Interoception and symptom reporting: disentangling accuracy and bias

Sibylle Petersen[1,2]*, Ken Van Staeyen[1], Claus Vögele[2], Andreas von Leupoldt[1] and Omer Van den Bergh[1]

[1] *Research Group on Health Psychology, KU Leuven, Leuven, Belgium,* [2] *Institute for Health and Behaviour, University of Luxembourg, Luxembourg, Luxembourg*

Anxiety and anxiety sensitivity are positively related to accuracy in the perception of bodily sensations. At the same time, research consistently reports that these traits are positively related to bias, resulting in the report of more and more intense symptoms that poorly correspond with physiological dysfunction. The aim of this study was to test the relationship of accuracy and bias in interoception. Furthermore, we tested the impact of individual differences in negative affect and symptom report in daily life on interoceptive accuracy and bias. Individuals higher in symptom report in daily life and negative affect were marginally more accurate in an interoceptive classification task in which participants were asked to identify different respiratory stimuli (inducing breathing effort) as belonging to a high or low intensity category. At the same time, bias in overestimating intensity of stimuli was significantly increased in participants higher in symptom report and negative affect, but only for more ambiguous stimuli. Results illustrate that interoceptive accuracy and bias need to be considered independently to understand their interaction with psychological factors and to disentangle (mis)perception of bodily sensations from liberal or conservative perceptual decision strategies.

Keywords: classification, interoception, bias, accuracy, decision strategies, sensitivity, negative affect, symptom report

Introduction

Nothing is closer to us than our own body but few things seem so elusive than the perception of bodily sensations. The brain is not passively 'measuring' signals from the body, but these signals interact with emotions and influence decision-making, behavior, and attention which in turn can change (perception of) bodily signals (Damasio, 1994; Craig, 2004). It is thus little surprising that correlations between self-reported bodily sensations and physiological changes are usually low in healthy individuals and in patients with somatic disease (e.g., Banzett et al., 2000; Petersen et al., 2011). Research consistently reports that negative affect, anxiety, and anxiety sensitivity are related to stronger nocebo effects (Van den Bergh et al., 1997, 2004) and higher levels of symptom report (Watson and Pennebaker, 1989; Petersen et al., 2011). Furthermore, in anxiety disorders, elevated self-report of somatic sensations is combined with a tendency for catastrophizing interpretations (e.g., Barlow, 1988; Barsky et al., 1994; Carleton et al., 2014).

In contrast to these low correlations between self-report and physiology particularly in individuals high in negative affect, a meta-analysis reports medium to strong effect sizes for a positive relationship between accuracy in heartbeat detection and psychological variables such as

anxiety sensitivity and anxiety (Domschke et al., 2010). Higher anxiety sensitivity is also related to lower detection thresholds for respiratory stimuli (Petersen and Ritz, 2011). Furthermore, a study measuring respiratory-related evoked potentials (RREPs) in participants watching negative emotional pictures found increased later components of RREPs in individuals higher compared to lower in anxiety (von Leupoldt et al., 2011). The authors interpreted these increases as higher motivated attention toward respiratory signals in more anxious individuals in negative affective contexts.

Thus, while individuals higher in negative affective states and traits seem to be more biased in interoception, they seem at the same time to be more accurate. Attempts to reconcile these seemingly paradox findings are faced with a methodological challenge. Research on interoceptive accuracy uses mostly heartbeat detection tasks (e.g., Domschke et al., 2010). In one type of heartbeat detection task, participants are asked to decide whether a sound signal is matching their heartbeat (e.g., Asmundson et al., 1993; Harver et al., 1993). Data collected in such a task could be used to differentiate bias and accuracy in a signal detection approach. Unfortunately, performance levels are around chance in the majority of participants, raising questions about the usefulness of this paradigm (Jones, 1994; Domschke et al., 2010). Most studies on interoception use another type of heartbeat detection task and follow a mental tracking paradigm in which participants are instructed to count their heartbeats during intervals of different length (Schandry, 1981). This paradigm, however, cannot differentiate sensitivity and bias.

Confounding sensitivity and bias means to lose crucial information. At the same level of sensitivity, individuals can apply liberal or conservative decision strategies which may result in very different forms of coping behavior following either a liberal "better safe than sorry" or a conservative "wait and see" approach (Macmillan and Creelman, 2004; Lynn and Barrett, 2014).

Making perceptual decisions and classifying bodily sensations, for example, as symptom or as benign sensation, is an inherent part of interoception (Petersen et al., 2014) and may present a challenge (Carleton et al., 2014). A headache may be painful, but *perhaps* not painful enough to take medication; breathlessness may be strong, but *perhaps* not strong enough to signal that we should stop exercise; heartbeat may be elevated, but *probably* not a sign of a heart disease. Although these sensations are clearly above detection threshold they are ambiguous regarding their category. We cannot be 100% sure about classification of ambiguous sensations at the border of two interoceptive categories, but we can optimize decision strategies by weighting the risk of missing a symptom against the risk of false alarms, that is, the risk of classifying a benign sensation as pathological (Griffiths et al., 2008). These processes of forming probabilistic beliefs about bodily sensations have been suggested to underlie interoception at every stage of information processing and may not always be deliberate or even consciously accessible (Edwards et al., 2012).

