Available online at www.sciencedirect.com
SCIENCE DIRECT
International Journal of Psychophysiology 59 (2006) 97 – 106
INTERNATIONAL JOURNAL OF PSYCHOPHYSIOLOGY
www.elsevier.com/locate/ijpsycho

Validation of the thoracic impedance derived respiratory signal using multilevel analysis

Jan H. Houtveen[a,*], Paul F.C. Groot[b], Eco J.C. de Geus[b]

[a]Department of Health Psychology, Utrecht University, P.O. Box 80140, 3508 TC Utrecht, The Netherlands
[b]Department of Biological Psychology, Vrije Universiteit, Amsterdam, The Netherlands

Received 9 November 2004; received in revised form 20 December 2004; accepted 20 February 2005
Available online 11 May 2005

Abstract

The purpose of the current study was to validate the change in thoracic impedance (dZ) derived respiratory signal obtained from four spot electrodes against incidental spirometry. Additionally, a similar validation was performed for a dual respiratory belts signal to compare the relative merit of both methods. Participants were 38 healthy adult subjects (half male, half female). Cross-method comparisons were performed at three (paced) respiration frequencies in sitting, supine and standing postures. Multilevel regression was used to examine the within- and between-subjects structure of the relationship between spirometric volume and the respiratory amplitude signals obtained from either dZ or respiratory belts. Both dZ derived respiratory rate and dual belts derived respiratory rate accurately reflected the pacing frequencies. For both methods, fixed factors indicated acceptable but posture-specific regression on spirometric volume. However, random factors indicated large individual differences, which was supported by variability of gain analyses. It was concluded that both the dZ and dual belts methods can be used for measurement of respiratory rate and within-subjects, posture-specific, changes in respiratory volume. The need for frequent subject-specific and posture-specific calibration combined with relatively large measurement errors may strongly limit the usefulness of both methods to assess absolute tidal volume and minute ventilation in ambulatory designs.

Keywords: Validation; Thoracic impedance; Respiration; Tidal volume; Respiratory rate

1. Introduction

Respiratory measures–time and volume components–are of great interest to the field of psychophysiology. Nonetheless, many practical and methodological considerations have impeded their widespread use, particularly in naturalistic settings. Speech, posture and physical activity, for instance, present a much larger problem to the measurement of respiration than they do to cardiovascular recordings. Moreover, in selecting their instruments, researchers of respiration are faced with a delicate balance between precision and intrusiveness. Although relatively unintrusive techniques like thermistors and single bands or strain gauges have been used in the past for obtaining respiratory rate, these methods provide only crude and imprecise information on respiratory volume. Volume can be accurately measured by techniques such as spirometry and pneumotachography, but these techniques are relatively intrusive (e.g., they alter spontaneous breathing by adding dead space and resistance). Therefore, these techniques are not well suited for continuous or ambulatory monitoring of subjects in naturalistic settings.

When properly used and calibrated, chest motion sensors can be used as an alternative to measure respiratory volume (Wientjes, 1992; Martinez et al., 1996; Earthrowl-Gould et al., 2001). In his review published in 1992, Wientjes argued that measurements of the separate motion of the rib cage and the abdomen provide the most accurate unintrusive estimation of the volume components of respiration. However, frequent individual rib and abdominal calibration of the

* Corresponding author. Tel.: +31 30 253 9303; fax: +31 30 253 4718.
E-mail address: j.h.houtveen@fss.uu.nl (J.H. Houtveen).

0167-8760/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpsycho.2005.02.003

measurement system was strongly recommended (Gribbin, 1983; Wientjes, 1992). Respiratory inductance plethysmography is one of the currently available techniques to measure movements of the rib cage and the abdomen (Morel et al., 1983; Watson, 1980). This technique has been successfully used in a number of recent studies (e.g., by using the LifeShirt™ system, see Wilhelm et al., 2003).

Another recently used technique (not reviewed by Wientjes in 1992) is based on the respiratory signal that can be derived from impedance cardiography. de Geus et al. (1995) showed that band-pass filtered thoracic impedance signals obtained from four spot electrodes can reliable be used for the assessment of respiratory rate. Ernst et al. (1999) took this further and validated the full respiratory signal, as obtained from impedance cardiography, against respiratory signals obtained in parallel from a respiratory belt and continuous direct spirometry. Validation was done during baseline, paced breathing, abdominal and thoracic breathing and a verbal arithmetic task. Transfer function analyses were used to compute the coherence (i.e., the cross-correlation of the thoracic impedance and spirometric signals in the frequency domain, which is an index of similarity of waveform morphology). The transfer function analysis of Ernst et al. showed that almost 90% of the variance of the spirometric signal was explained by the thoracic impedance signal.

However, as argued by Bland and Altman (1986, 1999), correlational analysis demonstrates that signals obtained from two methods may have a (linear) relationship, but it does not give insight in the absolute differences between methods and the degree these methods vary across subjects, for instance as function of the mean. They suggested the use of difference against mean values plots (Bland and Altman, 1986). This approach requires both signals to be equally scaled and cannot be applied to demonstrate a deviation between the unscaled thoracic impedance amplitude and spirometric volume. To deal with this problem, Ernst et al. (1999) computed variability of gain parameters (referred to in their paper as 'consistency of gain'). Gains can be obtained by dividing the observed thoracic impedance amplitudes (the output values) by the corresponding spirometric volumes (the input values). These values can be scaled within-subjects by dividing them by the mean value across conditions and between-subjects by dividing them by the mean value across subjects and conditions. This results in scaled gain values that can be interpreted as proportions to the within- and between-subjects mean. Variability of gain is defined as the standard deviation of these scaled gain values. Ernst et al. (1999) reported larger within- and between-subjects variability of gain for a respiratory belt signal as compared to the impedance signal and concluded that impedance cardiography signals, derived from a standard tetrapolar band electrode configuration, provide a more accurate measure of respiration.

