Modality Differences in Timing: Testing the Pacemaker Speed Explanation

Emily A. Williams (Emily.A. Williams@Manchester.ac.uk) Andrew J. Stewart (Andrew.J. Stewart@Manchester.ac.uk) Luke A. Jones (Luke.Jones@Manchester.ac.uk) Division of Neuroscience and Experimental Psychology,

The University of Manchester, UK

Abstract

A classic effect in the timing field is that "sounds are judged longer than lights" (Goldstone, Boardman & Lhamon, 1959). Recently, judgements for tactile durations have been found to fall between the two (Jones, Poliakoff & Wells, 2009). These modality differences are commonly interpreted within scalar timing theory as the work of a central pacemaker which runs faster for sounds, then vibrations, and slowest for lights (Wearden, Edwards, Fakhri & Percival, 1998). We investigated whether verbal estimates and temporal difference thresholds are correlated within each modality, but found this not to be the case. This suggests that differences in pacemaker speed may not be the main driver for modality differences in thresholds. In addition, we investigated sensory bias as an alternative to the pacemaker explanation, but this was found not to correlate with modality differences in timing.

Keywords: Time perception; interval timing; sensory modalities; pacemaker-accumulator; sensory bias.

Introduction

The timing of stimulus duration by humans has historically been under-researched compared to other perceptual domains. One reason is that although humans possess a very sensitive discrimination for duration (with difference thresholds as low as 10 ms), there is no sensory organ for time. This forces explanations to draw on hidden processes more heavily than for other sensory systems, such as vision and hearing. To date the most successful models of human timing have centred on the idea that humans possess an internal clock of a pacemaker-accumulator type, such as in scalar timing theory (SET: Gibbon, 1977; Gibbon, Church & Meck, 1984). The pacemaker generates internal events ('pulses' or 'ticks') which are connected to an accumulator via a switch. The accumulator contents increase linearly with the duration being estimated and forms the basis for timing judgments (further memory and decision modules are also typically added to this clock model - see Gibbon et al., 1984).

Support for the idea of a pacemaker-accumulator internal clock comes from several sources. People (and animals) can stop and start timing like a stopwatch, even managing to 'pause' timing and continue after a short gap (Buhusi & Meck, 2009). In addition, people can perform ordinality judgements, express one duration as a proportion of another, and average durations together, again suggesting a linear relationship between perceived time and real time (Wearden & Jones, 2007).

Furthermore, it appears the speed of the internal clock can be altered. A key signature of a change in pacemaker speed is the 'slope effect'. In verbal estimation¹ tasks, when the stimulus durations are plotted against estimates of those durations, the difference between the experimental condition and control manifests itself as a difference in slope. This is consistent with a multiplicative increase in pacemaker speed, rather than a simple bias (which manifests as a difference in intercept). Such slope effects have been found for certain drugs (Meck, 1984), body temperature changes (for review see Wearden & Penton-Voak, 1995), repetitive stimulation (Penton-Voak, Edwards, Percival & Wearden, 1996), and filled versus unfilled durations (Wearden, Norton, Martin & Montford-Bebb, 2007).

It is known that durations of sounds are judged longer than lights (Goldstone et al., 1959). This effect manifests as a difference in slope (Wearden et al., 1998), where the auditory slope is steeper than the visual slope. It has been argued that the pacemaker runs at a faster speed for auditory than visual stimuli (Wearden et al., 1998; Penney, Gibbon & Meck, 2000). This has sparked a debate about whether there is a central pacemaker that runs at different speeds for different modalities, or separate pacemakers for each modality (See Grondin, 2010). These auditory-visual differences have been found to occur on a range of timing tasks, from temporal generalization (Wearden et al., 1998), to temporal bisection (Penney et al., 2000), to temporal difference thresholds (Jones et al., 2009), suggesting the effect is not task-dependent.