In this study, we tested the relationship between accuracy and bias in an interoceptive classification task in which participants were asked to correctly categorize different respiratory stimuli (respiratory loads increasing breathing effort) as belonging to either a low or a high intensity category and to indicate their place within these categories by labeling them as A1, A2, A3, A4, B1, B2, B3, and B4, with increasing numbers being related to increasing stimulus intensity (Petersen et al., 2014). A novel feature of this study is that we tested classification of a number of stimuli varying in intensity and assigned to two categories. Most signal detection paradigms test only one (usually low intensity) stimulus representing the signal against signal absent trials (for exceptions, see Kepecs et al., 2008; Yang et al., 2014). Interoceptive categories, however, usually subsume a range of different signals (e.g., signs of airway obstruction in asthma may come in different degrees of intensity and a variety of symptoms on multiple dimensions may fall in the category 'cold symptoms') and an important task is to classify ambiguous signals as belonging to one of various sensation classes.

We tested accuracy in the classification of multiple stimuli as belonging to one of two intensity categories. Furthermore, we tested how bias changes across the different stimuli within interoceptive categories. We expected that it would be more challenging to classify sensations closer to the border of two interoceptive categories than for more prototypical/central sensations, for example, a stimulus which is not clearly high or low, but moderate. It is important to note that also if non-categorization information would be given, an A1 stimulus would be more easily classified as low than an A4 stimulus. We tested whether interoceptive bias would be higher for more ambiguous sensations and whether individual differences in negative affect and symptom report in daily life would affect this bias. Higher increase in bias toward the shared category border could be interpreted a strategy to rather risk false alarms than to miss signals which indicated potential harm (Lynn and Barrett, 2014). Lower increase in bias toward the category border could be interpreted as the category border being incorporated into interoceptive decision-making, serving as a red flag and reducing misclassification particularly for stimuli at this category border.

We expected that higher bias for more ambiguous stimuli (better safe than sorry approach ignoring the category border) would lead to more accurate results in the classification task (as predicted by Lynn and Barrett, 2014), explaining the seeming paradox of higher bias and higher accuracy in individuals with negative affective expectations toward symptoms or high in general negative affect.

Materials and Methods

Participants

Participants were 54 women (mean age 21.04 years, SD = 1.8) without known chronic or acute disease. They were selected based on prescreens for high and low habitual report of bodily symptoms in daily life using the Checklist for Symptoms in Daily Life (Wientjes and Grossman, 1994). Participants completed this symptom checklist consisting of 39 bodily sensations (e.g., tingling in arms and hands, back pain, etc.) with regard to

how often they had experienced these sensations within the last year (scale 1 *never* −5 *very often*). The questionnaire has good reliability (Chronbach's alpha 0.70–0.90, Wientjes and Grossman, 1994). The pre-screen was filled in by 355 individuals and mean score was 91.95 (SD = 18.59). We invited participants with values in the symptom checklist of either higher than 80 or lower than 60 consecutively from this prescreen list.

Participants completed the same checklist also after the laboratory appointment. After the task, four participants from the high symptom reporter group did no longer reach high values on the symptom checklist and were excluded from the analysis. Thirteen of the participants in the low symptom report group reached values of 61–75 on the symptom checklist after the task (which induced symptoms), and were included in the low symptom report group. This resulted in $n = 25$ for the low symptom report group and $n = 25$ for the high symptom report group. Participants signed informed consent prior to participation. They received course credit or reimbursement for participation. The study was approved by the local ethics committee.

Instruments

We induced feelings of breathing effort by presenting eight different respiratory loads using the instrument Powerbreathe K5 (POWERbreathe International Ltd., Southam, UK). The instrument allows gradually increasing breathing resistance at inhalation by reducing the diameter of the breathing port. We used the software Breathelink to program presentation of loads. Exhalation was unrestricted. Resistances of the eight respiratory loads were chosen to ensure that differences between adjacent breathing loads could be distinguished by healthy volunteers, but were similar enough to leave an uncertainty margin for classification (7, 9, 11, 14, 18, 23, 28, and 37 cmH$_2$O).

Participants completed the Checklist for Symptoms in Daily Life (Wientjes and Grossman, 1994) twice during the experimental session. Firstly, they were asked to indicate which of the symptoms they had perceived in the last year to test whether results of the pre-screen were reliable. Secondly, we asked them to indicate which symptoms they had experienced during the actual task. We used the Positive and Negative Affect Schedule (Watson et al., 1988), a five point rating scale assessing negative and positive affect within the last 4 weeks with ten mood related adjectives as items for each subscale.

