Article:

Plaisier, MA, Holt, RJ orcid.org/0000-0002-1614-5795 and Kappers, AML (2020) Representing numerosity through vibration patterns. IEEE Transactions on Haptics, 13 (4). pp. 691-698. ISSN 1939-1412

https://doi.org/10.1109/toh.2020.2988211

Representing numerosity through vibration patterns

Journal:	*Transactions on Haptics*
Manuscript ID	TH-2019-11-0202.R2
Manuscript type:	Regular paper
Keywords:	Numerosity perception, Vibration pulse sequences, Perceptual judgment, Haptic perception

Representing numerosity through vibration patterns

Myrthe A. Plaisier, Raymond J. Holt and Astrid M.L. Kappers

Abstract—It can be useful to display information about numerosity haptically. For instance, to display the time of day or distances when visual or auditory feedback is not possible or desirable. Here we investigated the possibility of displaying numerosity information by means of a sequence of vibration pulses. From previous studies on numerosity perception in vision, haptics and audition it is known that numerosity judgment can be facilitated by grouping. Therefore, we investigated whether perception of the number of vibration pulses in a sequence can be facilitated by temporally grouping the pulses. We found that indeed temporal grouping can lead to considerably smaller errors and lower error rates indicating that this facilitated the task, but only when participants knew in advance whether the pulses would be temporally grouped. When grouped and ungrouped series of pulses were presented randomly interleaved, there was no difference in performance. This means that temporally grouping vibration sequences can allow the sequence to be displayed at a faster rate while it remains possible to perceive the number of vibration pulses accurately if the users is aware of the temporal grouping.

Index Terms—Vibration pulse sequences, Perceptual judgment, Numerosity perception, Haptic perception

1 Introduction

Vibration is used in many personal devices such as smartphones, smart watches and fitness bracelets. Often series of vibration pulses are used to convey information to the user. For instance, the Timebuzz app allows to display the time on a smart watch via a series of vibration pulses (www.timebuzz.nl). This app is especially useful for individuals with a combination of visual and auditory impairments. There are different ways to encode the time using vibration pulses. In the Timebuzz app several encodings are available. Telling time by vibration pulses can be improved with training [1], but reaching a hundred percent correct time recognition might take many hours of training. Another example is ActiVibe. This is a set of vibrotactile icons that consists of varying numbers of vibration pulses that convey information about fitness progress [2]. Understanding how vibration pulse sequences should be structured to facilitate perception can make applications more user friendly and easier to learn to use. In a series of pulses often the number of pulses is relevant as this can be used to haptically convey information about quantities or distances. Therefore, we focussed on the perception of the number of vibration pulses.

Visual judgment of large numbers of items (i.e. >4) can be facilitated by clustering these items together into smaller groups [3], [4]. For instance, the six dots on a die are distributed over two groups of three dots. This spatial arrangement could be recognised making it easier to determine the number of dots. A similar facilitation effect has been shown in haptic perception [5]. When participants were asked to judge a number of spheres in the hand, they were faster and more accurate when the spheres were distributed over the two hands instead of all grasped in the same hand. It has been suggested that this is because small numbers of items are processed through a more efficient cognitive mechanism than counting. This mechanism has been labeled subitizing [6]. Subitizing was first found in visual perception, but more recently it has also been shown to exist in haptic perception [7], [8] and there is evidence that a similar mechanism also occurs in auditory perception [9]. See [10] for a recent review of subtitizing in the various perceptual modalities. Generally the idea is that each subgroup can be subitized and added to the running total. This results in faster response times and lower error rates than counting all items sequentially.

Not all types of stimuli, however, can be subitized. In that case, counting is used also for the smallest numerosities. In a recent haptic study, it was found that spatially grouping items also facilitated numerosity judgment for stimuli that cannot be subitized [11]. This is an example that grouping of items can also facilitate numerosity judgment due to other reasons than the occurrence of subitizing. It has been suggested that spatial pattern recognition may play a role in this facilitation. Especially in the case of using fixed dot patterns such as on a die people might, of course, recognise the pattern [12]. Often this is studied in the context of a spatial pattern, however temporal patterns might also facilitate numerosity judgment. While in the aforementioned studies all the items to be enumerated were always presented simultaneously, one can also enumerate sequentially presented items. In the current study we are interested in numerosity judgment of series of vibration pulses, which means that items are presented sequentially. This is quite similar to how auditory stimuli are often delivered.

• *M.A. Plaisier is with the Dynamics and Control section, Department of Mechanical Engineering of Eindhoven University of Technology, Eindhoven, The Netherlands, R.J.Holt is with the School of Mechanical Engineering, University of Leeds, UK and A.M.L.Kappers is with the sections Dynamics and Control, Control Systems Technology and Human Technology Interaction of Eindhoven University of Technology, The Netherlands. E-mail: {m.a.plaisier, a.m.l.kappers}@tue.nl, r.j.holt@leeds.ac.uk*

In audition, sequential presentation of a number of tones has been used to investigate numerosity perception in several studies [9], [13], [14]. When presenting items sequentially, participants might automatically resort to counting. To prevent counting, the items are presented with only short intervals in between. Some of these studies have reported low error rates and/or fast response times for up to 2 or 3 items suggesting that a subitizing-like cognitive process might play a role here. A more recent study, however, indicates that spatially distributed stimuli can be subititized but not sequentially presented stimuli [15]. Nonetheless, it has been found that grouping by varying the frequency of the sound facilitated numerosity judgment of sequentially presented tones [13]. This indicates that indeed grouping can facilitate numerosity judgement of sequentially presented items regardless of whether sequentially presented stimuli can be subitized. An important aspect of sequentially presented items is rhythm. As rhythm is a temporal pattern, it could facilitate numerosity judgment in a similar way as spatial pattern recognition has been suggested to play a role in visual numerosity judgment [16].