In the study of Ernst et al. (1999), however, all breathing manipulations were limited to sitting conditions. This hampers the generalizability of the obtained results to situations where subjects frequently change posture, e.g., during ambulatory recording. Ideally, all cross-method comparisons should be performed across a range of different postures. The pioneering study of Ernst et al. (1999) can be extended on a number of further points. Their eight subjects (only one of which was female) were continuously breathing through the mouthpiece of a spirometer. This influences dead space and resistance and may therefore have created deviant breathing patterns. Furthermore, only one respiratory belt was used (i.e., at the level of the lower rib cage), while two are needed, according to Wientjes (1992), to optimize this method.

Here, we report on a validation study on 38 healthy adult subjects (half male, half female) that used three paced breathing conditions and incidental spirometry to validate thoracic impedance derived respiratory rate and amplitude. During incidental spirometry, subjects only breathed through a spirometer for the duration of a single breath. This was repeated twice, interspersed by a minute of unimpeded breathing. By generalizing from the single breaths, this procedure yielded the spirometric criterion, whilst leading to only minimal disturbance of the overall breathing pattern. In parallel to the thoracic impedance recording, a dual respiratory belt signal was recorded in order to compare both methods. Because these methods do not yield absolute respiratory volumes (the validity of such a calibration was exactly the main issue examined here), cross-method comparison was performed by comparing the relative amount of explained variance in the spirometric criterion.

To extend the generalizability of the findings (as compared to Ernst et al., 1999), cross-method comparisons were performed separately for a sitting, supine and standing posture combined with the paced breathing manipulations in a factorial design. Multilevel analysis was used to analyze the between- and within-subjects regression of the respiratory amplitudes from both explanatory signals (thoracic impedance and dual belts) on the spirometric volume. Within- and between-subjects gain and variability of gain values were additionally computed as a measure of the difference in performance between the two methods.

2. Methods

2.1. Subjects

The participants of the study were 38 healthy subjects aged 18–50, 20 men (age $M=30.1$, S.D.$=10.5$) and 18 women (age $M=29.8$, S.D.$=8.8$). Due to technical problems, only partial data are available for two participants resulting in 36 complete data sets. The study was presented as an investigation of breathing patterns. The participants received 10 euros after the experiment. All participants gave written informed consent.

2.2. Posture and paced breathing conditions

The experiment consisted of 12 different breathing conditions, each of which had a 2 min duration. In three different postures (sitting, supine and standing), the order of which was randomized, subjects breathed at four paced frequencies, the order of which was fixed: (a) spontaneous breathing without pacing, (b) paced breathing at a rate of 25 cpm without pause, (c) paced breathing at a rate of 15 cpm combined with 0.4 s pause and (d) paced breathing at a rate of 5 cpm combined with 1 s pause. A visual pacing signal was presented on a 15 in. monitor, positioned at 80 cm in front of them and consisted of a respiratory-like sinusoid signal with adjustable amplitudes (highest during slow breathing, lowest during fast breathing).

2.3. Physiological recordings

Recording of the changes in thoracic impedance (dZ) was performed by the Vrije Universiteit Ambulatory Monitoring System (VU-AMS) from a four spot electrode configuration (de Geus et al., 1995; de Geus and van Doornen, 1996). Two additional electrodes were used by this system for ECG measurements (results not reported here). The used four spot thoracic impedance electrode configuration was slightly different from that used by Qu et al. (1986). We have previously found that the Qu et al. (1986) configuration misses lower thorax movements due to respiration. Our configuration yields a clear ICG while simultaneously keeping the respiration signal intact, even during belly-breathing. Two electrodes on the back were used to continuously send a high frequency current of 50 kHz, 350 μA through the thorax. Two electrodes on the chest were used to measure impedance. The upper measuring electrode was placed at the jugular notch of the sternum between the collarbones. The lower measuring electrode was placed at the tip of the sternum (xiphoid process). The upper current electrode on the back was placed at least 3 cm above the horizontal plane of the measuring electrode. The lower current electrode on the back was placed at least 3 cm below the horizontal plane of the lower measuring electrode. The thoracic impedance signal was amplified and (analogue) band-pass filtered to obtain the impedance changes (dZ) due to respiration.

Changes in thoracic and abdominal circumference were measured with two BioPac TSD201 respiratory effort transducers (see http://www.biopac.com). The two respiratory belts were attached over (1) the rib cage at the level of the fifth thoracic vertebrae and (2) the abdomen at the level of the navel. The transducers have a variable resistance output between 50 and 125 kΩ. These signals were low-pass filtered (1 Hz), sampled at 25 Hz and registered by a BIOPAC MP150 system and the acqknowledge v3.7.3 software package.

Incidental spirometry was performed using the Micro plus Spirometer unit (Micro Medical, Rochester, UK). This device has an accuracy of ± 3%, which was verified by the 1000 ml ± 0.1% AEGER manual calibration pump. During each 2 min breathing condition, two incidental spirometry measurements took place by breathing through a 3 cm cannula after 30 and 90 s. Digitized tidal volumes from the spirometer device were sent to the registration computer through an RS-232 interface cable. A synchronization signal was received from the spirometer at the start of each incidental spirometric measurement. This signal was fed to the registration computer and to the VU-AMS for off line synchronization with the dZ and respiratory belts signals.

A capnometer (TG-951T CO_2 sensor Kit; Nihon Kohden Corporation) was connected to the spirometer device to control for episodes of hyperventilation during the paced breathing sessions (results not reported here).