Protocol

At the start of the protocol, participants signed an informed consent form giving information about the study. In the first block of the interoceptive classification task, we presented the eight loads in random order, each load four times for two breaths. During presentation, the label of the load was presented on a computer screen (label A1 together with the 6 cmH$_2$O load, A2 – 9 cmH$_2$O, A3 – 11 cmH$_2$O, A4 – 14 cmH$_2$O, label B1 – 18 cmH$_2$O, B2 – 23 cmH$_2$O, B3 – 28 cmH$_2$O, and B4 – 37 cmH$_2$O). Participants were instructed that in the following block, they would have to solve a categorization task. They were asked to memorize the sensation and the label so that they would be able in a second block, when loads were

presented without label, to indicate the correct category (A or B) and the location of the stimulus within its category (1–4). In previous studies, we found that participants are able to distinguish and label eight different loads assigned to two categories after this training procedure above chance level (Petersen et al., 2014). Furthermore, in a study comparing a group of participants receiving category information and a control group receiving the information that stimuli were labeled with numbers only (increasing consecutively from lowest to highest stimulus), we found that categorization and an arbitrary category boundary indeed change perception of interoceptive stimuli. Differentiation between stimuli falling into the same categories was reduced and discrimination between categories was more pronounced in the experimental compared to the control group (Petersen et al., 2014, Study 1). Similar results were found in studies on visual perception using arbitrary category boundaries in a similar paradigm (Tajfel and Wilkes, 1963; Corneille et al., 2002), and in a study using an odor classification paradigm where mice were trained to classify six odor stimuli as belonging to one of two similar odor categories (Kepecs et al., 2008).

In Block 2, loads were presented again in pseudo-randomized order with 18 presentations per load, that is, 144 load presentations overall, with two breaths per load presentation. We presented the loads in a way that each load was preceded at least once by every other load to reduce the impact of order effects. Participants were asked to classify each load by assigning the correct labels (A1–B4). They did not receive feedback on their performance. After Block 2, participants completed the PANAS, the symptom checklist for symptoms last year, and the symptom checklist for symptoms directly after the task. Finally, we asked them to answer demographic questions on age, height, weight, and chronic or acute disease.

Data Analysis

We used SPSS 20 for all analyses and used the SPSS syntax for c and d' indices proposed by Stanislaw and Todorov (1999). We calculated d'_{class} measures reflecting how accurate participants were in distinguishing categories A and B. In contrast to signal detection, classification is a choice between two types of signals and not between signal and noise. Still, the formula is identical, only that A or B is treated as 'noise' and the other category as 'signal.' We calculated one mean d'_{class} index as mean of a d'_{Aclass} (treating A as signal and B as noise) and d'_{Bclass} (treating B as signal and A as noise).

Furthermore, we used signal detection analysis to calculate classification criteria c_{class}. We calculated $c_{class} = -(z(H) + z(F))/2$, with $z(H)$ as z-transformed hit rates of rating A if a load of Category A was presented, and $z(F)$ as z-transformed false alarm rates of reporting A1 if a load from B was presented. This resulted in four c_{class} indices per category ($c_{classA1}$, $c_{classA2}$, $c_{classA3}$, $c_{classA4}$, $c_{classB1}$, $c_{classB2}$, $c_{classB3}$, and $c_{classB4}$). Thus, c_{class} reflects bias, taking into account classification errors and correct classification with c indices below zero indicating a liberal bias and c indices above zero indicating a conservative bias.

In a classification paradigm, the meaning of the terms liberal and conservative depends on the standard used. A tendency to require little evidence to classify a stimulus as B is liberal in a

"better safe than sorry" approach, if missing B is regarded to be more costly than missing A. We calculated c_{class} indices in a way that the term 'liberal' refers to a decision strategy that favors classifying A as B for stimuli actually belonging to A, or classifying B as A for stimuli actually belonging to B. In other words, a liberal tendency in our analysis is a strategy that favors misclassification. We expect this liberal tendency to be strongest for more ambiguous stimuli and to become weaker for clearly high or clearly low stimuli.

We tested in a univariate analysis of covariance (ANCOVA) whether d'_{class} would differ between groups of high and low habitual symptom reporters, controlling for negative affect and symptoms experienced during the task as covariates. In further repeated-measures ANCOVAs we tested whether c_{class} values would be significantly reduced as functions of closeness to the shared category border between A and B. The within-individual factor was Classification with four levels of the within-individual factor Classification per category ($c_{classA1}$, $c_{classA2}$, $c_{classA3}$, $c_{classA4}/c_{classB1}$, $c_{classB2}$, $c_{classB3}$, $c_{classB4}$). We included Symptom Report (high/low) as between-individual variable to test whether individuals higher in habitual symptom report would show more liberal tendencies to assign stimuli to the B category. Again, we included negative affect and symptoms experienced during the task within the ANCOVA model as covariates.

As alternative way of analysis, we performed an individual regression slopes analysis, following the procedure described by Pfister et al. (2013). In this analysis, we tested whether the individual regression slopes of participants indicating a decline toward a more liberal bias from A1 to A4 and an increase to more conservative bias from B1 to B4 would be steeper for participants in the High Symptom Report group. We compared the single slopes in a one way ANCOVA with negative affect as covariate and symptoms reported directly after the task included as control variable. Furthermore, we calculated the absolute value for the slopes across Categories A and B and compared, whether steepness of slopes would differ between A and B in a repeated-measures ANCOVA with negative affect as covariate, controlling for symptoms experienced during the task. Furthermore, in the supplemental material we present data on accuracy of differentiation within categories.