Most studies on haptic numerosity judgment used simultaneously presented stimuli. In one of the few previous studies on haptic numerosity judgment using vibration pulse sequences the influence of the time interval between pulses was studied. It was found that judging numerosity became more difficult with smaller time intervals between vibration pulses [17], [18]. When the time between pulses was less than 100 ms participants made many mistakes indicating that they could no longer reliably count the pulses in the sequence. So increasing the time between vibration pulses is a way to lower error rates of the numerosity perception. Another way would be to use multisensory stimuli because Philippi and colleagues found that sequentially presented tactile taps are systematically underestimated and that this underestimation can be reduced by using multisensory stimuli [19].

In the current study we set out to investigate whether grouping by means of adding rhythm to a sequence of vibration pulses facilitates perception of the number of pulses. To this end we presented fast series of vibration pulses and varied the time intervals between the pulses.

2 Experiment 1

2.1 Method

2.1.1 Participants

A total of 11 participants participated in this experiment (one female, age range 22 – 29, five left-handed). The data of one of these participants were omitted from the analysis because this participant had a 0% error rate for all trial types in the practice phase and performed the experiment with the vibration pulse sequences at a higher pace than the others. This means that the data are not comparable. All participants were students of Eindhoven University of Technology and received financial compensation for their participation. They signed informed consent prior to the start of the experiment and the study was approved by the ethical committee of the Human Technology Interaction group of Eindhoven University of Technology.

2.1.2 Setup and stimuli

The hardware consisted of a micro controller (Arduino nano) and a vibration motor (Adafruit mini motor disc). The vibration motor was 10 mm diameter and 2.7 mm thick. The vibrator was switched on by providing 5V to the digital pin of the micro controller to which it was connected.

The vibrator was taped to the middle of the volar side of the left forearm of the participant. A vibration pulse lasted 100 ms and the duration of the break depended on the condition (minimum 60 ms and maximum 107 ms). These intervals was chosen such that each vibration pulse was clearly noticeable, while the sequence was too fast to be counted reliably. This ensured the task was difficult enough to be able to use the error rates as performance measure. Between groups of vibrations a 150 ms break was introduced. Different numbers of vibration pulses were presented (1, 2, 3 ..9 pulses). There were three trial types: ungrouped, ungrouped-equal duration and grouped. In this context "ungrouped" means that all vibration pulses in the sequence were delivered with homogeneous temporal spacing. In the "grouped" case we introduced variation in the temporal spacing to divide the sequence into subgroups. In the ungrouped trials the break between pulses was alway 60 ms. In the grouped trials the break between pulses was 60 ms and the break between groups was 150 ms. Finally, in the ungrouped-equal duration the break between pulses was chosen such the total duration of the sequence for a specific number of pulses was the same as the duration of the grouped equivalent. The break between pulses in the ungrouped-equal duration trials varied with the number of pulses and is shown in Table 1. In the grouped sequences the size of a group was 2 or 3 pulses. The different grouping arrangements are shown in Table 1. For example, a grouped trial with four items in Experiment 1 meant that one pulse of 100 ms was displayed followed by a break of 60 ms and a second pulse of 100 ms. After the second pulse a break of 90 ms was introduced followed by a 100 ms pulse, a break of 60 ms and a fourth 100 ms pulse. Due to a programming error the ungrouped-equal duration trials for 7, 8 and 9 pulses were not exactly the same duration as the grouped trials. The breaks for these trials were maximally 15 ms (18%) shorter than they should have been. The breaks in Experiment 2 were the correct duration for all numerosities. Each numerosity was presented 10 times for each trial type. The order of the groups was randomly shuffled, so 5 pulses could be 3-2 or 2-3. For the ungrouped trials 1, 2, 3 ... up to 9 pulses were presented. For the grouped and ungrouped-equal duration 4 ,5 .. up to 9 pulses were presented. All trial types were presented randomly interleaved in a blocked fashion (in a block all trials types were presented in random order) to ensure that all trial types were homogeneously distributed over the experimental session.

In principle participants could distinguish 1, 2 and 3 pulses by the duration alone. To test whether they really perceived 2 or 3 separate pulses we included catch trials. These were trials in which a single pulse was presented with the same duration as the 2 or 3 pulse trials. Catch trials were repeated 10 times per pulse duration, so 20 catch trials in total.

TABLE 1
Grouping configuration of the grouped trials. In Experiment 1 the order of the groups was randomly shuffled. In Experiment 2 the order of the groups was always as indicated here. The time between pulses for each numerosity in the ungrouped-equal duration trials is indicated in the last two columns for Experiments 1 and 2, respectively.

Number of pulses	Group1	Group2	Group3	Break Exp 1 (ms)	Break Exp2 (ms)
4	2	2		90	107
5	3	2		82	100
6	3	3		78	96
7	3	3	1	75	107
8	3	3	2	73	103
9	3	3	3	71	100

2.1.3 Experimental procedure and design

Participants were instructed to verbally report the number of vibration pulses that were presented. The experimenter entered the response into the computer and started the next trial. Participants wore earmuffs during the experiment to remove sounds that were produced by the vibrator. The vibrator and the vibrations were not visible due to the tape with which it was attached to the arm. Participants were informed about what the maximum number of pulses was and that the trials could be temporally grouped or not. Prior to starting the experiment participants performed a practice session in which each numerosity in each trial type occurred twice in blocked random order just as in the main experiment. During the practice sessions they received feedback on what the correct answer was. They did not receive such feedback during the main experiment. The experimenter notified the participant prior to starting a new trial during both the practice trails as wel as the main experiment.

