Supplementary Materials

Study 1

[S1] Oblimin Rotation pattern matrix

	Factor 1	Factor 2
Urinate (Item 5)	.739	
Hungry (Item 2)	.686	
Thirsty (Item 4)	.632	
Defecate (Item 6)	.622	
Heart (Item 1)	.543	
Pain (Item 17)	.505	
Breathing (Item 3)	.490	
Taste (Item 7)	.479	
Temperature (Item 11)	.455	
Muscles (Item 15)	.422	
Affective touch (Item 19)	.395	
Sexual arousal (Item 12)	.388	
Vomit (Item 8)	.354	
Itch (Item 21)		-.765
Tickle (Item 20)		-.659
Cough (Item 10)		-.639
Burp (Item 14)		-.633
Blood Sugar (Item 18)		-.608
Bruise (Item 16)		-.596
Sneeze (Item 9)		-.505
Wind (Item 13)		-.463

Study 4

[S2] Additional control measures

Body Mass Index

BMI was calculated using the following equation: mass(kg)/(height(m))^2.

Resting heart rate & heart rate variability

From the baseline measure of 120 seconds taken at the beginning of the experiment resting heart rate and heart rate variability were calculated from the photoplethysmograph capture and resulting RR intervals (time between two consecutive peaks in the QRS complex). Resting heart rate was calculated as the average heart rate over the 120 seconds. From the same interval, the root mean squared of successive differences was calculated and used as a measure of heart rate variability.

[S3] Additional control analyses

Additional control analyses for the simple correlation between the IAS and HCT

As the HCT data in Study 4 was collected remotely a number of additional data validation steps were taken.

First, trials for which the actual number of recorded heartbeats indicated an extremely low resting heartrate (<28bpm), possibly indicating a recording error caused by participant movement, were removed. This resulted in the removal of trials for three participants (two participants had one trial removed and one participant had two trials removed). This had little influence on the relationship between the IAS and the HCT when these individuals were excluded (standard scoring system with people that overestimated removed: $r(50) = .343$, $p=.013$; alternative scoring system: $r(54) = .324$, $p=.015$) or when the invalid trials were removed and the participant retained by averaging across the remaining valid intervals

(standard scoring system with people that overestimated removed: $r(52) = .298$, $p=.028$; alternative scoring system $r(57) = .274$, $p=.035$).

Second, in addition to the above, trials that indicated a resting heartrate that did not fall within the UK's National Health Service (NHS) guidelines of 40-120bpm were removed (NHS 'how do I check my pulse?', 2018). Along with extremely low (<28bpm) trials detailed above, this resulted in the removal of one trial for four participants. For one participant, the removal of this trial along with two extremely low trials (<28bpm), resulted in only one valid trial. As such, this participant was removed from further analyses. This additional control step had little influence on the relationship between the IAS and HCT when these individuals were excluded (standard scoring system with people that overestimated removed: $r(47) = .371$, $p=.009$; alternative scoring system: $r(51) = .351$, $p=.010$) or when the invalid trials were removed and the participants retained by averaging across the remaining valid intervals (standard scoring system with people that overestimated removed: $r(52) = .320$, $p=.018$; alternative scoring system $r(56) = .301$, $p=.022$).

Third, for each individual participant extreme trials (trials that deviated >25bpm +/- from the average resting heartrate of all valid trials between 40-120bpm) were removed. This resulted in the removal of one trial for five participants. Along with the above controls, this additional control step had little influence on the relationship between the IAS and HCT when these individuals were excluded (standard scoring system with people that overestimated removed: $r(43) = .379$, $p=.010$; alternative scoring system: $r(46) = .357$, $p=.013$) or when the invalid trials were removed and the participant retained by averaging across valid intervals (standard scoring system with people that overestimated removed: $r(52) = .305$, $p=.025$; alternative scoring system $r(56) = .288$, $p=.029$).

Additional regression analyses

To control for a number of confounds, entry method regressions were conducted predicting HCT accuracy from the IAS, controlling for age group (0 = 18-24, 1 = 25-34, 2 = 35-44, 3 = 45-54, 4 = 55-64, 5 = 65-74), gender (0 = Female, 1 = Male; no individuals identified as non-binary in this sample), TET performance, BMI, Resting HR, and HR variability. Using the standard scoring method for both the HCT and TET (with over estimators removed) the overall model was significant ($F(7, 45) = 2.745$, p=.018) and the IAS was a significant predictor of HCT performance (β = .355, t = 2.664, p=.011). Using the alternative scoring method for both HCT and TET, had little influence on the pattern of results obtained. The overall model was significant ($F(7, 51) = 2.882$, p = .013) and the IAS was a significant predictor of HCT performance (β = .371, t = 2.970, p = .005). The use of robust regressions conducted in Matlab using the default tuning function did not alter the relationship between the IAS and the HCT using either scoring method. Moreover, for both scoring schemes, the stringent data quality checks described above did not change the relationship between the IAS and the HCT for all simple linear regression analyses.