Results

Table 1 summarizes mean values across groups of high and low symptom reporters for self-report scales. We tested group differences in a t-test for independent groups.

Accuracy of Classification

High symptom reporters were marginally more accurate classifying loads as A or B, main effect for Symptom Report $F(1,49) = 3.60$, $p = 0.064$, $\eta_p^2 = 0.073$ (**Figure 1**). Effects for negative affect and symptoms reported during the task were not significant, $Fs < 1$.

Misclassifications and Response Bias

Classification bias c_{class} per load became less conservative toward the shared category border between A and B in both groups, but this decrease was significant only for Category A $F(3,44) = 18.27$, $p < 0.001$, $\eta_p^2 = 0.555$, and not for Category B $F(3,44) = 2.06$, $p = 0.119$, $\eta_p^2 = 0.123$ (**Figure 2**, statistics are reported with the covariate Negative Affect). Please note that the lack of
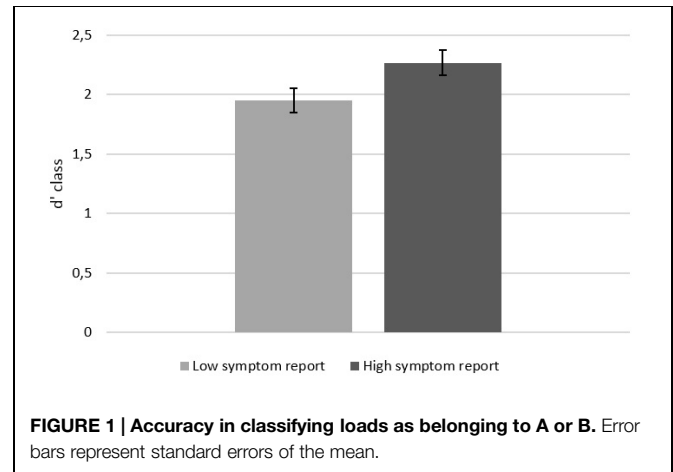

FIGURE 1 | Accuracy in classifying loads as belonging to A or B. Error bars represent standard errors of the mean.
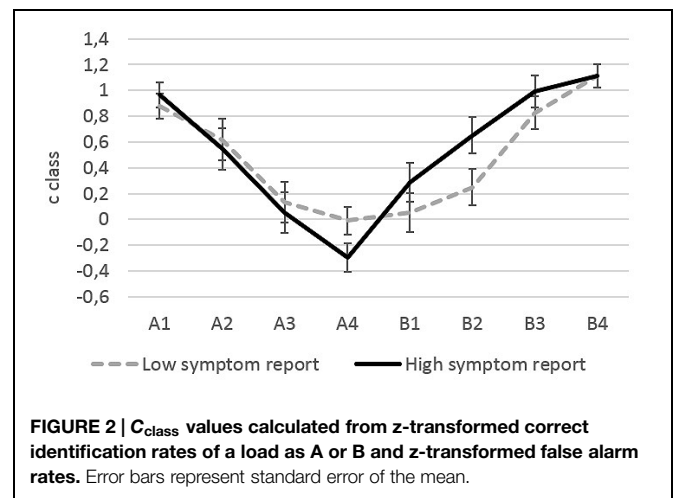

FIGURE 2 | C_{class} values calculated from z-transformed correct identification rates of a load as A or B and z-transformed false alarm rates. Error bars represent standard error of the mean.

TABLE 1 | Participants' characteristics (standard deviation in parentheses).

	Symptom experienced last year	Negative affect	Symptoms experienced during the task
Low symptom reporters	60.12* (SD = 7.05, range: 42–75)	16.84* (SD = 7.10, range: 10–34)	57.04* (SD = 9.24, range: 41–75)
High symptom reporters	99.88 (SD = 14.06, range: 81–138)	24.48 (SD = 7.84, range: 11–28)	78.68 (SD = 16.47, range: 41–107)

*$p < 0.001$.

significant results for Category B does not necessarily imply that the effect for Category A was larger than for Category B. In the next section, we report a comparison of effects with individual slope analysis. The decrease in Category A, was significantly stronger in high symptom reporters, interaction Classification X Symptom Report $F(3,44) = 3.97$, $p = 0.014$, $\eta_p^2 = 0.213$. This interaction was non-significant for category B, $F(3,44) = 1.65$, $p = 0.192$, $\eta_p^2 = 0.101$. In other words, across Category A, participants high in habitual symptom report showed a more pronounced change from a conservative to a liberal criterion than participants low in habitual symptom report. Also after exclusion of the covariate Negative Affect from this model, the interaction effect of Classification X Symptom Report remained significant, $F(3,44) = 3.18$, $p = 0.033$, $\eta_p^2 = 0.175$.

