breathlessness assessments, this comparison is essentially exploring whether the “best choice” for sensory quality (MDP) was **ever** volunteered or selected in the remaining three breathlessness assessments.
- 5) **Participants rating more than item with the same highest score (whether highest intensity rating could be a proxy for best match item)** - Only two assessments required participants to rate items (D-12 and MDP). In both the MDP and D-12, participants frequently rated more than one sensory quality item /item group with the same high intensity rating. For the seven sensory quality items (physical) in the D-12, few participants rated a single item with score higher than any other item (breathlessness in daily life, n= 20 (24%); end exercise n = 25 (30%)). Whereas, for the five sensory quality items in the MDP, a greater percentage of participants rated a single item higher than any other item (daily life: n= 45 (54%); end exercise n= 51 (61%)). In both instruments, there was not a clear majority (>75%) where their highest intensity rating was limited to a single sensory quality item. In this sample at least, highest intensity ratings could not be confidently used as proxy for best match sensory quality.
- 6) Comparing the associations (r , r^2) between MDP and D-12 sensory quality intensity ratings provides group association but not individual consistency and is limited by the number of comparable items between the D-12 and MDP

Given these issues, we explored individual consistency across breathlessness assessments using 2x2 contingency tables for four items which shared similarities between the MDP and D-12 (Banzett et al 2015 - unpleasantness /uncomfortable, depressed, air hunger, work/effort) for both focal periods (daily life or end exercise) using the following processes:

- Descriptors from Mahler et al’s (1996) list were grouped according to the method reported in our previous work (Chest 2010; 138: 315-322) – Table 10 below.
- Three items in the D-12 are identical/near identical to descriptor statements in Mahler et al’s list but “My breath does not go in all the way” was not categorised as Air hunger in our grouping of Mahler’s statements but is grouped as a partial match with Air hunger statements in MDP sensory quality word grouping (Banzett et al 2015). Hence, two D-12 items were used: Air Hunger (“I cannot get enough air”) and Work/Effort (My breathing requires work). Unpleasantness /uncomfortable or depressed items not available in descriptor list (Mahler et al)
- For MDP and D-12 items scoring ≥ 1 (participant ever rated an item rather than MDP SQ forced choice or highest score).
- Consistency was calculated as the number of individual participants that selected or rated (≥ 1) the comparable item /item group across instruments with prescribed items (descriptor list, D-12 and MDP) expressed as a percentage (denominator = total sample minus the number of participants that did not select/rate the item in any instrument).
- Differences in the number of participants consistently indicating that the item applied (rated ≥ 1) in both the MDP and D-12 were assessed with McNemar’s test

	Unpleasant /uncomfortable	Air hunger /suffocating	Work/effort	Tight/constricted	Depressed
Volunteered	Examples: Uncomfortable, discomfort, annoying	e.g Difficulty breathing in, short of breath, cannot breathe, cannot take a deep breath, suffocating,	e.g Labour , tired, hard work	e.g tight, constricted	Examples: Helpless, depressed, regret
Descriptor list Mahler et al 1996	No items	I feel that I am smothering I feel that I am suffocating I cannot get enough air I feel out of breath I feel hunger for air	My breathing requires work My breathing requires effort	My chest feels constricted My chest feels tight	No items
Dyspnoea-12 Single items (rate at least "Mild")	My breathing is uncomfortable	I cannot get enough air	My breathing requires work	No items	My breathing makes me feel depressed
MDP Rate ≥ 1	Unpleasant or discomfort (A1)	I am not getting enough air or I am smothering or I feel hunger for air	My breathing requires muscle work or effort	My chest or lungs feel tight or constricted	Depressed

Table S10: Descriptors /items for each of the breathlessness assessments.