Additional control analyses for the simple correlation between the IAS and HCT confidence ratings

The additional data quality control steps were also performed for confidence ratings.

First, for all HCT accuracy trials where resting heart rate fell below <28bpm we also removed the confidence ratings from the trials that were excluded. When we removed these participants entirely, the relationship between confidence ratings and the IAS was unchanged in the total sample ($r(54) = -.033$, $p>.250$), as was the relationship between confidence and HCT accuracy (standard method: $r(50) = .532$, $p<.001$; alternative scoring system: $r(54) = .541$, $p<.001$). Likewise, when we averaged over valid trials to retain participants, there was

no change in the relationship between confidence and the IAS in the total sample ($r(57) = -.023$, $p>.250$) or the relationship between confidence and HCT accuracy (standard method: $r(52) = .461$, $p<.001$; alternative scoring system: $r(57) = .457$, $p<.001$).

Second, when all HCT accuracy and confidence trials where resting heartbeat was outside the normal range of 40-120bpm were removed, again there was no change in this relationship. When these participants were removed, the IAS was not correlated with confidence in the total sample ($r(51) = -.034$, $p>.250$) and confidence was correlated with HCT accuracy (standard method: $r(47) = .557$, $p<.001$; alternative scoring system: $r(51) = .558$, $p<.001$). When we averaged over valid trials to retain participants, the IAS was still uncorrelated with confidence in the total sample ($r(56) = -.001$, $p>.250$) and confidence was correlated with HCT accuracy (standard method: $r(52) = .449$, $p<.001$; alternative scoring system: $r(56) = .448$, $p<.001$).

Finally, when we removed outlying trials in addition to the above, again there was no change in the relationship. When these participants were removed, the IAS was not correlated with confidence in the total sample ($r(46) = -.014$, $p>.250$) and confidence was correlated with HCT accuracy (standard method: $r(43) = .552$, $p<.001$; alternative scoring system: $r(46) = .554$, $p<.001$). When we averaged over valid trials to retain participants, the IAS was still uncorrelated with confidence in the total sample ($r(56) = .005$, $p>.250$) and confidence was correlated with HCT accuracy (standard method: $r(52) = .459$, $p<.001$; alternative scoring system: $r(56) = .461$, $p<.001$).

Additional regression analyses

To control for a number of confounds, entry method regressions were conducted predicting HCT accuracy from confidence ratings, controlling for age group (0 = 18-24, 1 = 25-34, 2 = 35-44, 3 = 45-54, 4 = 55-64, 5 = 65-74), gender (0 = Female, 1 = Male; no

individuals identified as non-binary in this sample), Time estimation, BMI, Resting HR, and HR variability. Using the standard scoring method for both the HCT and TET (with over estimators removed) the overall model was significant ($F(7,45) = 3.985$, $p=.002$) and confidence ratings significantly predicted HCT performance ($\beta = .484$, $t = 3.760$, $p < .001$). Using the alternative scoring method for both the HCT and TET, the overall model was significant ($F(7, 51) = 3.117$, $p=.008$) and confidence ratings significantly predicted performance ($\beta = .412$, $t = 3.194$, $p=.002$). The use of robust regressions conducted in Matlab using the default tuning function did not alter the relationship between the IAS and the HCT using either scoring method. Moreover, the stringent data quality measures described above did not alter the relationship between the HCT and confidence ratings in any of the simple linear regression analyses.

Additional control analyses for the simple correlation between the IAS and HCT insight

The additional data quality control steps were also conducted for interoceptive insight.

First, for all HCT accuracy trials where resting heart rate fell below <28bpm we also removed the confidence ratings from the trials that were excluded. Therefore, these trials were also removed from the calculation of interoceptive insight. When we removed these participants entirely, using the standard scoring system interoceptive insight (high scores represent poor insight) was only correlated with confidence ratings ($r(50) = -.449$, $p<.001$), not the IAS ($r(50) = .051$, $p>.250$) or HCT accuracy ($r(50) = .088$, $p>.250$). Using the alternative scoring method, none of these factors were correlated with insight (all $p>.250$). Retaining participants by using average scores based on valid trials did not change the pattern of results reported above or in text. Using the standard scoring method confidence ratings were correlated with insight ($r(52) = -.402$, $p=.003$) but the IAS and HCT accuracy were

uncorrelated with insight (both p>.250). Using the alternative scoring method, none of these factors were correlated with insight (all p>.235).

Second for all HCT accuracy trials where resting heart rate fell outside of the range of 40-120bpm we also removed the confidence ratings from the trials that were excluded. Therefore, these trials were also removed from the calculation of interoceptive insight. When we removed these participants entirely, using the standard scoring system interoceptive insight (high scores represent poor insight) was only correlated with confidence ratings (r(47) = -.449, p=.001) not the IAS or HCT accuracy (both p>.250). Using the alternative scoring method, none of these factors were correlated with insight (all p>.250). Retaining participants by using average scores based on valid trials did not change the pattern of results reported above or in text. Using the standard scoring method confidence ratings were correlated with insight (r(52) = -.407, p=.002) but the IAS and HCT accuracy were uncorrelated with insight (both p>.250). Using the alternative scoring method, none of these factors were correlated with insight (all p>.250).