For Category A, the interaction of the within-individual factor Classification with the covariate Negative Affect was marginally significant, $F(3,44) = 2.29$, $p = 0.092$, $\eta_p^2 = 0.135$, but not for Category B, $F(3,46) = 2.06$, $p = 0.120$, $\eta_p^2 = 0.123$. None of the other effects was significant (all $Fs < 1.18$).

Post hoc tests revealed that none of the single c_{class} indices differed significantly between groups regardless whether negative affect was included in the model or not (all $ps > 0.103$). We performed explorative *post hoc* tests with *t*-tests for independent groups to test whether the c_{class} index for A4 indicated indeed a *liberal* bias or only a *less conservative* bias. This explorative *t*-test testing differences from zero, suggests that only for high habitual symptom reporters, c_{class} for A4 was significantly smaller than zero, that is, only in this group bias changed significantly from conservative to liberal $t(24) = -2.327$, $p = 0.027$.

Individual Slopes Analysis

Individual slopes analysis revealed that the shift toward a liberal bias (as expressed in a negative slope for Category A) was stronger in the High Symptom Report group (mean: -0.386, SD $= 0.143$) compared to the Low Symptom Report group (mean: -0.357, SD $= 0.126$), $F(1,47) = 4.59$, $p = 0.037$, $\eta_p^2 = 0.089$. The effect of negative affect was non-significant and we report statistics without including negative affect as covariate. For Category B, the shift to a more conservative bias from B1 to B4 did not differ significantly between groups of Low (mean: 0.366, SD $= 0.159$) and High Symptom Report (mean: 0.293, SD $= 0.184$), $F(1,47) = 2.34$, $p = 0.128$, $\eta_p^2 = 0.049$. Comparing the absolute value of single slopes (ignoring the direction of slope) in a repeated-measures ANCOVA revealed that the steepness of slope was marginally higher for Category A (mean: 0.375, SD $= 0.125$) than for Category B (mean: 0.339, SD $= 0.155$), main within-individual effect $F(1,46) = 3.27$, $p = 0.077$, $\eta_p^2 = 0.065$. The interaction with the between-individual factor Symptom Report was significant, $F(1,47) = 4.64$, $p = 0.036$, $\eta_p^2 = 0.090$. While for participants from the Low Symptom Report group the absolute amount of slopes for Category A (mean: 0.357, SD $= 0.126$) and B (mean: 0.366, SD $= 0.158$) were not significantly different ($p = 0.828$), the absolute amount of slopes differed significantly for the High Symptom Report group between Category A (mean: 0.392, SD $= 0.125$)

and B (mean: 0.311, SD $= 0.150$) with a steeper slope for Category A. In other words, change in bias across categories was stronger for category A than for Category B, but only in the High Symptom Report group. Again negative affect had no significant effect and including it did not change patterns of significance and we report results without including this covariate.

Correlations between Bias and Accuracy

Correlations (Pearson's correlation coefficient, two-tailed, **Figure 3**) between d'_{class} and c_{class} were negative for category A, but significant for stimulus A4 only. In other words, higher accuracy was related to a more liberal decision criterion for the stimulus at the category border of A. This 'better safe than sorry' approach for this particularly ambiguous stimulus was related to overall better classification results. Furthermore, d'_{class} was significantly positively related to c_{class} for stimulus B1 and B2, that is, higher accuracy was related to a more conservative decision criterion (being reluctant to decide that stimuli B1 and B2 belonged to A) for stimuli at the border of category B. None of the other correlations was significant, $p > 0.053$.

Discussion

Participants higher in symptom report in daily life scored higher on negative affect and reported more symptoms during the experimental task. They were also marginally more accurate in the interoceptive classification task. This is consistent with results of prior research showing a relationship between interoceptive accuracy/awareness and negative affective traits and states (e.g., Domschke et al., 2010; Petersen et al., 2011; von Leupoldt et al., 2011). A novel finding of this study is that classification strategies changed across categories. Bias became more liberal with increasing closeness to the border between categories. This change from conservative to more liberal decision strategies was significantly stronger in high symptom reporters and positively related to classification sensitivity d'_{class}. Furthermore, for Category A (but not for B), the decline of individual slopes was steeper in the High compared

FIGURE 3 | Correlation coefficients between c_{class} indices and d'_{class} ($p < 0.01$).
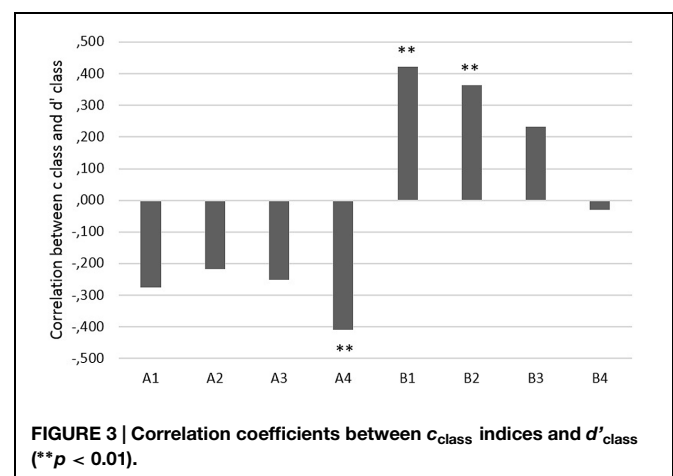
to the Low Symptom Report group. Thus, the information that there was a category border between A4 and B1 was effectively reducing bias toward the category border, but only in participants who did not have strong affective expectations about stimuli.