2.4. Procedure

The experimental sessions took place between 10 a.m. and 4 p.m. and lasted approximately 45 min. The six recording electrodes were attached and connected to the thorax impedance measuring device. Next, the two respiratory belts were attached and the signals were checked. Subjects received a general explanation of the experimental protocol and were seated in a sound shielded and dimly lit cabin. Next, a practice session took place to familiarize the participants with paced breathing and incidental breathing through the spirometer. A nose clip was worn continuously during all breathing conditions. Next, the three 8 min paced breathing sessions (i.e., four times 2 min: normal breathing, 25 cpm, 15 cpm and 5 cpm) took place in sitting, supine and standing posture. Each paced breathing session was followed by a short break. At the end of the three paced breathing sessions, all equipment was disconnected and participants were debriefed and paid.

2.5. Signal analysis and data reduction

For an optimal comparison, both continuous respiratory signals (dZ and upper and lower respiratory belts) were processed by the same respiratory signal scoring software package (i.e., AMSRES; a manual is available at the VU-AMS website http://www.psy.vu.nl/vu-ams). This also allowed an easy interactive visual alignment of the spirometer signal with the dZ and respiratory belts signals. The dZ and upper and lower respiratory belts signals were all band-pass filtered using a FIR filter 0.05–0.5 Hz (3–30 cpm).

2.5.1. Amplitude and frequency from the continuous recordings

The continuous dZ, upper and lower respiratory belts signals were visually inspected and labeled for segments free of artifacts due to clipping. Interactive scoring of the peaks and throughs in these segments yielded the respiratory frequencies and amplitudes on a breath-to-breath basis.

These were averaged to yield a single mean respiratory rate and respiratory amplitude value for each breathing condition. The two breaths obtained during incidental spirometry were not included in these condition means. All continuous respiratory amplitude values were $\log^{10}(x+1)$ transformed to obtain normal distributions.

2.5.2. Amplitudes from incidental spirometry

Separate respiratory amplitudes were computed during spirometry in each of the 12 posture-by-breathing conditions. This was done for dZ, the upper and the lower belts signals. The amplitudes were visually inspected and amplitudes with artifacts or clippings were coded as missing. Clipping or artifacts were detected in 0.45% of the dZ observations, 2.1% of the upper belt observations and 4.8% of the lower belt observations. The majority of clippings and artifacts in the respiratory belt signals occurred during the 5 cpm paced breathing manipulation: 6.5% for the upper belt and 13.9% for the lower belt. For the other conditions, a mean value of 0.62% was found for the upper belt and 1.76% for the lower belt. The artifact-free breaths for the respiratory dZ and belts amplitudes were averaged (i.e., two incidental breaths per condition) and $\log^{10}(x+1)$ transformed to obtain normal distributions.

2.6. Statistical analysis

Repeated measures pacing by posture and pacing by posture by method ANOVA tests (implemented as MANOVA in SPSS 12) were performed on the respiratory rates obtained from the dZ and respiratory belts recordings to compare both methods, to test if the paced breathing manipulation was successful and to test whether the effects of the pacing manipulations were modulated by posture. Because a direct comparison of the unscaled peak-to-through amplitudes obtained from the dZ and the respiratory belts recordings to the spirometric volume is not meaningful, only the separate pacing by posture ANOVA tests were performed on the mean peak-to-through amplitudes obtained from the dZ and the respiratory belts signals. A final pacing by posture ANOVA test was performed on the volumes obtained from incidental spirometry.

Validity of the dZ and dual belts derived respiratory amplitudes was assessed by multilevel regression analysis, which is the most informative method, and by gain and variability of gain analysis, which allows us to compare the differences between the two methods and to compare our results directly to those reported by Ernst et al. (1999). The multilevel regression analysis (MLwiN1.10) was performed on *Z*-transformed values. The so-obtained standardized fixed factors provide the between-subjects regression coefficients, while random factors provide the variances of the intercepts and slopes on the subject level and the within-subjects residual variance. The incidental measured spirometric volume signal was predicted during spirometry, using as explanatory variables: (a) the dZ respiratory amplitude signal, (b) the upper belt respiratory amplitude signal, (c) the lower belt respiratory amplitude signal, (d) both upper and lower belts respiratory amplitude signals and (e) the mean belts respiratory amplitude signal. These analyses were performed for the posture-combined data set and separately for the sitting, supine and standing postures.

Finally, posture-combined and posture-specific gain and variability of gain values were computed (using the same data set but without the *Z*-transformations). The gains were computed as the ratios of the volumes obtained from spirometry and the amplitudes of dZ or the mean amplitudes of the two respiratory belts. The variances of the gain values were scaled (for the posture-combined and posture-specific values) in two ways: (a) between-subjects by using the across-subjects mean gain level and (b) within-subjects by using subject-specific mean gain levels. Thus, between and within variability of gain parameters were computed similar to Ernst et al. (1999), but they were computed and scaled for the posture-combined data set and separately for the sitting, supine and standing postures. Variability of gain values provide additional information to the random factors of the multilevel regression analysis because they are based on the original and unstandardized data set; the interpretation is therefore more meaningful in terms of the absolute differences between methods.

3. Results

3.1. Repeated measure ANOVA

The mean respiratory rates and respiratory amplitudes obtained during continuous recording without spirometry are shown in Figs. 1A,B and 2A,B, respectively. The mean spirometric volumes (in liters) and respiratory amplitudes (arbitrary units) across breaths with incidental spirometry are shown in Fig. 3A,B,C. Because the posture-specific fixed factors of the two respiratory belts were almost similar (as illustrated below in Table 1), we choose to display the mean belts values only, which was computed by averaging the upper and lower belts respiratory rates and amplitudes.

