

**Contribution of Time Estimation and Knowledge to Heartbeat Counting Task
Performance under Original and Adapted Instructions**

Olivier Desmedt^{a,b}, Olivier Corneille^a, Olivier Luminet^{a,b}, Jennifer Murphy^c, Geoffrey
Bird^{d,e} & Pierre Maurage^{a,b}

^a Psychological Science Research Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium

^b Fund for Scientific Research – Belgium (FRS-FNRS)

^c Department of Psychology, Royal Holloway, University of London

^d Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK

^e Department of Experimental Psychology, University of Oxford, UK

Corresponding author. E-mail address: olivier.desmedt@uclouvain.be (O. Desmedt).

To be published in Biological Psychology

This is the preprint version

Abstract

Interoceptive accuracy is frequently assessed using the Heartbeat Counting Task (HCT), requiring participants to count the number of times their heart beats. The HCT validity has been questioned, as participants may perform the task by estimating, rather than counting, their felt heartbeats. Participants could estimate the time or use their knowledge of their heart rate. Some research ruled out the contribution of time estimation in HCT performance. However, we believe these studies relied on a problematic analytic rationale. We revisited this question by relying on new analytic strategies, and examining the role of estimation in HCT performance, while varying task instructions. The findings support the role of time and knowledge-based estimations under original instructions. They also highlight the critical impact of instructions on HCT validity. Given the many limitations of the HCT, we urge researchers to test the robustness of published effects and to reconsider the interpretation of replicable results.

Keywords: Heartbeat Counting Task; Interoception; Interoceptive Accuracy; Time Estimation.

Introduction

Interoception can be broadly defined as the processing of internal bodily stimuli by the nervous system (Cameron, 2001; Khalsa et al., 2017). Interoceptive accuracy (IAcc; Garfinkel et al., 2015) is the term used to describe the objective capacity to detect these stimuli.

Individual differences in IAcc are thought to play a central role in a variety of everyday abilities (e.g., adaptive responses, feelings, drives, emotions), and to predict mental (Khalsa et al., 2017; Paulus & Stein, 2010; Pollatos et al., 2008; Schaefer et al., 2012) and physical (Herbert & Pollatos, 2014; Verdejo-Garcia et al., 2012) health.

In order to test adequately the relationship between IAcc and any outcome variable, valid measures are required. The validity of the most frequently used measure of IAcc, the Heartbeat Counting Task (HCT; Dale & Anderson, 1978; Schandry, 1981), however, has been increasingly questioned (for a recent discussion, see Corneille et al., 2020). In the HCT, participants are requested to count their heartbeats during various time intervals. Their reports are then compared to their objectively measured heartbeats. In the original description of the measure, participants were allowed to estimate their heartbeats (for a discussion see Brener & Ring, 2016; Schandry, 1981). For this task to be valid, however, performance should reflect the ability of participants to detect, rather than estimate, their heartbeats.

Contrary to this requirement, many studies have suggested that HCT performance reflects heart rate estimation rather than true detection (at least for some participants). Modifications of the instructions have been designed to limit the influence of estimation strategies. Specifically, while the original instructions ask participants to count or estimate their heartbeats (Schandry, 1981), adapted instructions ask them to only count their felt heartbeats (Desmedt et al., 2018; Ehlers et al., 1995; Murphy et al., 2018). Importantly,

previous research shows that results obtained using the HCT differ depending on which instructions are given (Desmedt et al., 2018; Ehlers et al., 1995).

In the present study, we investigated the contribution of time estimation strategies and knowledge about their heart rate (i.e., number of heartbeats per minute) to HCT performance, under both original and adapted instructions. In doing so, we also relied on more accurate analytic strategies. Below, we provide the theoretical and empirical background for our research question. Then, we report two studies that addressed this question. Finally, solutions to improve the HCT's validity are discussed.

Interoceptive accuracy as measured by the HCT

The HCT is the most frequently used measure of IAcc (Dale & Anderson, 1978; Schandry, 1981). In the HCT, participants are asked to count their heartbeats during different time intervals (e.g., 25s, 35s and 45s) without feeling their pulse using their hands. Smaller proportional absolute differences between the number of heartbeats reported by participants and objectively measured heartbeats result in higher HCT performance. Because of its ease of use, this task has been widely used during the last decades. Most of these studies (even the most recent ones, e.g., Abrams et al., 2018; Weineck et al., 2019), while acknowledging some limitations of the task, still treat HCT performance as a valid IAcc measure. However, the identification of core limitations of the task has cast doubts on HCT performance interpretation, and on how it relates to outcome variables.

The questionable validity of the HCT has been acknowledged for a long time (e.g., Flynn & Clemens, 1988; Jones, 1994; Katkin & Reed, 1988; Pennebaker, 1981; Weisz et al., 1988; Yates et al., 1985). One of its major limitations is that participants can reach high HCT

performance by estimating, rather than feeling, their heartbeats (Desmedt et al., 2018; Phillips et al., 1999; Ring et al., 2015; Ring & Brener, 1996; Windmann et al., 1999). Participants could perform this estimation by counting at a pace close to their estimated heart rate, or by combining their judgement of the elapsed time with their judgement of their heart rate.

If participants are trying to estimate their heart rate during the HCT, they are presumably relying on time estimation as they are blind to the time intervals¹. As a result, it could be inferred that HCT performance should be related to time estimation accuracy (Dunn et al., 2010). Following this rationale, many studies have tested the correlation between time estimation and HCT performance. While some studies found a significant correlation (up to $r=.40$), the vast majority documented low or non-significant correlations (between $r=-.001$ and $.40$; Ainley et al., 2014; Craske et al., 2001; Dunn et al., 2007, 2010; Knoll & Hodapp, 1992; Murphy et al., 2017; Shah, Catmur, et al., 2016; Shah et al., 2017; Shah, Hall, et al., 2016; Zoellner & Craske, 1999). The absence of a substantial correlation seemingly runs against the claim that participants use time estimation during the HCT, and this would support its validity (e.g., Ring & Brener, 2018; Weineck et al., 2019).