2.1.4 Vibration pulse characterisation

The vibration motors used in this study were cheap low quality motors. These are used in many commercial applications because they are cost effective. We want our findings to be easily extrapolated to existing devices with a vibration motor incorporated. Furthermore, this study was part of a large scale European project (SUITCEYES, www.suitceyes.eu) in which an application is being developed that will incorporate the specific vibration motors used here. Because the specs of these vibration motors are largely unknown measurements using an accelerometer were performed.

The vibration motor was taped to the forearm in the same way it would be during the experiments and an accelerometer (Adafruit ADLX345) was attached to the top of the vibration motor sampling at 667 Hz. Several sequences of vibration pulses that were representative of the ones used in the experiment were provided while collecting data from the accelerometer. Fast Fourier Transform (FFT) was used to determine the vibration frequency and the waveform envelope was determined using a Hilbert Transform. The waveform envelope was subsequently smoothed using a sliding window of 16 ms. The envelope wraps each of the vibration pulses. Each vibration pulse is considered to be an item to be counted by the participants. Examples of the waveform for a grouped, ungrouped and ungrouped-equal duration sequence are shown in Figure 1A. Shown are accelerations in the direction along the forearm. In the example in Figure1A it can be seen that the envelope has 4 peaks, which means that the correct answer to the number of vibration pulses would be 4. The average vibration frequency during a pulse was determined to be 105 ± 2.8 (SD) Hz. The power spectral density is shown in Figure1B for a grouped sequence with 4 pulses. It can be seen that the peak is slightly above 100 Hz. From the determined waveform shown by the blue line in Figure1A it was estimated what the duration of the break between pulses was. For shorter breaks, the motor did not actually come to a rest before being started up again. Accordingly, to estimate the break duration a threshold value was chosen as the value of the acceleration below which the motor was considered off. Because the choice of this threshold is arbitrary and the calculated break duration will be influenced by the choice of the threshold value, we calculated the duration of the break for a range of threshold values. Figure 1C shows the actual duration of the break as a function of the intended duration for different threshold values. Each target duration was sampled a minimum of 9 times and maximum of 57 times. This differed per target duration because we measured a subset of the signals as they were actually presented in the experiments. Linear regression showed that the targeted break duration and measured break duration followed a linear pattern and the slope was comparable for the different thresholds.

2.2 Results

The error rates are shown in Figure 2A for all trial types. The error rate is defined as the proportion of trials that was answered incorrect. This was determined per numerosity for each participant. It can be seen that for most participants the error rates increased rapidly after 2 or 3 pulses. However, the median error rate for 3 pulses is already at 25% and the error bar for 2 pulses indicates that there were participants who also made already quite some mistakes for 2 pulses. The catch trials were, however, always correctly identified as 1 pulse (error rate of 0%). The trials with two pulses were mistaken for being only 1 pulse in on average 1% of the trials and 3 pulses were never mistaken for being 1 pulse. This indicates that participants could reliably distinguish a single long pulse from 2 or 3 separate pulses. Overall, the error rates for the grouped trials were somewhat lower than for the ungrouped trials, but the error bars also indicate that there was quite some variability between participants. As error-rates do not follow a normal distribution, non-parametric statistical tests were used. A paired Wilcoxon signed rank test comparing the error-rates for the grouped and ungrouped-equal duration trials showed no significant difference ($V = 13, p = 0.15$). Also a comparison between