3.1.1. Continuous recordings, respiratory rate

Changes in paced breathing frequency were accurately reflected in respiratory rate obtained from the dZ signal ($F(3,31)=1312.27$, $p<0.001$) and from the mean belts signal ($F(3,31)=26{,}599.03$, $p<0.001$). For both signals, within-subjects contrast tests indicated differences between all three pacing frequencies (all p's<0.001). The no pacing condition, however, did not significantly differ from the 15 cpm condition. No gender or age effects on respiratory rate were found in either signal. However, a main effect of method ($F(1,36)=32.12$, $p<0.001$) and interaction effects between method and pacing ($F(3,34)=10.06$, $p<0.001$) and method and posture ($F(2,35)=6.68$, $p<0.01$) indicated that

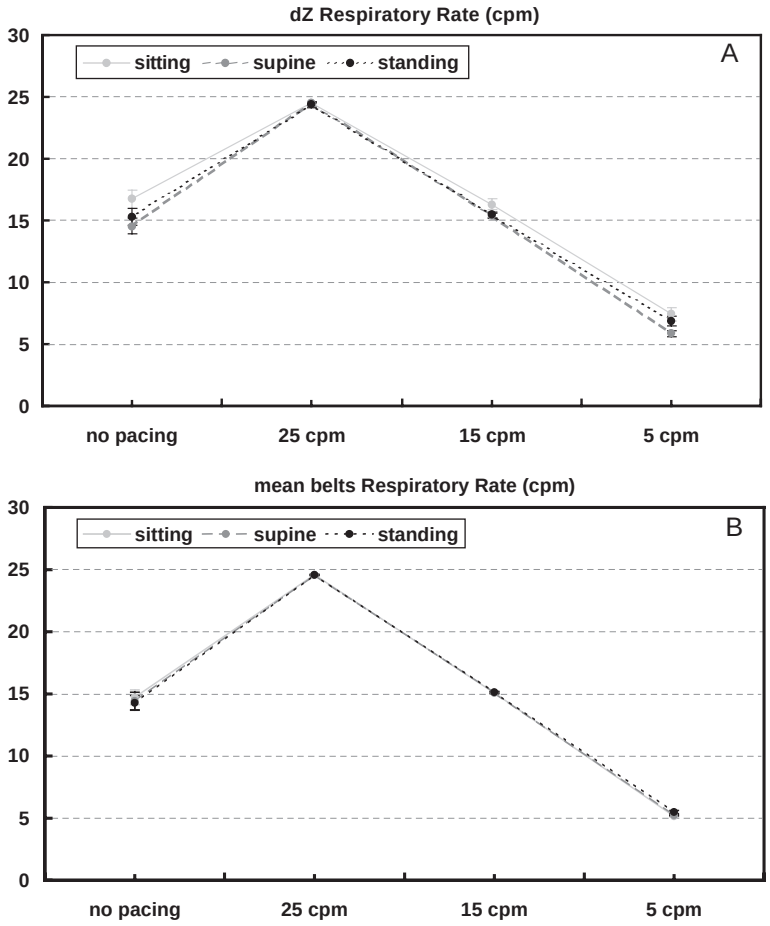

Fig. 1. (A, B) Mean (± S.E.M.) impedance (A) and dual belts (B) derived respiratory rate across the continuous recordings (in cycles per minute).

the methods do behave differently. A posture by pacing effect was found for the dZ signal ($F(6,28)=3.10$, $p<0.05$). Further decomposition showed that a higher rate was found during sitting than during supine or standing in the no pacing and 5 cpm conditions, but not in the 15 and 25 cpm conditions (p's<0.01). No such posture by pacing effects were found for the mean belts signal.

3.1.2. Continuous recordings, respiratory amplitude

Experimental manipulation of respiratory rate was faithfully reflected in the mean belts respiratory amplitude and dZ respiratory amplitude signals as shown by significant Pacing main effects ($F(3,31)=64.32$, $p<0.001$) and ($F(3,31)=63.47$, $p<0.001$). For both signals, within-subjects contrast tests indicated differences between all three pacing frequencies (all p's<0.001). The no pacing condition, however, again did not significantly differ from the 15 cpm condition. A main Posture effect was found for the dZ amplitude signal ($F(2,32)=35.38$, $p<0.001$): within-subjects contrast tests indicated differences between all three postures (all p's<0.01). A posture by pacing effect was also found for the dZ signal ($F(6,28)=5.29$, $p<0.01$; see Fig. 2A). No posture or posture by pacing effects were found for the mean belts signal. No gender or age effects were found for both signals in respiratory amplitude. Notice that method effects could not be tested for the amplitude values because both methods produced arbitrary amplitude units.

3.1.3. Incidental spirometry, volumes and amplitudes

In both the dZ and mean belts signals respiratory amplitude was increased for the breaths with spirometry as compared to continuous recordings without spirometry (p's<0.001; Fig. 3B,C as compared to Fig. 2A,B). As shown by main effects for pacing, the experimental manipulation of respiratory rate was significantly reflected in the spirometric volumes ($F(3,30)=92.61$, $p<0.001$), in the dZ amplitudes ($F(3,31)=55.11$, $p<0.001$) and in the mean belts amplitudes ($F(3,31)=34.21$, $p<0.001$). For all three signals, within-subjects contrast tests indicated differences between all pacing frequencies (all p's<0.001). The no pacing condition again did not significantly differ from the 15 cpm condition for the dZ and mean belts amplitude signals. An unexpected significant effect of posture was found on the spirometric volumes ($F(2,31)=11.94$, $p<0.001$). Within-subjects contrast tests indicated that less air was exhaled in the supine posture (p's<0.01; see Fig. 3A). Posture effects were also found for mean belts