Of critical importance, however, the influence of time estimation on HCT performance may not be correctly estimated by merely correlating HCT performance with time estimation accuracy. This is because HCT performance is computed on the basis of actual heart rate, and heart rates vary across individuals. To illustrate, at an individual level, given optimal time estimation (i.e., 60sec/min), HCT performance for participants counting seconds with a heart rate of 60bpm would be perfect. Conversely, if a participant's heart rate was 80 bpm, and they

¹ It should be noted that some authors use a time estimation task to control for the effect of other variables on HCT performance, such as attention, fatigue, motivation (Murphy et al., 2018).

reported the number of seconds elapsed on a trial with perfect accuracy, they would be considered to have low HCT accuracy as their report of the number of heartbeats (based on their estimate of the time elapsed) would not match the actual number of heartbeats. At the group level, time estimation accuracy would predict HCT performance only if heart rate is constant across individuals, which is an unreasonable assumption.

Instead of the correlation between time estimation accuracy and HCT performance, we believe the correct test consists of measuring the correlation between *the raw number of heartbeats* reported in the HCT (independent of actual heart rate) and the number of seconds estimated in the time estimation task, for identical time intervals. If participants rely on an identical strategy in the time estimation and HCT (i.e., estimate time in both tasks), then a significant correlation should be observed between reported heartbeats and reported seconds. As well as a time estimation strategy, participants could use their knowledge about their heart rate to adapt their counting rate. These may have been acquired through listening to another person's heart rate (e.g., in a movie) and/or perceiving their own heart rate (e.g., during sport or via heart rate monitoring devices). If participants indeed use their knowledge about their heart rate, a positive correlation should be observed between the raw number of heartbeats reported in the HCT and the reported knowledge about heart rate in beats per minute (bpm). Additionally, if they combine time estimation and knowledge about heart rate (i.e., by adjusting their second counting on the basis of their knowledge), a positive correlation between reported heartbeats in the HCT and the product between reported seconds and reported knowledge about heart rate [i.e. number of heartbeats in a minute (according to knowledge) X (number of estimated second in the time interval / 60)] should be observed.

The aim of the present study was therefore to test the contribution of time estimation and knowledge about heart rate to HCT performance under both original and adapted task instructions, using these new analytic strategies. Analyses were conducted on two separate data sets from the United Kingdom (Murphy et al., 2018) and Belgium (Desmedt et al., 2018), where both studies used adapted task instructions. In the UK study, participants were explicitly instructed not to count seconds or to guess. In the Belgian study, participants performed the HCT under original and adapted task instructions. The latter instructions made it clear to the participants that they should avoid estimating their heart rate and should exclusively report their genuinely felt heartbeats. In line with our rationale, the correlation between reported heartbeats and estimated seconds, as well as estimated heart rate (i.e., their knowledge) and their product should be reduced under the adapted task instructions (i.e., for the HCT in Study 1 and adapted HCT in Study 2), as these instructions (contrary to the original instructions) prevent reliance on heart rate estimation strategies.

Study 1

Method

Participants and Procedure

This study concerns a subset of a larger combined dataset from Murphy et al. (2018). This subset includes 69% of the participants included in the larger dataset for whom the raw data for both the HCT and the time estimation task were available. In this study, 198 volunteers (139 females, $M_{age} = 44.9$, $SD_{age} = 22.39$) were recruited through local advertisements and recruitment databases, and compensated for their participation. Ethical approval was granted by the local ethics committee. All participants gave informed consent

and were fully debriefed upon task completion. Participants performed the HCT and a time estimation task in a counterbalanced order.

Materials

HCT. Participants' heart rate was measured by a pulse oximeter. Participants were asked (full instructions in Appendix A) to silently count their heartbeats over different time intervals.

They were instructed to avoid counting seconds or guessing. Instructions were thus adapted as compared to the original ones by Schandry (1981, see Appendix B). For the sake of clarity, the gist of the instructions used in this first adaptation is: "Count your heartbeats without physically monitoring it using your hands. It is very important that you do not count seconds or guess". Four durations were examined, half of participants having shorter (35s, 25s, 45s, 100s) and the other half longer (28s, 38s, 48s, 103s) durations. Duration order was counterbalanced across participants.

Time estimation task. Participants had to count seconds during various time intervals (25s, 35s, 45s, 100s or 28s, 38s, 48s, 103s). Duration order was counterbalanced across participants. The durations (e.g., 25s vs. 28s) were counterbalanced across HCT and time estimation tasks. The instructions were as follows: "In this task you will be asked to count seconds from when the experimenter indicates you should begin until you hear a beep telling you to stop counting at which point you should tell the experimenter how many seconds you counted. Do you have any questions about these instructions?"

Results

HCT performance was computed with the following formula: $1/4 \sum (1 - (|\text{actual heartbeats} - \text{reported heartbeats}|) / \text{actual heartbeats})$, with higher scores indicating higher accuracy. Participants with extreme heart rates and extreme heart rate differences between time intervals (i.e., heart rate and mean difference of heart rate between intervals of more than

three times the interquartile range) were excluded from the analyses as these values may indicate reliability issues with the heart monitor devices. Statistical assumptions were tested before each analysis (i.e., Pearson correlations and paired sample t -tests). As univariate outliers (values more than three times the interquartile range) and non-normal distributions were detected, both parametric (Pearson, r) and non-parametric (Spearman, r_s) analyses were performed. It should, however, be noted that, with large enough sample sizes (> 40) as used here, parametric procedures can be used even when the data are not normally distributed (Elliott & Woodward, 2007). Only the parametric analyses will therefore be presented. Spearman correlations can be found in the Supplementary Materials. The results do not substantially change according to analyses type. When reporting correlation coefficients, we have used the interpretation thresholds proposed by Cohen (1992), i.e. small ($r = .1$), medium ($r = .3$) and large ($r = .5$).