Recently, the temporal judgement of tactile stimuli has been investigated. Jones et al. (2009) found that verbal estimation slopes and temporal difference thresholds for tactile stimuli fall between those for auditory and visual stimuli. Additionally, the two tasks share an inverted pattern, where estimation slopes are highest (most accurate) and thresholds are lowest (most sensitive) for auditory stimuli, for example. It has been suggested that a faster pacemaker is a more accurate pacemaker (Troche & Rammsayer, 2011), which appears to be the case, but this assertion has yet to be empirically investigated. Therefore, the present series of studies will begin with a replication of Jones et al. (2009), but analysis will also examine whether estimation slopes and difference thresholds correlate with each other for auditory, visual, and tactile stimuli.

In contrast to the pacemaker speed explanation, it has been suggested that modality differences could be due to

¹ 'Verbal estimation' is a misnomer stemming from experiments where participants verbalized their estimates of duration, before the introduction of computers.

intrinsic differences between the different sensory systems (Yuasa & Yotsumoto, 2015), e.g. some combination of differences in transduction rates (Zampini, Shore & Spence, 2003) and attentional biases (Spence, Shore & Klein, 2001). Therefore, we will investigate whether these aspects of sensory bias or salience correlate with differences between auditory, visual and tactile stimuli. The present study will operationalise sensory bias as the 'point of subjective simultaneity' on a temporal order judgement task, i.e. the duration that one modality has to precede another by, in order for the two modalities to be judged as simultaneous.

In summary, the aim of the current work is to investigate whether verbal estimates and temporal difference thresholds correlate within each modality, and whether modality differences can be alternatively explained by a measure of sensory bias.

Experiment 1a: Verbal Estimation

Method

Participants 52 right-handed participants (staff and students of the University of Manchester and some members of the general population) completed all three tasks in a counterbalanced order and received $\pounds 10$ for their time.

Apparatus and Materials Participants were seated at a table in a dark room, with their chin on a chin rest. A PC presented the experiments, written in E-Prime (Psychology Software Tools, Pittsburgh, PA). A 17" Samsung Syncmaster monitor stood at a distance of 60 cm. Participants' eyes were level with the top of the monitor and the fixation cue and questions were displayed 20° below eye level. A black foam grip (5.5 x 9.5 x 4.5 cm) was secured to the table 30cm in front of participants in the centreline. Behind the grip was a Sony speaker, which presented the auditory stimuli (500 Hz sine wave tones), and to the left of the grip was a numerical keypad (8.5 x 12 cm) for use with the left hand.

The grip housed an Oticon-A (100 Ohm) bone conductor with vibrating surface 1.6 cm x 2.4 cm. The bone conductor was inset into the foam in the index finger position when gripped with the right hand, and driven by a 500 Hz sine wave signal through a TactAmp 4.2 amplifier (Dancer Design). Visual stimuli were presented via a 6 mm green LED light (87 cd/m²), embedded in a black plastic casing (4 x 4 x 1.75 cm) and attached on top of the foam block. The LED was 16° below the fixation cue (36° below eye level) and 32 cm in front of participants.

Participants wore 3M Peltor ear protectors with inset earphones, which played white noise (56 dB) during the tasks to mask the sound of the vibrations.

Procedure On each trial participants estimated the duration of a stimulus. The task contained 150 trials, where ten stimulus durations (77, 203, 348, 461, 582, 767, 834, 958, 1065, and 1183 ms) were presented in each modality

(auditory, visual, and tactile) five times. Trials were grouped into three counterbalanced blocks by modality.

Each trial began with the presentation of a fixation cross for 500-1000 ms, which was followed by the stimulus. Participants were prompted on-screen to type in their estimate in milliseconds and were reminded that 1 second = 1000 ms. The task lasted approximately 17 minutes.

Results

Outliers were defined as estimation functions that were invariant to stimulus duration (identified as linear regressions not significantly different to 0), which suggested an inability to perform the task. This led to the exclusion of one individual, leaving a sample of 51 participants. See Figure 1 for the mean verbal estimates for each modality.

The hypothesis that verbal estimates would be highest for auditory stimuli and lowest for visual stimuli was examined using a factorial ANOVA with two repeated measures factors: modality (auditory, visual and tactile) and stimulus duration.



Figure 1: Mean verbal estimates for each modality against stimulus durations.