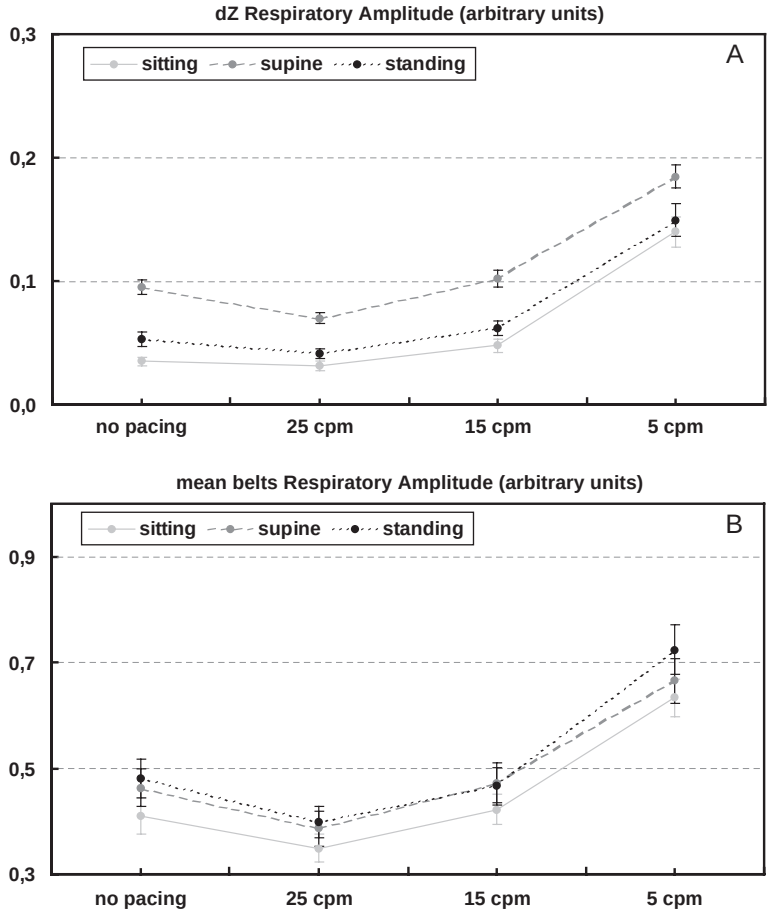

Fig. 2. (A, B) Mean (± S.E.M.) impedance (A) and dual belts (B) derived respiratory amplitude across the continuous recordings. Scaling of the amplitude changes in both signals is arbitrary. The *Y*-axis displays the $\log^{10}(x+1)$ transform of the recorded voltages.

amplitude ($F(2,32)=5.16$, $p<0.01$), but here within-subjects contrast tests indicated that less air was expired during both the sitting and supine positions (p's<0.01; see Fig. 3C). Posture effects on dZ amplitude were present but not significant ($F(2,32)=2.75$, $p=0.79$; Fig. 3B), possibly due to the larger residual variance of this method. No posture by pacing effect and no gender or age effects were found for spirometric volume, dZ amplitude and for mean belts amplitude values during incidental spirometry.

3.2. Multilevel regression analyses

The results of the multilevel regression analyses are shown in Table 1. The standardized β values can be used to compare the models across postures (posture-combined) and within each posture (posture-specific) condition. The indices of fit (deviances) and the residual variances can only be compared within a posture condition because the total variance (model 0) and the numbers of valid observations differed between conditions. Notice that the overall posture-combined values for the respiratory belt amplitude values cannot be considered very accurate: averaging the belts across posture was problematic because a significant ($p<0.001$) increase was found for the mean upper belt respiratory amplitude for sitting as compared to supine and standing, and a significant ($p<0.001$) decrease was found for the mean lower belt respiratory amplitude for sitting as compared to supine and standing. It should also be noticed that, within a posture condition, the model with the lowest Akaike's Information Criterion (AIC index = index of fit + $2q$, q is number of (additional) estimated parameters) can be considered the superior model; no additional statistical test is available (Hox, 2002). For model 4, the value 6 should be added to this index to compensate for the additional three parameters that are estimated.

The standardized β values were found to be much higher for the separate sitting, supine and standing regression analyses as compared to the overall values that include all postures. Thus, the posture-specific models performed superior as compared to the posture-combined model.

Comparison of the models within the posture conditions yielded mixed results. Posture-specific standardized β-values were almost similar for dZ respiratory amplitude as compared to the mean belts respiratory amplitude. Comparison of the index of fit (deviance) values between models indicated superior fits for model 1 (dZ) as compared to models 2–5 (belts). However, the residual variances were not always lower. In general, superior performances were