Results confirm hypotheses proposed recently in signal detection research. Lynn and Barrett (2014) have suggested that increased uncertainty about visual stimuli will lead to more extreme decision strategies in signal detection tasks. These more extreme strategies for more ambiguous stimuli are suggested to *optimize* decision-making and behavior (Lynn et al., 2012). Lynn and Barrett (2014) give the example of a person who walks more carefully in a dimly lit room compared to a brightly lit room (i.e., under higher or lower ambiguity of visual signals). Optimizing speed relative to sight allows avoiding injury or breaking objects. They suggest that "extreme bias may reflect not an impairment, but a normal adaptive mechanism that offsets the single impairment, poor sensitivity" (Lynn and Barrett, 2014, p. 1670). Calibrating bias to increased uncertainty for ambiguous sensations (closer to category borders) may be highly adaptive. More extreme decision strategies under higher ambiguity (closeness to category borders) were more successful in the present study. Higher classification accuracy (d'_{class}) was related to *more* bias at the borders of categories and increased bias in this laboratory task was related to bias in symptom report in daily life.

It is important to note that the border of the two neutrally labeled categories A and B was not intrinsically meaningful, as it would be the case for a border that marks a transition between a sensation and a symptom category (e.g., increased airway resistance which is either benign or indicates an oncoming asthma attack). It could be questioned whether arbitrary category boundaries as such affected perception. Prior research using a design with an experimental group receiving category information and a control group for which stimuli were numbered consecutively from lowest to highest without categories has found that arbitrary category boundaries between categories A and B affect interoception (Petersen et al., 2014, Study 1). Differentiation between categories was increased compared to differentiation within categories in the experimental group (receiving category information) compared to differentiation between identical stimuli in control groups (not receiving category information). Similar effects have also been found for visual perception (e.g., Tajfel and Wilkes, 1963; Corneille et al., 2002). Furthermore, a study testing classification decisions in mice which were trained to classify six odor stimuli as belonging to two categories (which shared a boundary and for which similarity between stimulus A3 and B1 was the same as between A3 and A2) found that detection of higher ambiguity for stimuli closer to category boundaries does not require meta-cognition and that ambiguity determines speed of decisions, overall accuracy, and is correlated with prefrontal cortex activity (Kepecs et al., 2008). For meaningful category labels and borders, or a paradigm which would trigger meta-cognition in participants about the costs of misclassifying stimuli, results might be even stronger than observed in this minimal paradigm using neutral labels.

It is tempting to speculate that the observed increasingly liberal bias for ambiguous stimuli at category borders in high symptom reporters may be a first step toward interoceptive threat-generalization. If a stimulus at a category border is consistently misclassified as belonging to a more intense category, this may lead in the long run to establishing a new and lower category border. This process may continue until more and more stimuli, which were initially in a low (or "safe sensation") category, are subsumed in a higher (or "dangerous symptom") category. Following this lead, fear-generalization in anxiety disorders (Lissek et al., 2008) may be interpreted as increasingly more liberal strategies spreading from category borders to more and more stimuli within a category.

Ambiguity of sensations because of their location close to a category border and their no longer clearly low or clearly high magnitude may be clinically as relevant as ambiguity related to a poor signal to noise ratio of sensations (i.e., minimal stimulation such as heartbeat at rest). Research on the role of uncertainty in panic disorder found that intolerance of uncertainty was substantially and significantly related to symptom report in panic disorder even after controlling for anxiety sensitivity (Carleton et al., 2014). Results from this correlation study suggest that patients seemed to find uncertainty about the decision whether a clearly detectable sensation is a sign of pathology (or not) at least as aversive as the sensation as such. Probing deeper into interoceptive classification strategies in anxiety disorder, future research should test whether the relationship of negative affective states and traits and interoceptive accuracy and bias is mediated by increased feelings of aversiveness related to uncertainty about classification of sensations as pathological or benign. Trait constructs such as intolerance for ambiguity which are closely related to anxiety may be interesting in that regard (Birrell et al., 2011).

Negative affect was increased in individuals high in habitual symptom report. Negative affect has been suggested to be related to a general lack of inhibition and not specifically to interoception (Bogaerts et al., 2015), but again, our results suggest that this effect is not a general effect across all types of stimuli within a category, but higher for more ambiguous stimuli. The relationship between negative affect and perceptual bias has been confirmed in a large number of studies (e.g., Bar-Haim et al., 2007; Cisler and Koster, 2010), but testing decision strategies under uncertainty may help to shed light on the processes underlying this relationship.