The ANOVA found a main effect of stimulus duration, $F_{(2.60, 130.17)} = 750.70$, p < .001, $\eta_p^2 = .938$. Post hoc analyses revealed that each of the 10 stimulus durations were estimated as significantly differently from each other (p < .001 for all comparisons).

There was also a main effect of modality, $F_{(2, 100)} = 7.50$, p = .001, $\eta_p^2 = .131$. Post hoc analyses revealed that participants estimated auditory stimuli to be significantly longer than visual (p = .006) and tactile (p = .012) stimuli. However, estimates for visual and tactile stimuli did not significantly differ (p = .909).

The interaction between stimulus duration and modality was also significant, $F_{(8.39, 419.32)} = 4.914$, p < .001, $\eta_p^2 = .089$. In order to investigate this interaction, linear regressions were conducted to extract the slope and intercept values of each participant's verbal estimation function for each modality. See Figure 2 for mean slope values.



Figure 2: Mean slope values for auditory, visual and tactile stimuli. Error bars denote standard error.

A repeated measures one-way ANOVA comparing the slopes across modalities found a significant difference between them, $F_{(2, 100)} = 12.76$, p < .001, $\eta_p^2 = .203$. Post hoc analyses confirmed that auditory slopes were significantly higher than visual slopes (p < .001), but not significantly different to tactile slopes (p = 1.00). In addition, the tactile slopes were significantly higher than visual slopes (p = .001).

Discussion

Verbal estimation slopes for auditory stimuli were significantly and multiplicatively higher than those for visual stimuli, with tactile slopes falling between the two. As perfect estimates would have a slope of 1, this suggests that people are more accurate when estimating durations of sounds and vibrations than lights, but tend to underestimate all three modalities. This is the same pattern of results found as in Jones et al. (2009).

Auditory and tactile estimates differed significantly in the first ANOVA, but further analysis on slopes (as pacemaker speed differences are said to manifest as slope effects) found the slopes not to differ. The significant difference between the two in the first ANOVA was perhaps due to a difference in intercept.

Experiment 1b: Temporal Difference Thresholds

Method

Participants The same participants completed this experiment as in Experiment 1a.

Apparatus and Materials The same apparatus and materials were used as in Experiment 1a.

Procedure Participants completed a 50-trial threshold task in each of the three modalities in a counterbalanced order. The test stimuli were the same as in Experiment 1a.

Each trial began with the presentation of a fixation cross for 500-1000 ms, which was followed by the stimuli. The first stimulus (the standard) was always 700 ms, while the second stimulus (the comparator) began at 1000 ms in duration. A 500-1000 ms delay occurred between the two stimuli, and a 125-250 ms delay followed the second stimulus. The order of the standard and the comparator was counterbalanced between trials. Participants pressed '1' or '2' on the keypad depending on whether they thought the first or the second stimulus was longer.

This task used a weighted 3-up 1-down staircase method (Kaernbach, 1991), which allowed for the calculation of the difference in stimulus durations that participants can discriminate 75% of the time. The step size was 15 ms for the first 30 trials, then 10 ms for the last 20 trials. Thresholds were calculated as the mean difference between the standard and comparator durations across the last 20 trials. The task took approximately 18 minutes to complete.

Results

Outliers were defined as thresholds greater than 600 ms (twice the starting difference) which suggested an inability to perform the task. However, no participant had thresholds above this value, giving a full sample of 52 participants.

Figure 3 shows the mean difference between the standard and comparator durations across the 50 trials for the three modalities. The resulting temporal difference thresholds can be seen in Figure 4.



Figure 3: Mean difference between the standard and comparator across the 50 trials for each modality. The vertical dashed line separates the last 20 trials over which the temporal difference thresholds were calculated.

The hypothesis that thresholds would differ according to the modality of the stimuli was examined using a one-way repeated measures ANOVA. This test found a significant difference between thresholds for the different modalities, $F_{(2, 102)} = 30.89$, p < .001, $\eta_p^2 = .377$. Post hoc analyses

confirmed that thresholds for auditory stimuli and tactile stimuli were significantly lower than thresholds for visual stimuli (p < .001 for each comparison). However, thresholds for auditory and tactile stimuli did not significantly differ (p = .079).