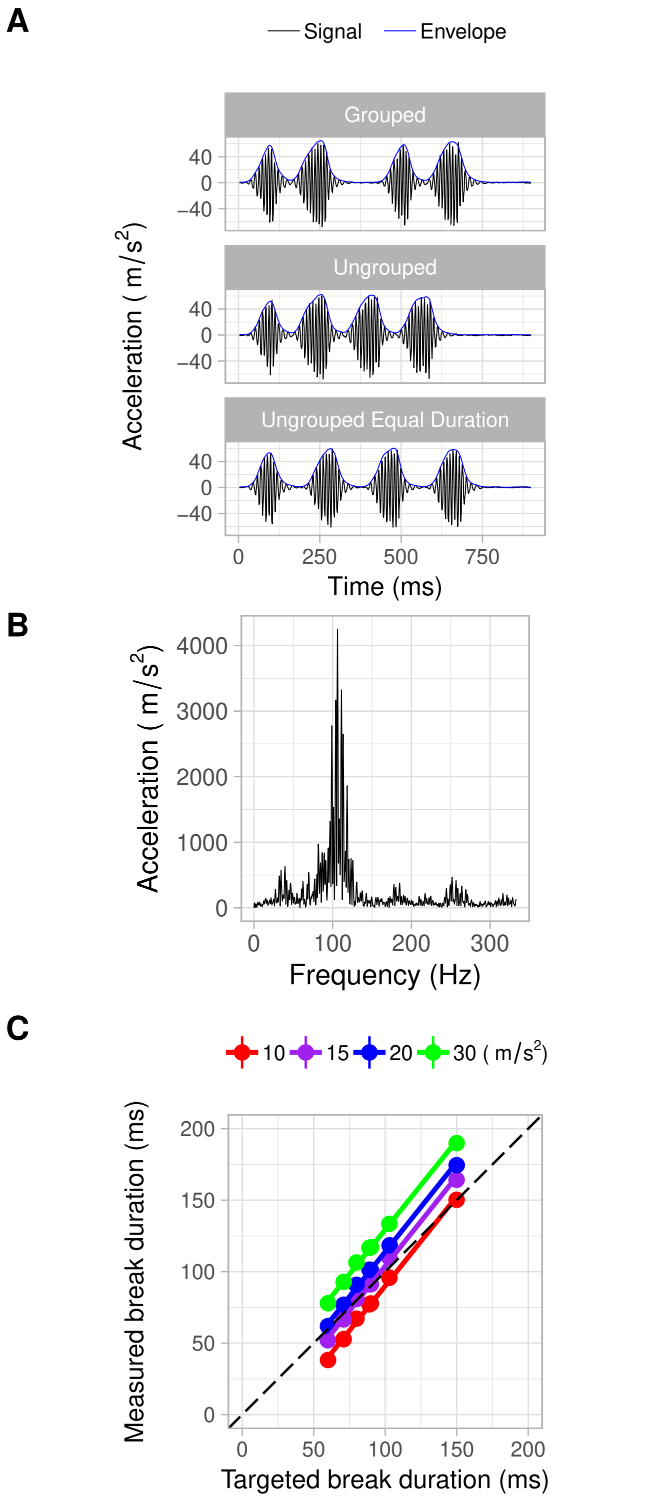

Fig. 1. Characterisation of the vibration pulses. A) shows examples of grouped, ungrouped and ungrouped-equal duration sequences with the blue line indicating the determined envelope. B) Power spectral density for a grouped sequence with four vibration pulses. C) Measured break duration as a function of targeted break duration for different threshold values. Threshold values represent the value of the acceleration below which the motors was classified as off. The error bars indicate the standard deviation. The black dashed line indicates the unity line. Coloured lines indicate linear regression.

the ungrouped trials and the ungrouped-equal duration trials showed no significant difference ($V = 44, p = 0.11$).

For each participant we calculated the median response as a function of the number of vibration pulses. The median was used here as it is less sensitive to outliers than the mean. In Figure 2B the average responses as a function of the number of pulses are shown. For numerosities larger than 3 the responses were systematically below the identity line indicating that participants underestimated the number of pulses. No clear differences between the trial types are visible. Underestimation is a very common finding in when participants are counting, as participants are more likely to fail to count an item than to count an extra item.

Next, we calculated the absolute error (i.e. unsigned error). It is important to look at the absolute errors because they reflect how close participants' answers were to the presented numerosity. When looking at signed errors, under- and over-estimations can average out resulting in a small average error while their answers were quite far off from the presented numerosity. The absolute error that participants made is shown in Figure 2C as a function of the number of items. The absolute error for the grouped trials seems to be systematically smaller than for the other trial types, but this difference is small and variability is considerable. In Figure 2D the absolute error is shown pooled over all numerosities. Because the absolute error does most likely not follow a normal distribution we opted for non-parametric tests. Two paired Wilcoxon signed rank tests were performed on these data to test whether there were differences between the trial types. First it was tested whether the ungrouped and ungrouped-equal duration differed. In the ungrouped-equal duration trials the breaks between pulses were longer and this may have facilitated the task. The paired Wilcoxon signed rank test showed that the absolute error was significantly smaller in the ungrouped-equal duration trails ($V = 5, p = 0.02$). A second paired Wilcoxon signed rank test was performed to test whether there was a difference between the ungrouped-equal duration trials and the grouped trials. This test did not show a significant difference ($V = 40, p = 0.23$).

2.3 Discussion

The results of Experiment 1 did not reveal a clear advantage for temporally grouped vibration pulses. Inspection of the individual participants' data revealed that for 3 out of the 10 participants the error rates for the grouped trials were systematically lower than for the ungrouped trial types. However, for the other participants there was either no clear difference or it was even in the opposite direction. Upon debriefing these participants indicated that they did not like the grouped trials because these were harder to sequentially count and they did not know in advance whether a trial would be grouped or not. In an applied setting grouped and ungrouped sequences would not be randomly interleaved. In fact, the user would most likely be aware of how the pulses would be grouped. Therefore, we decided to perform a second experiment in which grouped and ungrouped trials were presented in a blocked fashion.

Another thing that stood out in the single participants' data was that some participants already had a considerable

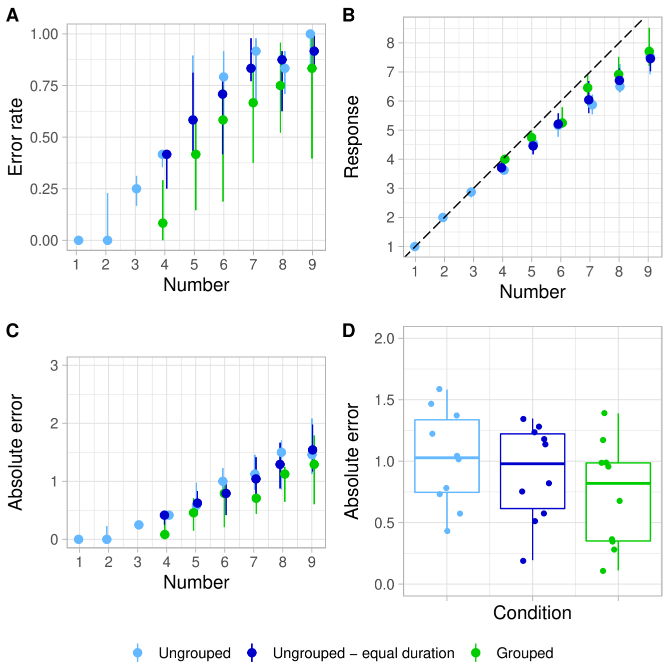

Fig. 2. Results of Experiment 1. Error bars indicate the 25% to 75% intervals. A) The median error rates as a function of the number of pulses B) The median of the average response given by each participant as a function of the number of pulses. The dashed line represents the identity line. C) The median of the average absolute error made by each participant as a function of the number of pulses. D) Boxplot of the average absolute errors pooled over the number of items. The thick horizontal line indicates the median, the box indicates the 25% to 75% interval and the whiskers indicate the 25% and 75% intervals minus or plus 1.5 times the inter-quartile range. Dots indicate the individual participants' data (dots are jittered for clarity).

error rate for 2 and 3 pulses. If they already have difficulty identifying these small numerosities, then grouping larger numerosities into groups of 2 or 3 pulses would not be expected to decrease the error rate. We decided to increase the time between vibration pulses in Experiment 2 to make it easier for participants to perceive 2 and 3 pulses.