Finally, in addition to the above we also removed outlying trials. When we removed these participants entirely, using the standard scoring system interoceptive insight (high scores represent poor insight) was only correlated with confidence ratings (r(43) = -.443, p=.002) not the IAS or HCT accuracy (both p>.250). Using the alternative scoring method, none of these factors were correlated with insight (all p>.240). Retaining participants by using average scores based on valid trials did not change the pattern of results reported above or in text. Using the standard scoring method confidence ratings were correlated with insight (r(52) = -.393, p=.003) but the IAS and HCT accuracy were uncorrelated with insight (all p>.250). Using the alternative scoring method, none of these factors were correlated with insight (all p>.230).

Study 5

[S4] Additional control measures

Body Mass Index

BMI was calculated using the following equation: mass(kg)/(height(m))^2.

Systolic blood pressure

Blood pressure was taken using an electronic upper arm monitor (Omron M2) whilst participants were seated. High scores indicate higher systolic blood pressure.

Resting heart rate & heart rate variability

Average resting heart rate was taken as a measure of resting heart rate. This was estimated from the last 60 seconds of the longest duration. Where this was not available, a comparable interval from one of the other trials was used as a replacement. As a proxy of heart rate variability (HRV), the root mean square of successive differences was calculated from the second by second pulse rate given by the pulse oximeter. Higher scores indicate higher resting heart rate or increased heart rate variability.

Knowledge of average resting heart rate

After the heartbeat counting task participants were asked to estimate the average person's resting heart rate "how many times do you think the average person's heart beats in 60 seconds when they are at rest?". The absolute difference between the participant's estimate and average resting heart rate (reported in large studies of human physiology; 72.26; Agelink et al., 2001; Ramaekers, Ector, Aubert, Rubens, & Van de Werf, 1998) was taken as a measure of accuracy. This was favoured over asking participants to estimate their own heart rate to avoid effects of estimation on the HCT and vice versa. High scores on this variable indicate greater deviation between the participant's estimate and average resting heart rate, and therefore greater inaccuracy.

[S5] Results

To control for a number of confounds, entry method regressions were conducted predicting HCT accuracy from either the IAS, BPQ, TAS-20, ICQ or confidence ratings for the HCT, controlling for age (years), gender (0=female, 1 = male; no individuals identified as non-binary in this sample), Beliefs, Time estimation, Systolic Blood Pressure, BMI, Resting HR, and HR variability. Given the correlations between a number of these variables (e.g., the IAS, BPQ, TAS-20 and ICQ), they were not entered together, but instead entered into separate regression models including the same control variables. For the model predicting HCT scores from the TAS-20, depression and anxiety scores were also entered, given that alexithymia often co-occurs with these factors.

When TAS-20 was the predictor variable, only higher HRV (b =.490, t = 2.232, p=.036) predicted better HCT performance, whereas a trend was observed for TAS-20 to predict poor performance (b =-.518, t = -1.832, p=.080). All other predictors were non-significant (all p>.230). In the model including the BPQ as a predictor, no predictors (including BPQ) were significant (all p>.13). In the model including the ICQ as a predictor, no predictors were significant (all p>.05) but a trend was observed for higher interoceptive confusion scores to predict poorer HCT performance (b =-.428, t = -1.881, p=.072). When IAS was the predictor variable, a similar pattern to the TAS-20 was observed; higher HRV (b =.440, t = 2.190, p=.038) and the IAS (b =.470, t = 2.385, p=.025) predicted better performance, with all other factors non-significant (all ps>.32). In the final model predicting HCT performance from confidence ratings, only confidence ratings strongly predicted HCT performance (b =.850, t = 7.163, p<.001). The use of robust regression analyses (conducted in Matlab using the default tuning function) did not alter the pattern of significance reported above regarding the relationship between self-reported interoception and HCT performance.

Supplementary References

How do I check my pulse? (2018). Retrieved from https://www.nhs.uk/common-health-questions/accidents-first-aid-and-treatments/how-do-i-check-my-pulse/ (retrieved on: 25/09/2018)

Agelink, M. W., Malessa, R., Baumann, B., Majewski, T., Akila, F., Zeit, T., & Ziegler, D. (2001). Standardized tests of heart rate variability: normal ranges obtained from 309 healthy humans, and effects of age, gender, and heart rate. *Clinical Autonomic Research*, *11*(2), 99–108. doi:10.1007/BF02322053

Ramaekers, D., Ector, H., Aubert, A. E., Rubens, A., & Van de Werf, F. (1998). Heart rate variability and heart rate in healthy volunteers. Is the female autonomic nervous system cardioprotective? *European Heart Journal*, *19*(9), 1334–1341. doi:10.1053/euhj.1998.1084