Breathlessness in daily life

Table S11: Breathlessness in daily life: individual consistency for sensory qualities across breathlessness assessments (D-12 and MDP ever rate ≥ 1)

Number of participants selecting or rating ≥ 1	Descriptor list (Mahler et al 1996) versus D-12				Descriptor list (Mahler et al 1996) versus MDP				D-12 vs MDP				Individual consistency across D-list:D-12:MDP N = (% calculated from 84 – X “not at all”)				Individual consistency Between D-12:MDP N = (% calculated from 84 – X “not at all”)			
Uncomfortable Mahler = NA D-12 =58 MDP =83									MDP	D-12			NA				N= 58 (70%)			
											Y	N								
										Y	58	25								
		N	0	1																
McNemars test									Chi square = 25.0 P < 0.000001*											
Air hunger Mahler n =56 D-12 = n = 51 MDP n = 56	D-12				D-List				D-12				3/3 agree = 33 (45%) 2/3 agree = 23 (32%) 1/3 = 17 (23%) Not at all = 11				N= 42 (66%)			
	D-List		Y	N	MDP		Y	N	MDP		Y	N								
		Y	38	18		Y	42	13		Y	42	13								
		N	13	15		N	14	15		N	9	20								
McNemars test									Chi square =0.73 P=0.39											
Work/effort Mahler n =38 D-12 n = 66 MDP = 52	D-12				D-List				D-12				3/3 agree = 20 (26%) 2/3 agree = 39 (51%) 1/3 = 18 (23%) Not at all = 7				N= 46 (64%)			
	D-List		Y	N	MDP		Y	N	MDP		Y	N								
		Y	30	8		Y	23	29		Y	46	6								
		N	36	10		N	15	17		N	20	12								
McNemars test									Chi square = 7.54 P= 0.006*											
Depressed D-12 =31 MDP =26					D-12								NA				N= 21 (58%)			
					MDP		Y	N												
						Y	21	5												
						N	10	48												
McNemars test					Chi square = 1.67 P= 0.20															
Key																				
*		McNemars test $p \leq 0.006$ (Bonfernoni corrected p value for significance)																		
Vol		Volunteered descriptors																		
D-List		15 item descriptor list Mahler et al (1996)																		
D-12		Dyspnea-12 items																		
MDPSQ FC		Multidimensional Dyspnoea Profile sensory quality forced choice question																		
Not at all		Participant did not volunteer or select item in any breathlessness assessment																		

Table S11: Breathlessness at end exercise: individual consistency for descriptors across breathlessness assessments.

Number of participants volunteering or selecting	Descriptor list (Mahler et al 1996) versus D-12	Descriptor list (Mahler et al 1996) versus MDP	D-12 vs MDP				Individual consistency across D-list:D-12:MDP N = (% calculated from 84 – X “not at all”)	Individual consistency Between D-12:MDP N = (% calculated from 84 – X “not at all”)						
Unpleasant Vol =11 D-12 =47 MDP =77			D-12				NA	N= 47 (61%)						
			MDP		Y	N								
				Y	47	30								
				N	0	7								
McNemars test			Chi square = 30.0 P= <0.00001*											
Air hunger Vol =30 Mahler =32 D-12 =40 MDP =35	D-12				D-List				D-12		3/3 agree = 16 (31%) 2/3 agree = 23 (44%) 1/3 = 13 (25%) Not at all = 32	N= 32 (74%)		
	D-List		Y	N	MDP		Y	N	MDP				Y	N
		Y	22	10		Y	17	18		Y			32	3
		N	18	34		N	15	34		N		8	41	
McNemars test			Chi square = 2.27 P = 0.13											
Work/effort Vol =31 Mahler =36 D-12 =57 MDP = 47	D-12				D-List				D-12		3/3 agree = 21(30%) 2/3 agree = 27 (38%) 1/3 = 23 (32%) Not at all =13	N= 37 (55%)		
	D-List		Y	N	MDP		Y	N	MDP				Y	N
		Y	27	9		Y	26	21		Y			37	10
		N	30	18		N	10	27		N		20	17	
McNemars test			Chi square = 3.33 P = 0.07											
Depressed Mahler - NA D-12 = 8 MDP =7							D-12		NA	N= 5 (50%)				
							MDP				Y	N		
								Y			5	2		
								N			3	74		
McNemars test			Chi square = 0.20 P= 0.65											
Key														
*	McNemars test p ≤ 0.006 (Bonfernoni corrected p value for significance)													
Vol	Volunteered descriptors													
D-List	15 item descriptor list Mahler et al (1996)													
D-12	Dyspnea-12 items													
MDPSQ FC	Multidimensional Dyspnoea Profile sensory quality forced choice question													
Not at all	Participant did not volunteer or select item in any breathlessness assessment													