Figure 4: Mean temporal difference thresholds for each modality. Error bars denote standard error.

Discussion

Thresholds for visual stimuli were significantly higher than both auditory and tactile stimuli, while auditory and tactile thresholds did not significantly differ. This suggests that people have greater sensitivity to the durations of sounds and vibrations than lights. This pattern of thresholds was reported previously by Jones et al. (2009).

Research Question 1: Do Estimates and Thresholds Correlate within Each Modality?

Results

The same outlier criteria were applied as in the previous sections, with the addition of values 2.5 SDs from the mean. Following removal, this left 51, 49 and 51 participants for auditory, visual and tactile correlations respectively.

Three Pearson's product-moment correlation coefficients found no correlations between estimation slopes and thresholds within each modality (See Table 1).

Table 1: Correlations between verbal estimation slopes and temporal difference thresholds within each modality.

| Modality | df | r | р |
|----------|----|-----|------|
| Auditory | 49 | 099 | .490 |
| Visual | 47 | 123 | .398 |
| Tactile | 49 | 130 | .362 |

Discussion

It had been argued that the differences between modalities in these two tasks were due to the pacemaker running at a faster rate for auditory stimuli and a slower rate for visual stimuli (Wearden et al., 1998; Jones et al., 2009) and that a faster pacemaker leads to greater accuracy and sensitivity (Troche & Rammsayer, 2011). However, accuracy in estimates (slopes) and sensitivity to duration (thresholds) did not correlate for within any modality. This poses a problem for applying the pacemaker explanation to both of these tasks. It could be argued that estimates (magnitude judgements) and thresholds (discrimination judgements) rely on different mechanisms and are of different levels of abstraction, but we expected small correlations despite the transformative nature of estimations.

Experiment 2: Sensory Bias, measured by PSS

This experiment will calculate sensory bias, as measured by the point of subjective simultaneity (PSS) on a temporal order judgement task. The PSS measures the duration that one modality has to precede another by, in order for the two modalities to be judged as simultaneous. This can be seen as a measure of relative salience between the different senses and is affected by the intrinsic properties of each sensory system, e.g. transduction rates (Zampini et al., 2003) and attentional biases (Spence et al., 2001).

Previous research has found sensory biases (measured by PSS) in favour of auditory stimuli when compared with visual (Zampini et al., 2003) and tactile stimuli (Zampini, Shore & Spence, 2005) and in favour of tactile stimuli when compared with visual stimuli (Spence et al., 2001). Therefore, PSSs appear to follow the same modality pattern as estimates and thresholds.

This measure of sensory bias will be investigated as an alternative explanation for the differences between auditory, visual and tactile performance on estimation and threshold tasks in the next section.

Methods

Participants The same participants completed this experiment as in Experiment 1a and 1b.

Apparatus and Materials The same apparatus and materials were used as in Experiment 1a and 1b.

Procedure Participants were presented with two crossmodal stimuli (15 ms each) in quick succession and were asked which occurred first. The task contained 300 trials, where participants were presented with three modality pairs (Aud-Vis, Aud-Tac, Vis-Tac), at 10 different stimulus onset asynchronies (SOAs, -400, -200, -90, -55, -20, +20, +55, +90, +200, and +400 ms), each repeated 10 times. Negative SOAs mean that the first-named stimulus in the pairing came first (e.g. auditory in the Aud-Vis stimulus pair), whereas positive SOAs indicate that the second-named modality came first. Trials were separated into three counterbalanced blocks by modality pair.

On each trial a fixation cross appeared on the screen after a 500 ms delay, where it remained for the rest of the trial. The first stimulus was presented following a random duration between 500-1000 ms. After the randomly selected SOA, the second stimulus was presented. For example, if the Aud-Vis modality pair was presented with the -400 ms SOA, participants heard a 15 ms tone, followed by a delay of 385 ms, and then saw the green LED illuminate for 15 ms. After a 125-250 ms delay, participants were then prompted to answer "Which stimulus came first?" and participants pressed '1' or '2' on the keypad. The task took approximately 15 minutes to complete.