All data are accessible at the following address:

https://osf.io/7b39n/?view_only=522315d79a9944bf97623825cbbcc4b6

Data exclusion and missing data: Four participants were removed for extreme (i.e., three times the interquartile range) mean differences of heart rate between time intervals. Actual heart rate during a time interval was absent for one participant. This participant was thus excluded from analyses including heart rate. The final sample size comprises 194 participants.

Reliability analyses: Internal consistency across time intervals is high for actual heart rate ($\alpha = .89$), HCT performance ($\alpha = .95$) and estimated seconds ($\alpha = .87$)

Descriptive statistics: In Table 1, we provide mean data of actual heartbeats, reported heartbeats and reported seconds for each interval and across intervals.

	Interval	Interval	Interval	Interval	Average
	25/28	35/38	45/48	100/103	beats/min

Actual HB	31	42.84	54.55	119.87	70.39
Reported HB	15.81	20.16	26.12	53.70	33.51
	Interval 25/28	Interval 35/38	Interval 45/48	Interval 100/103	Average sec/min
Reported sec	22.01	29.97	38.52	79.23	49.03

Table 1. Mean data of actual heartbeats, reported heartbeats and reported seconds for each interval and across intervals. Legend: HB = heartbeats; sec = seconds.

Involvement of time estimation in HCT performance

A small-to-medium positive correlation was found between the number of reported seconds and the number of reported heartbeats per minute ($r = .19, p = .009$).

For exploratory purpose, correlations were also computed for each time interval of the HCT and time estimation task. These correlations ranged from $r = .13$ ($p = .08$; interval 35/38) to $r = .25$ ($p = .001$; interval 45/48). The correlations observed in the 35/38 vs. 45/48 time intervals were significantly different ($z = -2.24, p = .02$) from each other, but no other pair of correlations was significant ($z < -1.9, p > .07$).

Study 2

Method

Data were from Desmedt et al. (2018). In this study, 123 healthy students (76 females, $M_{age} = 22.3, SD_{age} = 3.13$) were recruited through advertisements on mailing lists and compensated for their participation. Ethical approval was granted by the local ethics committee. All participants gave informed consent and were fully debriefed upon task completion. Participants first completed the International Physical Activity Questionnaire – short form (IPAQ-SF; Craig et al., 2003). Then, they performed the HCT under original task

instructions, immediately followed by questions about their anxiety and the strategies they used during the task. Then, they completed the HCT under adapted task instructions², followed by the time estimation task. The HCT was always completed under original task instruction first, as adapted instructions were more restrictive. Finally, participants were asked to report their usual heart rate at rest.

Materials

HCT under original and adapted task instructions. Participants' heart rate was measured by the Polar Watch RS800CX heart monitor, which has been shown to be valid (strong correlation with ECG methods) and reliable (moderate test-retest stability; Gamelin et al., 2006; Giles et al., 2016; Kingsley et al., 2005; Williams et al., 2017). Participants had to put their wrists on the electrodes of a belt, that was dampened with water. If participants moved/raised a wrist, the exercise was repeated, as the connection might have been broken. The signal was transferred from the belt to the watch in a wireless mode. The software Polar ProTrainer5 was used to examine the extracted actual heartbeats. As this procedure and the Polar Watch have recently been criticized for their potential lack of reliability (Ainley et al., 2019), participants with extreme heart rates and extreme heart rate differences between time intervals were excluded from the analyses (see Results). First, the HCT with original task instructions was administered, participants were instructed to report the counted number of heartbeats without feeling their pulses (for instructions, see Appendix B; Schandry, 1981). Next, participants performed the HCT under the adapted task instructions designed by Desmedt et al. (2018), where they were asked to count their felt heartbeats only (for instructions, see Appendix C). Three different time intervals (45s, 25s, 35s), each separated

² Following this task, questionnaires were administered to assess anxiety during the task, depression, anxiety, stress, alexithymia and awareness about the hypotheses. Results related to these variables are not reported here as they are irrelevant to the aims of the present research.

by a 20s break, were used. The reverse order in time intervals was used under adapted task instructions to promote independent judgments on the two tasks. An acoustic cue signaled the beginning and end of each time interval. For the sake of clarity, we summarize here the gist of the instructions used in the three tasks (one from Study 1 and two from Study 2): (1) Original (Schandry, 1981, Study 2): “Count your heartbeats without feeling (touching) your wrists or neck pulsations”; (2) first adaptation (Murphy et al., 2018, Study 1): “Count your heartbeats without physically monitoring it using your hands. It is very important that you do not count seconds or guess”; (3) second adaptation (Desmedt et al., 2018, Study 2): “It is very important that you only count the heartbeats you really feel, without trying to guess your heart rate”.

Time estimation task. Participants were asked to count the seconds for three time intervals (19s, 37s, 49s). The time intervals were different from those used for the HCT, to ensure independent judgments between the two tasks. The instructions given to participants were as follows: “I will ask you to estimate the time that passes during several intervals. You will hear a first signal telling you to start counting the seconds. When you hear a second signal, I ask you to stop counting and tell me how many seconds you counted.”

Participants' knowledge about their usual heart rate at rest was assessed by the following question: “What is your usual heart rate at rest (the number of heartbeats per minute)?”. If participants rely on their general knowledge to estimate their heartbeats during the HCT under original task instructions, then this knowledge should be correlated with the number of heartbeats reported in the HCT (given perfect time estimation). On the contrary, when participants are instructed to only count their felt heartbeats (under adapted instructions), their knowledge should hold no relation to their reported heartbeats during the HCT.