A limitation for generalization of the present results is that categories were distinct and defined only by one dimension (inspiratory resistance). Interoceptive categories are multidimensional and may overlap on some dimensions and be distinct on others. Sensations experienced during an asthma attack, for example, and sensations related to panic attacks belong to two distinct diagnostic categories, but are partly overlapping in how they are experienced by patients (Lehrer et al., 2002). Multidimensionality and overlap of categories in interoception may increase ambiguity of sensations, which in turn may lead to more extreme forms of bias for ambiguous sensations.

A further important limitation of this study is that we included only young female participants. Attention to interoceptive stimuli as well as expression of negative affect and fear may be higher in women than in men. Future research needs to address gender and age differences in interoceptive classification.

Conclusion

The relationship between sensitivity and bias is not uniform across interoceptive categories. To understand the relationship between interoception and individual difference variables, paradigms are needed which do not confound sensitivity and bias and vary the degree of ambiguity stimuli have regarding their classification.

References

Asmundson, G. J., Sandler, L. S., Wilson, K. G., and Norton, G. R. (1993). Panic attacks and interoceptive acuity for cardiac sensations. *Behav. Res. Ther.* 31, 193–197. doi: 10.1016/0005-7967(93)90071-2

Banzett, R. B., Dempsey, J. A., O'Donnell, D. E., and Wamboldt, M. Z. (2000). Symptom perception and respiratory sensation in asthma. *Am. J. Respir. Crit. Care Med.* 162, 1178–1182. doi: 10.1164/ajrccm.162.3.9909112

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., and van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and non-anxious individuals: a meta-analytic study. *Psychol. Bull.* 133, 1–24. doi: 10.1037/0033-2909.133.1.1

Barlow, D. (1988). *Anxiety and its Disorders: The Nature and Treatment of Anxiety and Panic.* New York, NY: Guilford Press.

Barsky, A. J., Cleary, P. D., Sarnie, M. K., and Ruskin, J. N. (1994). Panic disorder, palpitations, and the awareness of cardiac activity. *J. Nerv. Ment. Dis.* 182, 63–71. doi: 10.1097/00005053-199402000-00001

Birrell, J., Meares, K., Wilkinson, A., and Freeston, M. (2011). Toward a definition of intolerance of uncertainty: a review of factor analytical studies of the Intolerance of Uncertainty Scale. *Clin. Psychol. Rev.* 31, 1198–1208. doi: 10.1016/j.cpr.2011.07.009

Bogaerts, K., Rayen, L., Lavrysen, A., Van Diest, I., Janssens, T., Schruers, K., et al. (2015). Unraveling the relationship between trait negative affectivity and habitual symptom reporting. *PLoS ONE* 10:e0115748. doi: 10.1371/journal.pone.0115748

Carleton, R. N., Duranceau, S., Freeston, M. H., Boelen, P. A., McCabe, R. E., and Antony, M. M. (2014). "But it might be a heart attack": intolerance of uncertainty and panic disorder symptoms. *J. Anxiety Disord.* 28, 463–470. doi: 10.1016/j.janxdis.2014.04.006

Cisler, J. M., and Koster, E. H. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: an integrative review. *Clin. Psychol. Rev.* 30, 203–216. doi: 10.1016/j.cpr.2009.11.003

Corneille, O., Klein, O., Lambert, S., and Judd, C. M. (2002). On the role of familiarity with units of measurement in categorical accentuation: Tajfel and Wilkes (1963) revisited and replicated. *Psychol. Sci.* 13, 380–383. doi: 10.1111/j.0956-7976.2002.00468.x

Craig, A. D. (2004). Human feelings: why are some more aware than others? *Trends Cogn. Sci.* 8, 239–241. doi: 10.1016/j.tics.2004.04.004

Damasio, A. R. (1994). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 351, 1413–1420. doi: 10.1098/rstb.1996.0125

Domschke, K., Stevens, S., Pfleiderer, B., and Gerlach, A. L. (2010). Interoceptive sensitivity in anxiety and anxiety disorders: an overview and integration of neurobiological findings. *Clin. Psychol. Rev.* 30, 1–11. doi: 10.1016/j.cpr.2009.08.008

Author Contributions

SP and OVdB contributed equally to the development of the study concept and study design. AvL contributed conceptually to the study design and data analysis. Testing, data collection, and data analysis were performed by SP, and KvS in collaboration with OVdB as senior researcher. All authors, SP, OVdB, KvS, AvL, and CV drafted the paper. All authors approved the final version of the paper for submission.

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2015.00732/abstract

Edwards, M. J., Adams, R. A., Brown, H., Pareés, I., and Friston, K. J. (2012). A Bayesian account of 'hysteria'. *Brain* 135, 3495–3512. doi: 10.1093/brain/aws129

Griffiths, T. L., Kemp, C., and Tenenbaum, J. B. (2008). "Bayesian models of cognition," in *Cambridge Handbook of Computational Cognitive Modeling*, ed. R. Sun (Cambridge, MA: Cambridge University Press), 59–100.