Results

Cumulative Gaussian psychometric functions were fitted to participants' individual data, coded according to 'proportion auditory-first', for example. The PSS and just noticeable difference (JND) were extracted for each individual for each modality pair.

Participants' PSSs were inspected for outliers, identified as those with related JNDs greater than 400 ms (Zampini et al., 2003), which suggested an inability to complete the task. This resulted in the exclusion of eight individuals, leaving a sample of 44 participants. See Figure 5 for PSSs for each cross-modal comparison.





Figure 5: Mean Point of Subjective Simultaneity for auditory-visual, visual-tactile and tactile-auditory comparisons. Error bars denote standard error.

No significant sensory bias was found between auditory and visual stimuli, indicated by the PSS not departing from zero ($t_{(43)} = .94$, p = .354). However, the PSS for visualtactile comparisons was significantly above zero, $t_{(43)} =$ 3.24, p = .002. Participants were biased in favour of tactile stimuli in this comparison, and required visual stimuli to be presented 37 ms before tactile stimuli, for the pair to be judged as simultaneous. In addition, the PSS for auditorytactile comparisons was also significantly above zero ($t_{(43)} =$ 3.21, p = .003). Participants were biased in favour of auditory stimuli in this comparison, and required tactile stimuli to be presented 22 ms before auditory stimuli for subjective simultaneity.

Discussion

Significant sensory biases were found in favour of auditory stimuli when compared with tactile stimuli, and in favour of tactile stimuli when compared with visual stimuli, which concurs with previous research (Spence et al., 2001; Zampini et al., 2005). However, no significant sensory biases were found between auditory and visual stimuli. This was unexpected and is contrary to both previous research (Zampini et al., 2003), and our hypothesis that the large differences between auditory and visual estimates and thresholds may be due to large sensory biases.

Nevertheless, the next section will investigate whether these sensory biases correlate with the differences between modalities in estimates and thresholds.

Research Question 2: Do cross-modal PSSs correlate with the differences between modalities in estimates and thresholds?

Results

The same outlier criteria were applied as in the previous sections, leaving a sample 44 participants.

Six Pearson's product-moment correlation coefficients found no correlations between PSSs and estimation slopes or thresholds for any cross-modal pair (See Table 2).

Table 2: Correlations between cross-modal PSSs and slope and threshold differences for each cross-modal pair.

| Variable 1 | Variable 2 | r | р |
|-------------|---------------------|------|------|
| Aud-Vis PSS | Aud – Vis Slope | .004 | 1.00 |
| | Aud – Vis Threshold | .061 | 1.00 |
| Aud-Tac PSS | Aud – Tac Slope | 217 | .314 |
| | Aud – Tac Threshold | .288 | .116 |
| Tac-Vis PSS | Tac – Vis Slope | .323 | .066 |
| | Tac – Vis Threshold | 333 | .054 |

Discussion

Cross-modal sensory biases, as measured by PSSs, were found not to correlate with the differences in estimates and thresholds between each modality pair. This suggests that the differences in auditory, visual and tactile estimates and thresholds cannot be explained by the intrinsic sensory biases of the three different systems.

General Discussion

We aimed to investigate the pacemaker explanation for differences between auditory, visual and tactile estimates and thresholds, and discovered three main findings. Firstly, the pattern of differences between the modalities appears to be robust as we replicated these in Experiments 1a and 1b (with minor differences in magnitude). Secondly, estimates and thresholds do not correlate within each modality. This poses a problem for the idea that both the slopes in verbal estimation and the order of thresholds are mostly determined by pacemaker rate. Finally, the modality differences in estimates and thresholds did not correlate with sensory biases, potentially ruling out this alternative explanation. These findings generate several possible conclusions:

- 1. Pacemaker rate does not determine estimation slopes or threshold values
- 2. Pacemaker rate determines estimation slopes but not threshold values (or vice-versa)
- 3. Pacemaker rate contributes to both slopes and thresholds, but this contribution is washed out by other cognitive processes

Despite these theoretical uncertainties, the present research can state that the assertion that faster pacemakers give rise to smaller thresholds (Troche & Rammsayer, 2011) is flawed, if one assumes that the pacemaker underlies both estimation and threshold tasks.