3 Experiment 2

3.1 Method

3.1.1 *Participants*

A new group of 10 participants participated in Experiment 2 (5 female, age range 19 – 27, all right-handed). All participants were students of Eindhoven University of Technology and received financial compensation for their participation. They signed informed consent prior to the start of the experiment and the study was approved by the ethical committee of the Human Technology Interaction group of Eindhoven University of Technology.

3.1.2 *Stimuli and Procedure*

The same setup as in Experiment 1 was used. The stimuli were largely the same as for Experiment 1. However, in Experiment 2 we only included the ungrouped-equal duration trials and grouped trials. This time we included trials with 1, 2, and 3 pulses in the block of ungrouped-equal duration trials as well as the block of grouped trials. In these trials with 1, 2 or 3 pulses, the break between pulses was always 80 ms. In the grouped trials the break between pulses was also 80 ms, while the time between groups was 160 ms. To make the ungrouped-equal duration trials the same length as the group trials the time between pulses now was again varied and the time between pulses is in shown in Table 1. In contrast to Experiment 1 the groups were not randomly shuffled. For instance, the order of the groups for 5 pulses was always first 3 pulses followed by 2 pulses. Also, we did not include any catch trials in this experiment.

The procedure was the same as in Experiment 1 except that the different trial types were performed in separate experimental blocks. The order of the experimental blocks was counterbalanced over participants. Within the experimental blocks the trials were presented in blocked random order. This was also the case for the practice sessions. Also, the vibration motor was secured to the arm with an elastic band instead of taping it. Prior to starting the grouped condition the participants were shown a table indicating the grouping arrangement of the pulse sequences (e.g. 3 – 2 for 5 pulses). Like in Experiment 1 the participants performed a practice session with feedback. They practiced the grouped and ungrouped trials directly before starting the corresponding experimental block.

3.2 Results

The median error rates over participants for both trials types are shown in Figure 3A as a function of the number of pulses. It can be seen that the grouped trials consistently had a lower error rate than the ungrouped-equal duration trials. For 9 pulses the grouped trials even had a median error rate of about 15% while the ungrouped-equal duration trials had a median error rate of about 65%. The error bars indicating the 25% to 75% intervals show that also in this experiment there was quite some variability in the error rate between participants, but the error rates for the grouped trials were consistently lower than for the ungrouped-equal duration trials (paired Wilcoxon signed rank test, $V < 0.001, p = 0.009$). The median average response is shown in Figure 3B. It can be seen that there was systematic underestimation in the ungrouped-equal duration trials only. This makes sense because the error rates were quite low for the grouped trials. In Figure 3C it can be seen that the absolute error was consistently lower for the grouped trials than the ungrouped-equal duration trials for all numerosities. Figure 3D shows the same data pooled over all numerosities. A paired Wilcoxon signed rank test on these data showed that the absolute error was significantly lower in the grouped trials than in the ungrouped-equal duration trials ($V = 51, p = 0.02$).

3.3 Discussion

The results of Experiment 2 show a clear advantage for the grouped condition. The error-rates and absolute error were lower in the grouped condition than in the ungrouped condition. That we did not find a clear difference in Experiment 1 indicates that it is necessary to know in advance whether the sequence will be temporally grouped for there to be an advantage for the grouped sequences. This suggests that participants picked either a strategy of sequentially

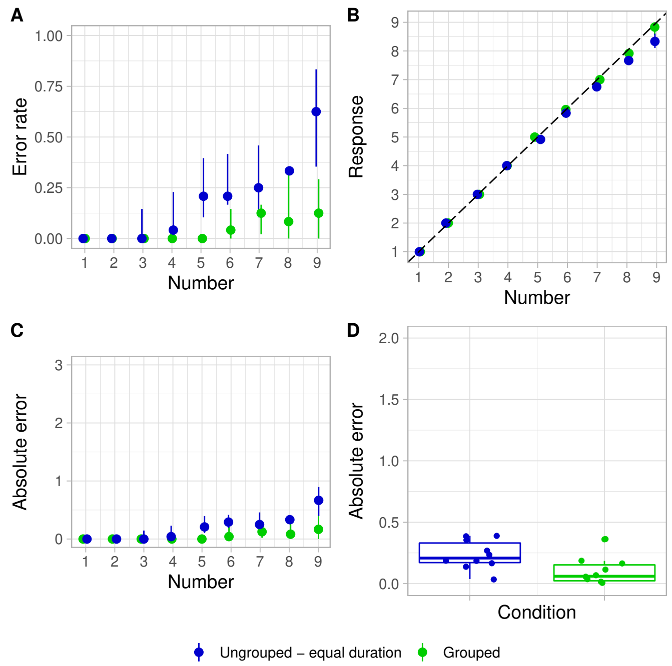

Fig. 3. Results of Experiment 2. Error bars indicate the 25% to 75% intervals. a) the median error rates as a function of the number of pulses b) The median of the average response given by each participant as a function of the number of pulses. The dashed line represents the identity line. c) The median of the average absolute error made by each participant as a function of the number of pulses. d) Boxplot of the average absolute errors pooled over the number of items. The thick horizontal line indicates the median, the box indicates the 25% to 75% interval and the whiskers indicate the 25% and 75% intervals minus or plus 1.5 times the inter-quartile range. Dots indicate the individual participants' data (dots are jittered for clarity).

counting, or counting pulses per group and add those to a running total. Apparently, they picked the strategy in advance. This is in line with what some participants in Experiment 1 reported upon debriefing as they said that they disliked not knowing in advance whether the sequence would be grouped.