Results

HCT performance was computed with the following formula: $1/3 \sum (1 - (|\text{actual heartbeats} - \text{reported heartbeats}|) / \text{actual heartbeats})$, with higher scores indicating higher accuracy. The same procedure (i.e., exclusion criteria, statistical assumptions) as in Study 1 was applied to Study 2. Spearman correlations can be found in the Supplementary Materials. The results do not change according to analyses type. When reporting correlation coefficients, we have used the interpretation thresholds proposed by Cohen (1992), i.e. small ($r = .1$), medium ($r = .3$) and large ($r = .5$).

All data are accessible at the following address:

https://osf.io/7b39n/?view_only=522315d79a9944bf97623825cbbcc4b6

Data exclusion and missing data: 2.4% of the data ($N=3$) were removed from the analyses due to problems in applying the instructions and extreme mean difference of heart rate between time intervals (i.e., three times the interquartile range). The final sample size comprises 120 participants. Six participants could not report any knowledge about their heart rate and were thus excluded from analyses including this variable.

Reliability analyses: Internal consistency across time intervals is high for actual heart rate under original instructions ($\alpha = .96$) as well as adapted ones ($\alpha = .96$), for HCT performance under original instructions ($\alpha = .91$) as well as adapted ones ($\alpha = .93$) and for estimated seconds ($\alpha = .96$).

Descriptive statistics: In Table 2, we provide mean data of actual heartbeats, reported heartbeats and reported seconds for each interval and across intervals.

	Interval 25	Interval 35	Interval 45	Average beats/min
Actual HB – Original HCT	32.13	46.03	57.77	77.68

Reported HB – Original HCT	20.23	28.60	33.63	47.47
Actual HB – Adapted HCT	31.21	44.11	56.50	75.28
Reported HB – Adapted HCT	9.75	12.94	15.10	21.91
	Interval 19	Interval 37	Interval 49	Average sec per minute
Reported Sec	15.99	28.99	37.63	47.86

Table 2. Mean data of actual heartbeats, reported heartbeats and reported seconds for each interval and across intervals. Legend: HB = heartbeats; sec = seconds.

Involvement of time estimation in HCT performance under original and adapted task instructions

A medium-to-large positive correlation was found between the number of reported seconds and the number of reported heartbeats per minute ($r = .41, p < .001$)³ in the HCT under original instructions, suggesting that some participants may try to guess their heart rate based on time estimation. However, no significant correlation was found between reported heartbeats and reported seconds per minute under the adapted task instructions used in Study 2 ($r = .08, p = .39$)⁴. The size of this correlation under adapted instructions was significantly smaller than the correlation observed under the original task instructions in Study 2 ($z = 3.59, p < .001$; 1-tailed). When comparing the correlation between reported heartbeats and reported seconds per minute in the adapted instructions condition of Study 1 with the original instructions of Study 2, a significant difference was also observed between the original

³ This correlation did not significantly differ depending on the time interval ($z < 0.46, p > 0.64$).

⁴ This correlation did not significantly differ depending on the time interval ($z < 0.95, p > 0.35$).

instructions (Study 2; $r = .41$) and adapted instructions (Study 1; $r = .19$), $z = 2.07$, $p = .02$; one-tailed). Likewise, the size of this correlation did not significantly differ between the two sets of adapted instructions in Study 1 ($r = .19$) and Study 2 ($r = .08$), $z = -0.955$, $p = 0.34$; one-tailed. This is consistent with the view that some participants have difficulties perceiving their heartbeats in the HCT under original instructions conditions, and might therefore rely on a time-based strategy for estimating their heart rate. The adapted instructions, removing the need to rely on an estimation strategy among participants who did not feel their heart beats, largely reduced HCT performance (i.e., original: $M = 0.61$, and adapted: $M = 0.30$; Desmedt et al., 2018).

Involvement of knowledge about one's own heart rate in HCT performance under original and adapted task instructions

In Study 2, participants were also asked to report what they believe is their usual heart rate at rest. This knowledge (i.e., the estimation that each participant has regarding his/her heart rate at basal level, independently of the actual heart rate) was positively correlated with the number of reported heartbeats under original ($r = .27$, $p = .004$)⁵ but not adapted ($r = .02$, $p = .85$)⁶ task instructions. These correlations were significantly different ($z = 2.629$, $p = .004$; one-tailed) from each other. This suggests that participants may rely on their beliefs for estimating their heartbeats in the HCT under original task instructions, but not adapted ones.

The combined use of time estimation and knowledge about heart rate

Participants could also combine time estimation and knowledge about their heart rate to perform the task. For example, if a participant counts 30 seconds and knows their heart rate

⁵ Similar results are found when the number of reported heartbeats is predicted by the number of reported seconds ($\beta = .42$) and knowledge about heart rate ($\beta = .27$) in a multiple linear regression.

⁶ Similar results are found when the number of reported heartbeats is predicted by the number of reported seconds ($\beta = .075$) and knowledge about heart rate ($\beta = .02$) in a multiple linear regression.

is 80 bpm, they will report 40 heartbeats. Consistently, a significant positive correlation was found between reported heartbeats in the HCT under original instructions and the product between reported seconds and reported knowledge about heart rate ($r = .45, p < .001$), which was not found under adapted task instructions ($r = .06, p = .52$). The size of these correlations was significantly different ($z = 4.17, p < .001$; one-tailed). However, when controlling for the main effects of the number of reported seconds and knowledge about heart rate, the interaction/product was not a significant predictor of original HCT performance anymore ($\beta = .02, p = .84$). This result does not support the hypothesis that some participants combine time-based and knowledge-based estimation strategies to perform the HCT under original instructions. Rather, time-based or knowledge-based estimation strategies appear to be used separately.