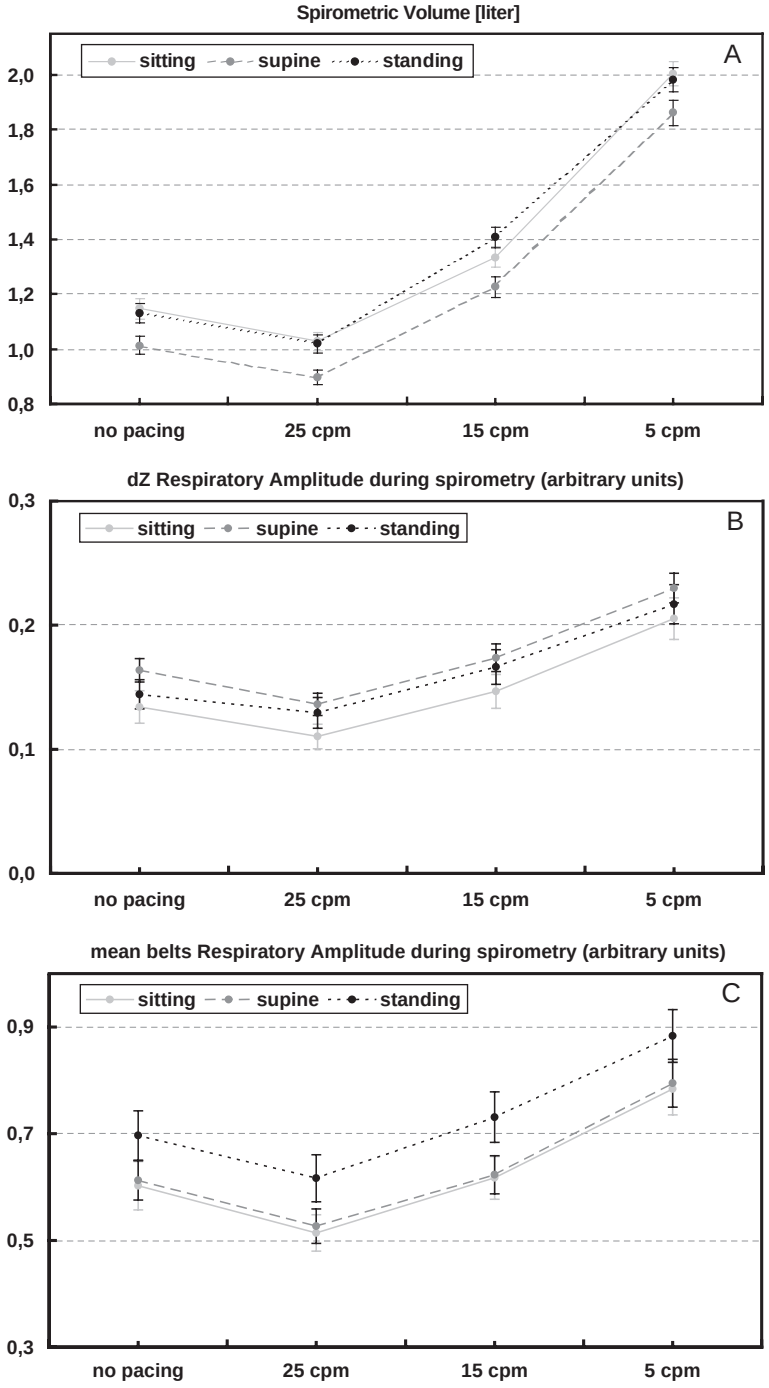

Fig. 3. (A, B, C) Mean (± S.E.M.) spirometric volume (A), impedance (B) and dual belts (C) derived respiratory amplitudes across breaths with incidental spirometry. Scaling of the amplitude changes in impedance and dual belts signals is arbitrary. The *Y*-axis in panels B+C displays the $\log^{10}(x+1)$ transform of the recorded voltages.

found for dZ respiratory amplitudes as compared to amplitudes from only one respiratory belt (i.e., model 1 versus models 2 and 3) and similar performances were found for dZ respiratory amplitudes as compared to the two belts respiratory amplitudes (i.e., model 1 versus models 4 and 5). Nonetheless, the random factor variances of intercepts and slopes at the subject level were found to be significant and relatively large for both methods.

In the multilevel analyses, we used the untransformed spirometric volumes as our prediction criterion, whereas the predictors, the amplitudes from the dZ and respiratory belts signals, had been $\log^{10}(x+1)$ transformed to obtain normality. To test for a possible effect of the transformation itself, we repeated the analyses using an identical $\log^{10}(x+1)$ transformation of the spirometric criterion as well. This did not meaningfully change the pattern of the results.

Table 1
Results of the multilevel regression analyses

	Posture-specific			Posture-combined
	Sitting	Supine	Standing	
Fixed factors model 1				
dZ	0.934	0.823	0.817	0.632
Fixed factors model 2				
Upper belt only	0.889	0.889	0.785	0.644
Fixed factors model 3				
Lower belt only	0.853	0.675	0.847	0.283
Fixed factors model 4				
Upper belt	0.681	0.559	0.443	0.751
Lower belt	0.640	0.607	0.695	0.296
Fixed factors model 5				
Mean belts	0.897	0.829	0.870	0.488
Random factors model 0				
Variance subject level	0.440	0.337	0.492	0.498
Residual variance	0.584	0.675	0.552	0.614
Index of fit	571	681	500	1685
Random factors model 1				
Intercept subject level	0.991	0.385	0.566	0.484
Slope–dZ subject level	0.254	0.119	0.265	0.075
Residual variance	0.160	0.236	0.159	0.343
Index of fit	394	483	339	1352
Random factors model 2				
Intercept subject level	1.672	1.148	1.017	1.013
Slope–upper belt subject level	0.463	0.683	0.590	0.413
Residual variance	0.274	0.393	0.289	0.423
Index of fit	507	642	447	1539
Random factors model 3				
Intercept subject level	0.945	0.784	1.106	0.506
Slope–lower belt subject level	0.381	0.252	0.323	0.084
Residual variance	0.282	0.429	0.228	0.518
Index of fit	489	629	411	1614
Random factors model 4				
Intercept subject level	1.871	1.023	1.247	1.520
Slope–upper belt subject level	0.351	0.195ns	0.143ns	0.536
Slope–lower belt subject level	0.302	0.229ns	0.181ns	0.098
Residual variance	0.160	0.329	0.198	0.319
Index of fit	441	599	396	1434
Random factors model 5				
Intercept subject level	1.872	1.153	1.350	0.773
Slope–mean belts subject level	0.397	0.344	0.305	0.169
Residual variance	0.186	0.329	0.192	0.424
Index of fit	446	597	393	1514

The response is the spirometric volume; the explanatory variables are: a constant (model 0), dZ respiratory amplitude (model 1), the upper belt respiratory amplitude only (model 2), the lower belt respiratory amplitude only (model 3), both belts respiratory amplitudes (model 4) and mean belts respiratory amplitude (model 5). The fixed factor intercepts are not displayed because they did not deviate significantly from zero. The values are standardized since all data were initially *Z*-transformed. The index of fit (deviance) is the − 2*log(likelihood) based on the IGLS estimate; ns=not significant.

3.3. Gain and variability of gain

The minimum, maximum and variance of the gain values are shown in Table 2. These values were scaled between-subjects (scaling-means computed across all within- and between-subject observations) and within-subjects (scaling-means computed separately for each subject). The values were again computed separately across postures (posture-combined) and within each posture (posture-specific). For both methods, variability of gain was smaller when scaled separately for each posture-specific condition as compared to posture-combined. For both dZ and the respiratory belts, the majority of variance was found to be at the subject level.