4 General Discussion

The results from both Experiments 1 and 2 showed that for the ungrouped vibration pulse sequences it became rapidly more difficult to judge the number of vibration pulses for more than 2 or 3 pulses. For the larger numerosities underestimation occurred. Underestimation of numerosities that can not be reliably counted is quite common in numerosity judgment studies in general and has also been reported in haptic judgment of sequentially presented items specifically [19]. In the same study it was shown that this underestimation decreased (i.e. judgement was more accurate) if a multisensory stimulus was used in which numerosity information was also presented via an auditory channel. Here we show that temporal grouping of vibration pulse sequences makes numerosity judgment more accurate. In Experiment 2 the error rates decreased on average from well above 50% to about 15% for 9 vibration pulses. For some participants, the error rates even dropped to zero for the grouped sequences. Consequently, also the systematic underestimation disappeared. This shows that temporal grouping facilitates numerosity judgment of vibration pulse sequences considerably.

Interestingly, we only found a statistically reliable advantage for numerosity judgement of grouped sequences in Experiment 2. Inspection of each individual participant's data in Experiment 1 showed that for 3 out of 10 participants there was a clear decrease in error rates for the grouped sequences. For the other participants there was no clear difference and one even seemed to perform worse for the grouped than the ungrouped sequences. The main difference between Experiments 1 and 2 was that in Experiment 2 the grouped and ungrouped conditions were performed in a blocked fashion instead of randomly interleaved. This means that participants anticipated the pulses being grouped and this helped to take advantage of the fact that the pulses were grouped. This suggests that the advantage was caused by the use of a cognitive strategy that was not automatic.

In a previous study numerosity judgment of vibration pulse sequences was investigated for signalling fitness progress [2]. It was found that numerosity (or amount/duration etc) is generally better encoded using sequences varying in the number of short pulses rather than using one pulse that varies in duration only. However, this also depends on whether several numbers are communicated in quick succession and also whether there was a distractor task [20]. Another study on numerosity judgment of vibration pulses did show that numerosity judgment was easier when vibration pulses were presented to fingers of the two hands instead of fingers of the same hand [17]. This might be considered as spatial grouping of the vibration pulses. However, they did not find any evidence of subitizing for these stimuli. Our current study was not designed to determine whether subitizing occurs for numerosity judgment of sequences of vibration pulses, but given previous findings it seems unlikely that temporal grouping facilitated numerosity judgement because it enabled subitizing of the subgroups.

One way in which temporal grouping facilitates numerosity judgment is that it introduces a rhythm. For audition it has been suggested that sequentially presented stimuli might be processed by 'sequential subititizing' [21]. So, by introducing rhythm it might be that the subgroups in the grouped trials in our study could be subitized and added to a running total. This would be faster and more accurate than counting all pulses in the sequence. Auditory studies have found that numerosity judgment becomes rapidly more difficult for numerosities above 2 or 3 suggesting a subitizing-like process for sequentially presented stimuli [9], [13], [14]. Recent research, however, suggests that sequentially presented auditory stimuli cannot be subitized [15]. Still, grouping sequentially presented tones using differences in pitch has been reported to facilitate numerosity judgment [13]. This facilitation might, however, not have been due to subitizing.

Although the current study was not specifically designed to detect whether subitizing occured, our results suggest it is not likely that it did. Subitizing is usually considered to be an automatic process, so it would be expected to occur also when grouped and ungrouped sequences were randomly interleaved. However, in Experiment 1 we did not find a difference between grouped and ungrouped trials. Only when we told participants beforehand whether trials would be grouped in Experiment 2, did we find a difference. This suggests a strategy that was cognitively penetrable and chosen before the stimulus was presented. One possibility is that rhythm allowed for keeping the groups of pulses in memory and enabled counting of the pulses after the sequence had been presented. This idea would also fit with our finding that temporal grouping only facilitates numerosity judgment if participants know beforehand that the pulses would be grouped. They would have to choose between trying to keep up with the pulses while counting them as they were presented or trying to memorise the rhythm and count later.

Another way in which temporal grouping might have facilitated numerosity judgment in our study is through pattern recognition. Each numerosity had a specific way in which it was grouped and thus a specific rhythm. For instance, in vision it has been shown that pattern recognition can play a role in numerosity judgment [12]. So, participants in our study might have recognized the specific rhythm for each numerosity. Rhythm possibly made the different numerosities more distinguishable. Vibration patterns are often used to design haptic icons. These are short tangible stimuli that can be associated with a certain meaning. It has been shown that using rhythm can help in designing large sets of icons, while keeping all of these items clearly distinguishable [22].

Overall, our results show that determining the number of vibration pulses in a sequence can be facilitated by temporally grouping the pulses under certain circumstances. Specifically, this is the case if participants are aware of the temporal grouping of the pulses. In the case that pulses are presented always using the same temporal grouping, grouped sequences can be presented faster than they can be sequentially counted while still allowing high precision in determining the number of pulses. This provides a faster way of communicating numerosity information to the user. Potentially, this might also lead to a lower cognitive load, making devices that present vibration pulse sequences less tiring to use.

Acknowledgment

We would like to thank Jordie Kolkman for assistance with data collection. This research was supported by funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 780814, project SUITCEYES.

References

[1] S. Töyssy, J. Raisamo, and R. Raisamo, "Telling time by vibration," in *Haptics: Perception, Devices and Scenarios*, M. Ferre, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 924–929.

[2] J. R. Cauchard, J. L. Cheng, T. Pietrzak, and J. A. Landay, "ActiVibe: Design and Evaluation of Vibrations for Progress Monitoring," in *34th Annual CHI Conference on Human Factors in Computing Systems, CHI 2016*, 2016, pp. 3261–3271.

[3] M. P. Van Oeffelen and P. G. Vos, "Configurational effects on the enumeration of dots: counting by groups." *Memory and Cognition*, vol. 10, pp. 396–404, 1982.

[4] ——, "Enumeration of dots: an eye movement analysis." *Memory and Cognition*, vol. 12, pp. 607–612, 1984.