Exploratory analyses

In addition to the use of adapted instructions, the contribution of time estimation and knowledge about heart rate in HCT performance should be assessed (as these may differ across populations) and their influence should be statistically controlled for. This can be done in a multiple regression analysis (or an ANCOVA) where HCT performance is the dependent variable. As independent variables, the following variables should be included: (1) the absolute difference between actual heartbeats and reported seconds, in a minute, and (2) the absolute difference between actual heartbeats and estimated heart rate (participants' beliefs), in a minute. Moreover, actual heart rate should also be controlled for. This leads to the following formula:

$$\text{HCT performance} = (1 - (|\text{actual heart rate} - \text{reported seconds}| / \text{actual heart rate})) + (1 - (|\text{actual heart rate} - \text{estimated heart rate}| / \text{actual heart rate})) + \text{actual heart rate}.$$

Applied to the dataset from Study 2, the following results were found: Under original instructions, HCT performance is significantly predicted by time estimation ($\beta=.39, p<.001$), marginally predicted by knowledge about heart rate ($\beta=.182, p=.054$), and not significantly predicted by resting heart rate ($\beta=-.024, p=.80$). Under adapted instructions, resting heart rate ($\beta=-.054, p=.572$), time estimation ($\beta=.163, p=.09$) and knowledge about heart rate ($\beta=.166, p=.08$) are not significant predictors of HCT performance anymore. These results suggest that participants use time or knowledge-based estimation strategies under original instructions, but adapted instructions reduce their influence on HCT performance.

Discussion

The past decade has been marked by a significant resurgence of interest in interoception. The most frequently used measure (HCT) of IAcc, however, suffers from considerable limitations impacting its validity (see Corneille et al., 2020). The use of an invalid task is likely to result in replicability issues (as past results could represent false positives) or in misleading conclusions, both of which have harmful consequences for theory and practice. The difficulty to replicate results coming from studies that used the HCT has already been observed in various contexts (Zamariola et al., 2018, 2019). Moreover, even if results are replicable, their interpretation may be misleading: an observed relation between HCT performance and any variable could be due to non-interoceptive factors (e.g., knowledge about heart rate or time estimation, or even attention and motivation).

The aim of the present study was to test the involvement of two confounding factors (i.e., time-based and knowledge-based estimation strategy) in HCT performance, under original and adapted task instructions. We argued that past studies exploring the influence of time estimation (by measuring the correlation between HCT performance and time estimation

accuracy) did not use adequate analyses. We proposed a new analytic rationale (i.e., computing the correlation between the number of reported heartbeats and the number of reported seconds for identical time intervals). Additionally, we examined the involvement of knowledge about heart rate, also with respect to its interaction with time estimation.

Regarding the involvement of time estimation, we found a positive correlation between reported heartbeats and reported seconds. While no causal interpretation can emerge from correlational analyses, this result is consistent with the view that at least some participants based their heartbeat counting response on time duration estimates under the original task instructions. An alternative explanation is that participants counted their heartbeats when estimating time (e.g., Wittmann, 2009). Our results run against this account. This is because the correlation between reported heartbeats and reported seconds decreased when (the adapted) instructions specified that seconds should not be counted during the HCT.

The highest correlation ($r = .41$) was found under the original task instructions used in Study 2, where participants were simply asked to count their heartbeats. Then, a weaker but still significant positive correlation ($r = .19$) was found in Study 1, where participants were asked not to count seconds or guess (i.e., first adaptation). This suggests that some participants might, even under the first adaptation of the instructions, guess their heart rate based on time estimation. Finally, no significant correlation ($r = .08$) was found under the adapted task instruction condition of Study 2 (i.e., second adaptation), where we made it even clearer that only the felt heartbeats should be counted.

Similar results were found, in Study 2, regarding the involvement of knowledge about heart rate in HCT. A positive correlation was found between reported heartbeats and estimated heart rate, as well as its product with reported seconds, under original instructions (r

= .27, $p = .004$; $r = .45$, $p < .001$), but not under adapted ones ($r = .02$, $p = .85$; $r = .06$, $p = .52$). As a set, these findings support the hypothesis that adapted instructions reduce participants' reliance on time and knowledge-based estimation strategies when performing the HCT. A critical asset of the present research is that our results were replicated across two studies, despite their strong methodological differences: The first study had a larger sample size and a higher mean age. The measure of actual heartbeats also differed, with a pulse oximeter in Study 1 and a Polar Watch in Study 2. Regarding HCT administration, instructions (see Appendices) and time intervals also differed. Moreover, contrary to Study 2, Study 1 counterbalanced HCT and time estimation task.

Hence, the current findings support the proposal that some participants may use time estimation or knowledge about heart rate when performing the HCT under original task instructions. Original instructions may indeed create an experimental demand prompting participants to use estimation strategies (Orne, 1962). A consequence of this is that participants who perceive that time is running fast are likely to show better HCT performance. This factor may be related to participants' mental or physical condition. For instance, individuals with depressive symptoms underestimate time duration (Bschor et al., 2004; Grinker et al., 1973; Tysk, 1984). The negative correlation between depression and HCT performance (Pollatos et al., 2009) may thus be explained by time underestimation. Future studies should determine the role of these variables (i.e., time estimation and knowledge about heart rate) in relationships observed between HCT performance and individual differences.