The scaled between-subjects gain and variability of gain estimates revealed superior minimum and maximum gain values (closer to 1) for the mean belts amplitude values as compared to the dZ amplitude values. Paired *t*-tests (standard deviation of gain values were computed within-subjects and tested across subjects) yielded reduced variance values for the mean belts amplitude values as compared to the dZ amplitude values for sitting and posture-combined (p's<0.05) and a trend for standing ($p<0.10$).

The scaled within-subjects gain and (residual) variability of gain estimates also revealed superior values for the mean belts amplitude values as compared to the dZ amplitude values: paired *t*-tests again yielded reduced variance values for the mean belts amplitude values as compared to the dZ amplitude values for sitting and posture-combined (p's<0.05) and a trend for standing ($p<0.10$). Thus, when interpreted in terms of the differences between methods, the

Table 2
Minimum, maximum and variance of gain values

	Posture-specific			Posture-combined
	Sitting	Supine	Standing	
dZ scaled between-subjects				
Minimum	0.16	0.32	0.22	0.19
Maximum	3.63	3.98	5.27	4.48
Variance	0.428	0.266	0.589	0.469
Variance subject level	0.281	0.149	0.314	0.190
Residual variance	0.147	0.117	0.275	0.279
Mean belts scaled between-subjects				
Minimum	0.21	0.20	0.18	0.17
Maximum	2.89	3.22	2.75	3.27
Variance	0.360	0.350	0.311	0.349
Variance subject level	0.312	0.273	0.260	0.237
Residual variance	0.048	0.077	0.051	0.112
dZ scaled within-subject				
Minimum	0.43	0.44	0.22	0.26
Maximum	2.48	2.23	2.24	3.19
Variance=residual variance	0.072	0.066	0.074	0.155
Mean belts scaled within-subjects				
Minimum	0.55	0.37	0.51	0.29
Maximum	1.82	2.01	1.67	2.31
Variance=residual variance	0.041	0.057	0.036	0.085

mean belts amplitude values performed superior as compared to the dZ amplitude values.

4. Discussion

Recording of changes in thoracic impedance (dZ) allows a relatively unintrusive assessment of respiratory rate and volume and can be used to monitor respiration even in naturalistic settings. Here we tested whether thoracic impedance derived respiratory rates correctly correspond to experimentally manipulated paced breathing frequencies and whether thoracic impedance derived respiratory amplitudes correspond to criterion volumes obtained from spirometry. Similarly, we compared respiratory rates and amplitudes based on the dual respiratory belts method with the experimentally induced paced breathing frequencies and with the spirometric criterion volumes. All testing was done across a number of paced breathing frequencies and in three different postures to make sure that validity would apply in the broad range of conditions likely to be encountered in naturalistic settings.

The results show that the thoracic impedance method tracks changes in respiratory rate very well, which corroborates previous reports on this method (Ernst et al., 1999; de Geus et al., 1995). A direct comparison with respiratory rate obtained from the dual respiratory belts method mildly favored the belts since the thoracic impedance signal was more inaccurate for very low respiratory rates and also showed more sensitivity to posture. The differences were subtle, however, and the between-subjects and within-subjects validity of thoracic impedance derived respiratory rate appears acceptable.

The use of the thoracic impedance derived respiratory amplitude signal to measure respiratory volumes proved to be much more complicated. In the multilevel regression models, thoracic impedance derived respiratory amplitudes (model 1) and dual belts derived respiratory amplitudes (models 4 and 5) were used to predict spirometric volumes. Based on the fixed (across-subjects) factors, both methods yielded moderate to high standardized regression coefficients on spirometric volume in all three postures. Variances of intercepts and slopes (on the subject level) were large for both methods. This is in accordance to Gribbin (1983) and Wientjes (1992) who recommend frequent subject-specific calibration. The original recommendation applied only to the dual belts method, but we can now extend it to the thoracic impedance method. Furthermore, standardized β values were much higher for each separate posture condition as compared to the overall posture-combined analyses. Thus, the results of the current study add to the existing recommendation by suggesting that calibration should not only be subject-specific but also posture-specific.

To compare the validity of thoracic impedance derived respiratory amplitudes and amplitudes based on the dual respiratory belts method against volumes obtained from spirometry, we used both multilevel regression and variability of gain analysis. Variability of gain values provides additional information regarding the degree in which these methods vary in validity across subjects. Note that we did not attempt to 'validate' the thoracic impedance derived respiratory signal against the belts directly, because a correlation of the errors could lead to faulty conclusions. Instead, we compared how well both signals covaried with the golden spirometric standard. We found the results from the multilevel analyses to be slightly superior for the thoracic impedance signal. The analyses of gain, however, showed that the amplitudes from the dual belts signal produced lower (residual) variability of gain values as compared to the amplitudes from the thoracic impedance signal. For both analyses, a very large variance component was found at the subject level. Thus, at the within-subjects level, changes in spirometric volume are tracked reasonably well by changes in thoracic impedance amplitudes and (slightly better) by changes in dual belts derived amplitudes, but both methods perform inferior at the between-subjects level.

These results appear to be more or less consistent with the results of Ernst et al. (1999) who reported better performance for the thoracic impedance derived respiratory signal as compared to a respiratory belt. Ernst et al. (1999), however, used only one belt and measured only in the sitting posture. For thoracic impedance, scaled across subjects and measured in the sitting posture (see Table 2), we found a standard deviation of gain of 0.65 at the subject level (i.e., the square root of 0.428) and 0.38 at the residual variance level (i.e., the square root of 0.147). These values are relatively large as compared to the values reported by Ernst and colleagues (0.32 between-subjects versus 0.20 within-subjects). Yet, our values for the mean of the two belts scaled across subjects (0.60, the square root of 0.360 and 0.22, the square root of 0.048) are lower as their reported single belt values (0.76 and 0.54). Thus, it can be concluded that amplitudes derived from the thoracic impedance respiratory signal may be more valid than amplitudes derived from a single belt to assess respiratory volumes, but that validity is inferior when compared to the mean amplitude obtained from two respiratory belts. However, based on the minimum and maximum gain values, the error made can be relatively large for specific subjects even with the dual belts method.