[5] M. A. Plaisier, W. M. Bergmann Tiest, and A. M. L. Kappers, "Grabbing subitizing with both hands: bimanual number processing," *Experimental Brain Research*, vol. 202, pp. 507–512, 2010.

[6] E. Kaufman, M. Lord, T. Reese, and J. Volkmann, "The discrimination of visual number," *American Journal of Psychology*, vol. 62, pp. 498–525, 1949.

[7] K. J. Riggs, L. Ferrand, D. Lancelin, L. Fryziel, G. Dumur, and A. Simpson, "Subitizing in tactile perception," *Psychological Science*, vol. 17, pp. 271–272, 2006.

[8] M. A. Plaisier, W. M. Bergmann Tiest, and A. M. L. Kappers, "One, two, three, many - subitizing in active touch," *Acta Psychologica*, vol. 131, pp. 163–170, 2009.

[9] V. Camos and B. Tillmann, "Discontinuity in the enumeration of sequentially presented auditory and visual stimuli," *Cognition*, vol. 107, pp. 1135–1143, 2008.

[10] N. Katzin, Z. Z. Cohen, and A. Henik, "If it looks, sounds, or feels like subitizing, is it subitizing? a modulated definition of subitizing," *Psychonomic Bulletin and Review*, 2019, article in Press.

[11] K. E. Overvliet and M. A. Plaisier, "Perceptual grouping affects haptic enumeration over the fingers," *Perception*, vol. 45, no. 1-2, pp. 71–82, 2016.

[12] G. Mandler and B. J. Shebo, "Subitizing: an analysis of its component processes." *Journal of Experimental Psychology: General*, vol. 111, pp. 1–22, 1982.

[13] G. Ten Hoopen and J. Vos, "Effect or numerosity judgement of grouping of tones by auditory channels," *Perception & Psychophysics*, vol. 26, pp. 374–380, 1979.

[14] B. H. Repp, "Perceiving the numerosity of rapidly occurring auditory events in metrical and nonmetrical contexts," *Perception and Psychophysics*, vol. 69, no. 4, pp. 529–543, 2007.

[15] K. L. Roberts, N. J. Doherty, E. A. Maylor, and D. G. Watson, "Can Auditory Objects Be Subitized?" *Journal of Experimental Psychology-Human Perception and Performance*, vol. 45, no. 1, pp. 1–15, 2019.

[16] H. Davis and R. Pérusse, "Numerical competence in animals: Definitional issues, current evidence, and a new research agenda," *Behavioral and Brain Sciences*, vol. 11, pp. 561 – 579, 12 1988.

[17] N. Iida, S. Kuroki, and J. Watanabe, "Comparison of tactile temporal numerosity judgments between unimanual and bimanual presentations," *Perception*, vol. 45, no. 1-2, pp. 99–113, 2016.

[18] J. Pasquero, S. J. Stobbe, and N. Stonehouse, "A haptic wristwatch for eyes-free interactions," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '11. ACM, 2011, pp. 3257–3266.

[19] T. G. Philippi, J. B. F. van Erp, and P. J. Werkhoven, "Multisensory temporal numerosity judgment," *Brain Research*, vol. 1242, pp. 116–125, 2008.

[20] J. R. Blum and J. R. Cooperstock, "Single-Actuator Vibrotactile Numeric Information Delivery in the Face of Distraction," in *IEEE World Haptics Conference (WHC)*, 2019.

[21] C. Davis and W. Roberts, "Lifting Movements In Size-Weight Illusion," *Perception & Psychophysics*, vol. 20, no. 1, pp. 33–36, 1976.

[22] D. Ternes and K. E. MacLean, "Designing large sets of haptic icons with rhythm," in *Haptics: Perception, Devices and Scenarios, Proceedings*, ser. Lecture Notes in Computer Science, Ferre, M, Ed., vol. 5024, 2008, pp. 199–208.

Myrthe A. Plaisier received her Masters degree in experimental physics in 2006 and her PhD in 2010, both from Utrecht University. She performed her PhD research on haptic perception in the group of prof. Astrid Kappers. Her thesis received the thesis award from the Dutch Psychonomics Society. In 2011 she received a Rubicon grant from the Netherlands organisation for scientific research (NWO) which allowed her to continue her research at Bielefeld University in Germany. In 2013 she received a VENI grant from NWO to investigate haptic perception of objects at the Department of Human Movement Sciences at Vrije Universiteit Amsterdam. In 2018 she joined Eindhoven University of Technology to work on the development of assistive technology for persons with deafblindness as part of the European project SUITCEYES.

Raymond J. Holt Raymond Holt received his BEng (Hons) degree and PhD in Mechanical Engineering from the University of Leeds in 2001 and 2005, respectively. Since 2005, he has been a lecturer in Product Design at the University of Leeds. He was appointed a Chartered Engineer and member of the Institute of Mechancal Engineers in 2016. At the university of Leeds, he is a member of the Immersive Cognition research group and the Centre for Disability Studies. His research focuses on technologies to support motor learning and rehabilitation in prehension, and assistive technologies to aid perception and navigation of the environment.

Astrid M.L. Kappers studied experimental physics at Utrecht University, the Netherlands. She received the PhD degree from Eindhoven University of Technology. From 1989 till September 2012, she was with the Department of Physics and Astronomy, Utrecht University. From 2008–2012, she was head of the Human Perception Group of the Helmholtz Institute. In September 2012, she moved with her whole group to the Department of Human Movement Sciences, Vrije Universiteit Amsterdam, the Netherlands. October 2018, she moved to Eindhoven University of Technology, where she works in three groups: Dynamics and Control, and, Control System Technology of the Department of Mechanical Engineering, and Human Technology Interaction of the Department of Industrial Engineering & Innovation Sciences. She was promoted to full professor in 2005. Her research interests include haptic and visual perception. In 2003, she won the prestigious VICI grant. She is/was a member of the editorial boards of Acta Psychologica (2006-present) and Current Psychology Letters (2000-2011) and is an associate editor of the IEEE Transactions on Haptics (2007-2011 and 2017 till present).