Future studies should overcome several limitations of the current research. First, a more precise evaluation of participant's heart rate knowledge is needed. Our question regarding knowledge about one's heart rate (i.e. "What is your usual heart rate at rest?") was

rather imprecise as it did not specify the meaning of “at rest”, which can vary as a function of different conditions (e.g., body posture, fatigue level, recent consumption of stimulants like caffeine or nicotine may influence this judgement). This might have reduced the precision of the measure. Second, despite our results clearly indicating an influence of heart rate knowledge (and time estimation) on HCT performance, our design did not allow us to determine the proportion of participants who actually relied on these factors to perform the task. We cannot exclude the possibility that some participants estimate their heart rate without using time estimation or knowledge about heart rate. Future research should further investigate this question (e.g., by explicitly questioning each participant, during post-experimental interviews, on the strategies used). Third, knowledge about heart rate could be inferred by the perception of actual heart rate, which would explain its correlation with reported heartbeats during the HCT. However, the observed decrease in the size of the correlation between knowledge about heart rate and reported heartbeats observed after instructions’ modifications prompting participants to only count their felt heartbeats suggests that this correlation is due to the use of estimation strategies during the HCT. On a more global note, the exact influence of the adapted instructions on participants’ strategy should be specified: our correlational results suggest that adapted instructions reduced the use of non-interoceptive strategies (as shown by the strong reduction of the correlations between HCT performance and time estimation/heart rate knowledge). However, they might also have led some participants to opt for a more conservative response bias (i.e., to increase their heartbeat detection threshold, independently of non-interoceptive strategies).

Beyond these limitations, the present research points to a couple of recommendations: (1) *the modification of HCT instructions*. The results show that adapted instructions reduce the contribution of estimation-based strategies in HCT performance. Therefore, future studies

should use adapted instructions. However, we suggest caution about this solution for two reasons. First, more studies testing the validity of HCT under adapted instructions are needed. Indeed, other confounding variables (e.g., decision threshold; Zamariola et al., 2018) may still influence HCT performance. Second, the extent to which adapted instructions might control for estimation strategies may depend not only on the type of instructions given, but also on the population (e.g., children, clinical populations).

(2) *the statistical control of confound variables*. In addition to the use of adapted instructions, we proposed to statistically control time estimation and knowledge about heart rate (as these may differ across populations). This can be done by applying the following model:

$$\text{HCT performance} = (1 - (|\text{actual heart rate} - \text{reported seconds}| / \text{actual heart rate})) + (1 - (|\text{actual heart rate} - \text{estimated heart rate}| / \text{actual heart rate})) + \text{actual heart rate}.$$

Applied to the dataset from Study 2, the results suggested that participants use time or knowledge-based estimation strategies under original instructions, but adapted instructions reduce their influence on HCT performance. Nevertheless, the statistical control of confound variables comes with multiple costs. Partialing out independent variables involves removing the variance from one variable that is shared with others. It allows for the investigation of the relation between an independent variable and a dependent variable, after taking into consideration the contribution of other independent variables. The most important limit of this strategy is multicollinearity, which gives rise to three problems: the compromised interpretation of the findings, the low stability of the findings, and the low computation accuracy of partial coefficients (Cohen et al., 1983). While the last two problems can be statistically solved (e.g., Morrow-Howell, 1994), the first one remains a challenge. Indeed, when the variance of one independent variable shared with the control variables is removed, it becomes difficult to know what construct this independent variable represents (Lynam et al.,

2006). Researchers should thus be cautious when controlling for confound variables, and keep in mind that the higher the number of control variables, the more compromised the interpretation of an observed relation. As a further important issue, if there is no consistency between studies in the control variables used, one may be tempted to add/remove covariates until significance is reached, therefore increasing false positives. Therefore, we recommend the preregistration of theoretically relevant covariates and replication of findings including non-preregistered covariates considered in exploratory analyses. Finally, the more covariates there are, the higher the sample size should be, which increases the risk of underpowered studies. As a whole, whilst statistical control may be the most practical approach in certain populations (e.g., children), one should consider its associated limitations. Adding covariates may (1) increase the risk of false positives if wide variations exist in the nature of the included covariates, (2) increase the risk of adulterating the variable of interest (e.g., Miller & Chapman, 2001) and, (3) increase the need of larger sample sizes to reach adequate statistical power. Finally, even if the aforementioned solutions are undertaken, the validity of the adapted or controlled HCT still needs to be established. Future research should thus work on additional modifications meant to improve the measurement of cardiac interoception abilities.

In sum, we discourage researchers from using the original task instruction when using the HCT. Furthermore, we call for a replication of past HCT findings using adapted instructions and control procedures.

Appendix A: First adaptation of HCT instructions.

In this task you will be asked to silently count your heartbeat without physically monitoring it using your hands. I will ask you to sit with your feet flat on the floor and both hands on the table with your eyes closed. You will be asked to count from when the experimenter indicates you should begin until you hear a beep telling you to stop counting at which point you should tell the experimenter how many heartbeats you counted. It is very important that you do not count seconds or guess, so if you feel no heartbeats you should give 0 as your answer. Do you have any questions about these instructions?

Appendix B: Original HCT instructions.

I'm inviting you to comfortably sit in your chair. Once you feel comfortable, try not to move during the task. In a moment, when you hear an acoustic signal, I will ask you to start silently counting your heartbeats without feeling (touching) your wrists or neck pulsations. When you hear a second signal, I will ask you to stop counting and tell me the exact number of heartbeats you counted. We will repeat this exercise several times.

Appendix C: Second adaptation of HCT instructions.

I'm inviting you to comfortably sit in your chair. Once you feel comfortable, try not to move during the task. The instructions for this task are simple but it is very important to follow them carefully. In a moment, I will ask you to try to sense your heartbeat without actively feeling (touching) your wrists or neck pulsations. You may feel them, you may not feel them at all or you may occasionally feel some. When you hear a signal, you will have to start silently counting the heartbeat you feel. When you hear a second signal, I will ask you to tell me the exact number of heartbeats you felt. We will repeat this exercise several times. For this task, it is very important that you only count the heartbeats you really feel, without trying to

guess your heart rate. I really insist that you only count what you feel. It could mean reporting no beat at all, some beats or all beats that actually happened.

Acknowledgments

Olivier Desmedt (PhD student), Olivier Luminet (Research Director) and Pierre Maurage (Senior Research Associate) are funded by the Fund for Scientific Research – Belgium (FRS-FNRS). Jennifer Murphy was supported by a doctoral studentship from the Economic and Social Research Council [1599941; ES/J500057/1]. GB was supported by the Baily Thomas Trust.