At present, there is no published model of how the scalar timing theory system operates in threshold tasks, unlike for temporal generalization (Droit-Volet, Clément & Wearden, 2001), temporal bisection (Wearden, 1991) and verbal estimation (Wearden, 2015). Additionally, the mathematical consequences of increasing pacemaker speed on timing performance (and the assumptions this is based on) in tasks are not explored or predicted in any great detail in the literature. Therefore, our future work will examine the role of the pacemaker in a model of threshold behavior and model mathematical implications of altering pacemaker rate.

Overall, the simple pacemaker speed explanation appears to fail and a more nuanced explanation is required.

Acknowledgments

Emily Williams was supported by a studentship from the Economic and Social Research Council.

References

- Buhusi, C. V., & Meck, W. H. (2009). Relativity theory and time perception: single or multiple clocks? *PloS* one, 4(7), e6268.
- Droit-Volet, S., Clément, A., & Wearden, J. (2001). Temporal generalization in 3- to 8-year-old children. *Journal of Experimental Child Psychology*, 80, 271-288.
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological review*, 84(3), 279-325.
- Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of sciences*, 423(1), 52-77.
- Goldstone, S., Boardman, W. K., & Lhamon, W. T. (1959). Intersensory comparisons of temporal judgments. *Journal* of Experimental Psychology, 57(4), 243-248.
- Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics, 72*(3), 561-582.

- Jones, L. A., Poliakoff, E., & Wells, J. (2009). Good vibrations: Human interval timing in the vibrotactile modality. *The Quarterly Journal of Experimental Psychology*, 62(11), 2171-2186.
- Meck, W. H. (1986). Affinity for the dopamine D 2 receptor predicts neuroleptic potency in decreasing the speed of an internal clock. *Pharmacology Biochemistry and Behavior*, 25(6), 1185-1189.
- Penney, T. B., Gibbon, J., & Meck, W. H. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *Journal of Experimental Psychology: Human Perception and Performance*, 26(6), 1770-1787.
- Penton-Voak, I. S., Edwards, H., Percival, A., & Wearden, J. H. (1996). Speeding up an internal clock in humans? Effects of click trains on subjective duration. *Journal of Experimental Psychology: Animal Behavior Processes*, 22(3), 307-320.
- Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, *130*(4), 799-832.
- Troche, S. J., & Rammsayer, T. H. (2011). Temporal information processing and mental ability: a new perspective. In *Multidisciplinary Aspects of Time and Time Perception*. Springer: Berlin Heidelberg.
- Wearden, J. H. (1991). Human performance on an analogue of an interval bisection task. *The Quarterly Journal of Experimental Psychology*, 43, 59-81.
- Wearden, J. H. (2015). Mission: Impossible? Modelling the verbal estimation of duration. *Timing & Time Perception*, 3(3-4), 223-245.
- Wearden, J. H., & Jones, L. A. (2007). Is the growth of subjective time in humans a linear or nonlinear function of real time? *The Quarterly Journal of Experimental Psychology*, 60(9), 1289-1302.
- Wearden, J. H., & Penton-Voak, I. S. (1995). Feeling the heat: Body temperature and the rate of subjective time, revisited. *The Quarterly Journal of Experimental Psychology*, 48(2), 129-141.
- Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *The Quarterly Journal of Experimental Psychology: Section B*, *51*(2), 97-120.
- Wearden, J. H., Norton, R., Martin, S., & Montford-Bebb, O. (2007). Internal clock processes and the filled-duration illusion. *Journal of Experimental Psychology: Human Perception and Performance*, 33(3), 716.
- Yuasa, K., & Yotsumoto, Y. (2015). Opposite distortions in interval timing perception for visual and auditory stimuli with temporal modulations. *PloS one*, *10*(8), e0135646.
- Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental* brain research, 152(2), 198-210.
- Zampini, M., Shore, D. I., & Spence, C. (2005). Audiovisual prior entry. *Neuroscience letters*, *381*(3), 217-222.