Editor Comments

Associate Editor
Comments to the Author:

The reviewers are happy with the revisions to the paper in light of earlier feedback. A few minor points still need to be addressed and then the paper will be ready for publication.

We are happy that the reviewers are positive.

The main point raised is with respect to additional clarification of statements in the discussion relative to the design of E1 and E2, of the accessibility of the temporal structure and participants internal hypotheses with respect to this source of information. Please carefully reconsider the construction and communication of this information in the text.

Please find our point-by-point response to the remaining comments of Reviewer #3 below.

Reviewer: 3

Recommendation: Author should prepare a minor revision

Comments:
The current study is well written and easy to read. It is clear from the introduction that the authors want to understand how the rate at which pulses are displayed plays a role in the perception of these pulses by the participants.

The review will be divided as follows: Main idea/Discussion; Procedure; Stimuli; Spelling mistakes.

Discussion:

Note that the following paragraph is intended to help the reader understand all the features of this study.

The conclusion seems to go a bit too far in terms of the results. As mentioned, the main difference between the two experiments (E1 and E2) is that the conditions involved were performed randomly mixed in E1 and separately in E2. It has also been said that this difference shows the importance of knowing in advance what the pace of the stimulus will be, which means that without this information, the participants had no possibility of obtaining better results with any temporal substructure. Here the participants have a huge hypothesis in E2. This shows the significant impact of the participant's expectation, visible in the comparison between the ungrouped-equal type in E1 and E2 and the grouped type in E1 and E2.
This information appears to be the most important feature of the current study, followed by the clustering effect. While the authors prefer to talk only about the grouping effect, it seems important not to hide the fact that the clustering effect only has an impact on performance when the participants know what they are going to experience (also notably through the training sessions). Therefore, the last paragraph of the discussion may be misleading for the reader as it only states that grouping pulses helps to have a better precision of the number of elements. In addition, the abstract

should also introduce this feature. From a personal point of view, this would more accurately reflect the work done on the two experiments without neglecting the fact that the determination of the number of impulses can be facilitated by grouping them together.

It was not our intention to hide the finding that grouping only leads to facilitation if participants can anticipate the sequence being grouped. We do understand the reviewer's point and agree that this should be improved. Following the suggestions of the reviewer we have added a statement about this to the abstract and the last paragraph of the Discussion has been revised as well.

Procedure:

2.1.3 "Prior to starting the experiment participants performed a practice session in which each numerosity in each trial type occurred twice"
Are these trials randomly ordered? It would be interesting to add the ordering in all cases (both in E1 and E2)

Practice trials were, like trials in the main experiments (E1 and E2), presented in block random order. So all trial types were presented in random order, and this was repeated until the intended number of repetitions was achieved. It was indeed not stated that trials in the practice sessions were block randomized like in the main experiments. In the Experimental Design and Procedure subsection of Experiment1 we now state that the practice trials were presented in blocked random order just as in the main experiment. In the Stimuli and Procedure section of Experiment 2 we added that that within the two separate experimental blocks the trials were presented in blocked random order.

Stimuli:

● Experiment 1
● It is not clear to me if the participants knew when the stimuli started.
The experimenter notified the participant each time a new trial was started. We have added this information to the Experimental Procedure and Design section of Experiment 1.
● "The breaks for these trials were maximally 15 ms shorter than they should have been."
○ The longer the duration of the inter-stimulation, the less impact this error will have on the final result. It might therefore be easier for the reader to have a percentage of time exceeded compared to the expected duration, rather than raw values.

We have added the requested percentage (18%).

● Fig1.A:
○ It really helps in understanding ungrouped and grouped types. It might be a good idea to also have the example of the ungrouped-equal type to visually show the difference between the three conditions, especially since E2 uses only the ungrouped-equal and grouped types. I feel that this would help remove any ambiguity between the types and make it easier/faster to understand for the reader.
The requested plot has been added to Fig.1A

● "The average vibration frequency during a pulse was determined to be 105 ± 2.8 (SD) Hz"

○ This actuator design is known to have both ramp-up and steady state. Its ramp-up lasts between 50 and 100 ms (depending on the orientation of the actuator, the position of the internal moving part, its manufacture,...) and its frequency content is generally not as clear as the authors have shown. It would be interesting for some readers if the authors could share the power spectral density of their characterization. It would give a much better representation of the actual stimulation.

The requested power spectral density plot has been added.

- Fig1.B
  - It is not clear to me why there are 4 different accelerations if the voltage is always the same (5V). Are the pulses different from each other during the experiments? I don't understand the importance of showing the 4 accelerations if only one was used. Is it to show why the authors chose the 100 ms stimuli rather than the others?

The cut-off values shown in Fig. B represent the cut-off values that we used for determining the beginning and ending of a pulse in the analysis. So, the 10 m/s^2 cut-off value means that the duration of a gap was defined as the time that the envelope was below this value. Since the gap duration will depend on this cut-off value we want to show what the differences are if slightly different cut-off values are chosen. Upon rereading the manuscript we realized that this explanation was actually missing. Therefore, we have added a better explanation of Fig1.B to the Vibration Pulse Characterization section. Moreover, we have changed the term "cut-off" to "threshold" for clarity.

Spelling mistakes
- Experiment 1
  - "Finally, in the ungrouped-equal duration [...] such the the total duration […]"
  - "For example, a grouped trial with for four items […]"
- Conclusion
  - "[...] than they can be sequentially counted while still allowing high precission precision in determining the number of pulses."

Spelling mistakes have been corrected.