Two potential limitations of the current study should be noted. First, our change in thoracic impedance signal was obtained from four spot electrodes, whereas others, including Ernst et al. (1999), have obtained this signal from tetrapolar circumferential mylar-band electrodes secured at standard cervical and thoracic sites. Notwithstanding this difference in methodology, our standardized regression coefficients for thoracic impedance were comparable to the coherence values reported by Ernst et al. (1999). This suggests that the thoracic impedance signal obtained from four spot electrodes does not substantially differ from the thoracic impedance signal obtained from tetrapolar circum-

ferential mylar-band electrodes. In support of this, Khalafalla et al. (1970) directly compared the thoracic impedance respiratory signal obtained from pairs of spot electrodes with spirometric values and found an almost linear relationship between changes in the thoracic impedance using spot electrodes and changes in lung volume.

A second limitation of the current study (and other validation studies) is the observation that spirometry has an effect on respiratory volume (see also Askanazi et al., 1980). In the current study, more air was exhaled during the breaths with incidental spirometry as compared to the periods without spirometry. We deliberately wanted to keep the breaths with spirometric breathing as identical as possible to the breaths in the rest of the condition. Because our spirometry device requires a nose clip, this meant that we had to use the nose clip throughout. This does not invalidate a direct comparison between spirometry and the thoracic impedance or dual belts amplitudes during these breaths, but it does limit the generalizability of our findings to 'normal' breaths. Although such a limitation would obviously apply to continuous spirometry as well, we had hoped it to be relatively small in our incidental spirometry approach.

To summarize, both the thoracic impedance derived respiratory signal obtained from four spot electrodes and respiratory belts can be used to validly measure respiratory rate, both within- and between-subjects. In a laboratory setting, thoracic impedance and respiratory belts can also be used to measure within-subjects changes in respiratory volume. The mean belts amplitude signal yielded smaller variability of gain values as compared to the thoracic impedance signal, but based on the regression analyses both methods performed more or less similar in their relationships to the spirometric criterion. Nonetheless, between-subjects comparison of respiratory volume is problematic for both methods due to the relatively large measurement errors. Without frequent subject-specific and posture-specific calibration, their usefulness as unintrusive techniques to assess (changes in) respiratory volume in ambulatory psychophysiology seems limited.

Acknowledgments

This study was funded by a grant from the Netherlands Organization for Scientific Research (NWO No. 452-02-011). The authors gratefully acknowledge the aid of Marte Kaan and Ferdinand Schulte for their assistance in data collection and Marjolein Raaijmaakers for her assistance in scoring.

References

Askanazi, J., Silverberg, P.A., Foster, R.J., Hyman, A.I., Milic-Emili, J., Kinney, J.M., 1980. Effects of respiratory apparatus on breathing pattern. Journal of Applied Physiology 48, 577–580.

Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet 8, 307–310.

Bland, J.M., Altman, D.G., 1999. Measuring agreement in method comparison studies. Statistical Methods in Medical Research 8, 135–160.

Earthrowl-Gould, T., Jones, B., Miller, M.R., 2001. Chest and abdominal surface motion measurement for continuous monitoring of respiratory function. Proceedings of the Institution of Mechanical Engineers 215, 515–5120.

Ernst, J.M., Litvack, D.A., Lozano, D.L., Cacioppo, J.T., Berntson, G.G., 1999. Impedance pneumography: noise as signal in impedance cardiography. Psychophysiology 36, 333–338.

de Geus, E.J.C., van Doornen, L.J.P., 1996. Ambulatory assessment of parasympathetic/sympathetic balance by impedance cardiography. In: Fahrenberg, J., Myrtek, M. (Eds.), Ambulatory Assessment. Hogrefe Huber Publishers, Seattle.

de Geus, E.J.C., Willemsen, G.H.M., Klaver, C.H.A.M., van Doornen, L.J.P., 1995. Ambulatory measurement of respiratory sinus arrhythmia and respiration rate. Biological Psychology 41, 205–227.

Gribbin, H.R., 1983. Using body surface movements to study breathing. Journal of Medical Engineering and Technology 5, 217–223.

Hox, J., 2002. Multilevel Analysis: Techniques and Applications. Lawrence Erlbaum Associates, London.

Khalafalla, A.S., Stackhouse, S.P., Schmitt, O.H., 1970. Thoracic impedance gradient with respect to breathing. IEEE Transactions on Biomedical Engineering 17, 191–197.

Martinez, J.M., Papp, L.A., Coplan, J.D., Anderson, D.E., Mueller, C.M., Klein, D.F., Gorman, J.M., 1996. Ambulatory monitoring of respiration in anxiety. Anxiety 2, 296–302.

Morel, D.R., Forster, A., Suter, M., 1983. Noninvasive ventilatory monitoring with bellows pneumographs in supine subjects. Journal of Applied Physiology 55, 598–606.

Qu, M., Zhang, Y., Webster, J.G., Tompkins, W.J., 1986. Motion artifact from spot and band electrodes during impedance cardiography. IEEE Transactions on Biomedical Engineering 33, 1029–1036.

Watson, H., 1980. The technology of respiratory inductive plethysmography. In: Stott, F.D. (Ed.), ISAM Proceeding of the Second International Symposium on Ambulatory Monitoring. Academic, London.

Wientjes, C.J.E., 1992. Respiration in psychophysiology: methods and applications. Biological Psychology 34, 179–203.

Wilhelm, F.H., Roth, W.T., Sackner, M.A., 2003. The lifeShirt. An advanced system for ambulatory measurement of respiratory and cardiac function. Behavior Modification 27, 671–691.