References

- Abrams, K., Cieslowski, K., Johnson, S., Krimmel, S., Bierlein-De La Rosa, G., Barton, K., & Silverman, P. (2018). The effects of alcohol on heartbeat perception: Implications for anxiety. *Addictive Behaviors*, 79, 151–158.
- Ainley, V., Brass, M., & Tsakiris, M. (2014). Heartfelt imitation: High interoceptive awareness is linked to greater automatic imitation. *Neuropsychologia*, 60, 21–28.
- Ainley, V., Tsakiris, M., Pollatos, O., & Herbert, B. M. (2019). *Comment on “Zamariola et al., (2018), Interoceptive Accuracy Scores are Problematic: Evidence from Simple Bivariate Correlations” - The Empirical Data Base, the Conceptual Reasoning and the Analysis behind this Statement are Misconceived and do not Support the Authors’ Conclusions* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/wdehr>
- Brener, J., & Ring, C. (2016). Towards a psychophysics of interoceptive processes: The measurement of heartbeat detection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1708). <https://doi.org/10.1098/rstb.2016.0015>
- Bschor, T., Ising, M., Bauer, M., Lewitzka, U., Skerstupeit, M., Müller-Oerlinghausen, B., & Baethge, C. (2004). Time experience and time judgment in major depression, mania and healthy subjects. A controlled study of 93 subjects. *Acta Psychiatrica Scandinavica*, 109(3), 222–229.
- Cameron, O. G. (2001). Interoception: The inside story—a model for psychosomatic processes. *Psychosomatic Medicine*, 63(5), 697–710.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (1983). Applied multiple regression. *Correlation Analysis for the Behavioral Sciences*, 2.

- Corneille, O., Desmedt, O., Zamariola, G., Luminet, O., & Maurage, P. (2020). A Heartfelt Response to Zimprich et al.(2019), and Ainley et al.(2019)’s Commentaries: Acknowledging issues with the HCT would Benefit Interoception Research. *Biological Psychology*, 107869.
- Craig, C. L., Marshall, A. L., Sjorstrom, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U. L. F., Yngve, A., & Sallis, J. F. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise*, 35(8), 1381–1395.
- Craske, M. G., Lang, A. J., Tsao, J. C., Mystkowski, J. L., & Rowe, M. K. (2001). Reactivity to interoceptive cues in nocturnal panic. *Journal of Behavior Therapy and Experimental Psychiatry*, 32(3), 173–190.
- Dale, A., & Anderson, D. (1978). Information variables in voluntary control and classical conditioning of heart rate: Field dependence and heart-rate perception. *Perceptual and Motor Skills*, 47(1), 79–85.
- Desmedt, O., Luminet, O., & Corneille, O. (2018). The heartbeat counting task largely involves non-interoceptive processes: Evidence from both the original and an adapted counting task. *Biological Psychology*, 138, 185–188.
- Dunn, B. D., Dalgleish, T., Ogilvie, A. D., & Lawrence, A. D. (2007). Heartbeat perception in depression. *Behaviour Research and Therapy*, 45(8), 1921–1930.
- Dunn, B. D., Stefanovitch, I., Evans, D., Oliver, C., Hawkins, A., & Dalgleish, T. (2010). Can you feel the beat? Interoceptive awareness is an interactive function of anxiety- and depression-specific symptom dimensions. *Behaviour Research and Therapy*, 48(11), 1133–1138.
<https://doi.org/10.1016/j.brat.2010.07.006>

- Ehlers, A., Breuer, P., Dohn, D., & Fiegenbaum, W. (1995). Heartbeat perception and panic disorder: Possible explanations for discrepant findings. *Behaviour Research and Therapy*, 33(1), 69–76.
- Elliott, A. C., & Woodward, W. A. (2007). *Statistical analysis quick reference guidebook: With SPSS examples*. Sage.
- Flynn, D. M., & Clemens, W. J. (1988). On the validity of heartbeat tracking tasks. *Psychophysiology*, 25(1), 92–96.
- Gamelin, F. X., Berthoin, S., & Bosquet, L. (2006). Validity of the polar S810 heart rate monitor to measure RR intervals at rest. *Medicine & Science in Sports & Exercise*, 38(5), 887–893.
- Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K., & Critchley, H. D. (2015). Knowing your own heart: Distinguishing interoceptive accuracy from interoceptive awareness. *Biological Psychology*, 104, 65–74.
- Giles, D., Draper, N., & Neil, W. (2016). Validity of the Polar V800 heart rate monitor to measure RR intervals at rest. *European Journal of Applied Physiology*, 116(3), 563–571.
- Grinker, J., Glucksman, M. L., Hirsch, J., & Viseltear, G. (1973). Time perception as a function of weight reduction: A differentiation based on age at onset of obesity. *Psychosomatic Medicine*.
- Herbert, B. M., & Pollatos, O. (2014). Attenuated interoceptive sensitivity in overweight and obese individuals. *Eating Behaviors*, 15(3), 445–448.
- Jones, G. E. (1994). *Perception of visceral sensations: A review of recent findings, methodologies, and future directions*.
- Katkin, E. S., & Reed, S. D. (1988). Cardiovascular asymmetries and cardiac perception. *International Journal of Neuroscience*, 39(1–2), 45–52.
- Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H. D., Davenport, P. W., Feinstein, J. S., Feusner, J. D., Garfinkel, S. N., Lane, R. D., & Mehling, W. E. (2017). Interoception and