Harver, A., Katkin, E. S., and Bloch, E. (1993). Signal-detection outcomes on heartbeat and respiratory resistance detection tasks in male and female subjects. *Psychophysiology* 30, 223–230. doi: 10.1111/j.1469-8986.1993.tb03347.x

Jones, G. E. (1994). "Perception of visceral sensations: a review of recent findings, methodologies, and future directions," in *Advances in Psychophysiology*, eds J. R. Jennings, P. K. Ackles, and M. G. H. Cole (London: Jessica Kingsley Publishers), 55–191.

Kepecs, A., Uchida, N., Zariwala, H. A., and Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature* 455, 227–231. doi: 10.1038/nature07200

Lehrer, P., Feldman, J., Giardino, N., Song, H. S., and Schmaling, K. (2002). Psychological aspects of asthma. *J. Consult. Clin. Psychol.* 70, 691–711. doi: 10.1037/0022-006X.70.3.691

Lissek, S., Biggs, A. L., Rabin, S. J., Cornwell, B. R., Alvarez, R. P., Pine, D. S., et al. (2008). Generalization of conditioned fear-potentiated startle in humans: experimental validation and clinical relevance. *Behav. Res. Ther.* 46, 678–687. doi: 10.1016/j.brat.2008.02.005

Lynn, S. K., and Barrett, L. F. (2014). "Utilizing" signal detection theory. *Psychol. Sci.* 25, 1663–1673. doi: 10.1177/0956797614541991

Lynn, S. K., Zhang, X., and Barrett, L. F. (2012). Affective state influences perception by affecting decision parameters underlying bias and sensitivity. *Emotion* 12, 726–736. doi: 10.1037/a0026765

Macmillan, N. A., and Creelman, C. D. (2004). *Detection Theory: A User's Guide.* Hove: Psychology press.

Petersen, S., and Ritz, T. (2011). Situational self-awareness and report bias in the detection of respiratory loads in healthy individuals: over-report is linked to external focus. *Br. J. Health Psychol.* 15, 401–416. doi: 10.1348/135910709X466676

Petersen, S., Schroijen, M., Moelders, C., Zenker, S., and Van den Bergh, O. (2014). Interoceptive categorization: the role of perceptual organization on processing sensations from inside. *Psychol. Sci.* 25, 1059–1066. doi: 10.1177/0956797613519110

Petersen, S., van den Berg, R., Janssens, T., and Van den Bergh, O. (2011). Illness and symptom perception: a theoretical approach towards an integrative measurement model. *Clin. Psychol. Rev.* 31, 428–439. doi: 10.1016/j.cpr.2010.11.002

Pfister, R., Schwarz, K. A., Carson, R., and Janczyk, M. (2013). Easy methods for extracting individual regression slopes: comparing SPSS, R, and Excel. *Tutor. Quant. Methods Psychol.* 9, 72–78.

Schandry, R. (1981). Heartbeat perception and emotional experience. *Psychophysiology* 18, 483–488. doi: 10.1111/j.1469-8986.1981.tb02486.x

Stanislaw, H., and Todorov, N. (1999). Calculation of signal detection theory measures. *Behav. Res. Methods Instrum. Comput.* 31, 137–149. doi: 10.3758/BF03207704

Tajfel, H., and Wilkes, A. L. (1963). Classification and quantitative judgement. *Br. J. Psychol.* 54, 101–114. doi: 10.1111/j.2044-8295.1963.tb00865.x

Van den Bergh, O., Stegen, K., and Van de Woestijne, K. P. (1997). Learning to have psychosomatic complaints: conditioning of respiratory behavior and somatic complaints in psychosomatic patients. *Psychosom. Med.* 59, 13–23.

Van den Bergh, O., Winters, W., Devriese, S., Van Diest, I., Vos, G., and De Peuter, S. (2004). Accuracy of respiratory symptom perception in persons with high and low negative affectivity. *Psychol. Health* 19, 213–222. doi: 10.1080/08870440410001675627

von Leupoldt, A., Chan, P. Y. S., Bradley, M. M., Lang, P. J., and Davenport, P. W. (2011). The impact of anxiety on the neural processing of respiratory sensations. *Neuroimage* 55, 247–252. doi: 10.1016/j.neuroimage.2010.11.050

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of a brief measure of positive and negative affect. *J. Pers. Soc. Psychol.* 54, 1063–1070. doi: 10.1037/0022-3514.54.6.1063

Watson, D., and Pennebaker, J. (1989). Health complaints, stress, and distress: exploring the central role of negative affectivity. *Psychol. Rev.* 96, 234–254. doi: 10.1037/0033-295X.96.2.234

Wientjes, C. J., and Grossman, P. (1994). Over-reactivity of the psyche or the soma? Interindividual associations between psychosomatic symptoms, anxiety, heart rate, and end-tidal partial carbon dioxide pressure. *Psychosom. Med.* 56, 533–540. doi: 10.1097/00006842-199411000-00009

Yang, F., Wu, Q., and Li, S. (2014). Learning-induced uncertainty reduction in perceptual decisions is task-dependent. *Front. Hum. Neurosci.* 8:282. doi: 10.3389/fnhum.2014.00282