- mental health: A roadmap. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 3, 501–503.
- Kingsley, M., Lewis, M. J., & Marson, R. E. (2005). Comparison of polar 810 s and an ambulatory ECG system for RR interval measurement during progressive exercise. *International Journal of Sports Medicine*, 26(01), 39–44.
- Knoll, J. F., & Hodapp, V. (1992). A comparison between two methods for assessing heartbeat perception. *Psychophysiology*, 29(2), 218–222.
- Lynam, D. R., Hoyle, R. H., & Newman, J. P. (2006). The perils of partialling: Cautionary tales from aggression and psychopathy. *Assessment*, 13(3), 328–341.
- Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, 110(1), 40.
- Morrow-Howell, N. (1994). The M word: Multicollinearity in multiple regression. *Social Work Research*.
- Murphy, J., Brewer, R., Hobson, H., Catmur, C., & Bird, G. (2018). Is alexithymia characterised by impaired interoception? Further evidence, the importance of control variables, and the problems with the Heartbeat Counting Task. *Biological Psychology*, 136, 189–197.
- Murphy, J., Geary, H., Millgate, E., Catmur, C., & Bird, G. (2017). Direct and indirect effects of age on interoceptive accuracy and awareness across the adult lifespan. *Psychonomic Bulletin & Review*, 1–10.
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American psychologist*, 17(11), 776.
- Paulus, M. P., & Stein, M. B. (2010). Interoception in anxiety and depression. *Brain Structure and Function*, 214(5–6), 451–463.

- Pennebaker, J. W. (1981). Stimulus characteristics influencing estimation of heart rate. *Psychophysiology*, 18(5), 540–548.
- Phillips, G. C., Jones, G. E., Rieger, E. J., & Snell, J. B. (1999). Effects of the presentation of false heart-rate feedback on the performance of two common heartbeat-detection tasks. *Psychophysiology*, 36(4), 504–510.
- Pollatos, O., Kurz, A.-L., Albrecht, J., Schreder, T., Kleemann, A. M., Schöpf, V., Kopietz, R., Wiesmann, M., & Schandry, R. (2008). Reduced perception of bodily signals in anorexia nervosa. *Eating Behaviors*, 9(4), 381–388.
- Pollatos, O., Traut-Mattausch, E., & Schandry, R. (2009). Differential effects of anxiety and depression on interoceptive accuracy. *Depression and Anxiety*, 26(2), 167–173.
<https://doi.org/10.1002/da.20504>
- Ring, C., & Brener, J. (1996). Influence of beliefs about heart rate and actual heart rate on heartbeat counting. *Psychophysiology*, 33(5), 541–546.
- Ring, C., & Brener, J. (2018). Heartbeat counting is unrelated to heartbeat detection: A comparison of methods to quantify interoception. *Psychophysiology*, e13084.
- Ring, C., Brener, J., Knapp, K., & Mailloux, J. (2015). Effects of heartbeat feedback on beliefs about heart rate and heartbeat counting: A cautionary tale about interoceptive awareness. *Biological Psychology*, 104, 193–198.
- Schaefer, M., Egloff, B., & Witthöft, M. (2012). Is interoceptive awareness really altered in somatoform disorders? Testing competing theories with two paradigms of heartbeat perception. *Journal of Abnormal Psychology*, 121(3), 719.
- Schandry, R. (1981). Heart beat perception and emotional experience. *Psychophysiology*, 18(4), 483–488.

- Shah, P., Catmur, C., & Bird, G. (2016). Emotional decision-making in autism spectrum disorder: The roles of interoception and alexithymia. *Molecular Autism*, 7(1), 43.
- Shah, P., Catmur, C., & Bird, G. (2017). From heart to mind: Linking interoception, emotion, and theory of mind. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 93, 220.
- Shah, P., Hall, R., Catmur, C., & Bird, G. (2016). Alexithymia, not autism, is associated with impaired interoception. *Cortex*, 81, 215–220.
- Tysk, L. (1984). Time perception and affective disorders. *Perceptual and Motor Skills*, 58(2), 455–464.
- Verdejo-Garcia, A., Clark, L., & Dunn, B. D. (2012). The role of interoception in addiction: A critical review. *Neuroscience & Biobehavioral Reviews*, 36(8), 1857–1869.
- Weineck, F., Messner, M., Hauke, G., & Pollatos, O. (2019). Improving interoceptive ability through the practice of power posing: A pilot study. *PloS One*, 14(2), e0211453.
- Weisz, J., Bálažs, L., & ÁDám, G. (1988). The influence of self-focused attention on heartbeat perception. *Psychophysiology*, 25(2), 193–199.
- Williams, D. P., Jarczok, M. N., Ellis, R. J., Hillecke, T. K., Thayer, J. F., & Koenig, J. (2017). Two-week test–retest reliability of the Polar® RS 800 CX™ to record heart rate variability. *Clinical Physiology and Functional Imaging*, 37(6), 776–781.
- Windmann, S., Schonecke, O. W., Fröhlig, G., & Maldener, G. (1999). Dissociating beliefs about heart rates and actual heart rates in patients with cardiac pacemakers. *Psychophysiology*, 36(3), 339–342.
- Wittmann, M. (2009). The inner experience of time. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1525), 1955–1967.

- Yates, A. J., Jones, K. E., Marie, G. V., & Hogben, J. H. (1985). Detection of the heartbeat and events in the cardiac cycle. *Psychophysiology*, 22(5), 561–567.
- Zamariola, G., Luminet, O., Mierop, A., & Corneille, O. (2019). Does it help to feel your body? Evidence is inconclusive that interoceptive accuracy and sensibility help cope with negative experiences. *Cognition and Emotion*, 1–12.
- Zamariola, G., Vlemincx, E., Luminet, O., & Corneille, O. (2018). Relationship between interoceptive accuracy, interoceptive sensibility, and alexithymia. *Personality and Individual Differences*, 125, 14–20.
- Zoellner, L. A., & Craske, M. G. (1999). Interoceptive accuracy and panic. *Behaviour Research and Therapy*, 37(12), 1141–1158.