Differences between measuring methods of time estimation

RIKU ASAOKA（朝岡　陸）[1] and YOICHI WATANABE（渡邊洋一）[2]

The present study compared the accuracy, consistency, and stability of measuring methods of time estimation, particularly the verbal estimation and time reproduction methods. We also investigated the relationship between memory and the two measuring methods. In each trial, participants were alternately presented with a dissyllable nonsense word and a blank screen. Their task was to memorize the presented nonsense words and estimate the duration of the memory task using the verbal estimation and time reproduction methods. The results showed that the accuracy and consistency of the two methods were almost the same, and that the reproduction method had a higher stability than the verbal estimation method. Moreover, performance on the memory task was related significantly to the estimation duration for the reproduction method, but not for the verbal estimation method. These findings suggest that the time reproduction method is a better method for time estimation, and is more strongly related to memory processing, than the verbal estimation method.

Key words: Time perception, Time estimation, Memory, Attention

When studying time estimation, the most important factor is probably the choice of measuring method. It has been considered that the measuring method is one of the most influential factors in time estimation (Hicks, Miller, & Kinsbourne, 1976; Zakay & Block, 1997). Several previous studies reported that using different measuring methods altered the pattern of time estimation, even in a single experimental setting (Brown, 1985; Grondin, 2008; Zakay, 1993). There are three principal methods for studying time estimation. The first method is referred to as the time reproduction method (RP); the experimenter presents a target interval in some way (such as a continuous sound or flash) and the participant reproduces the length of the interval. Second, there is the method of verbal estimation (VE). After the presentation of a target interval, the participant is asked to provide a verbal estimation of its duration, using temporal units, such as seconds or minutes. The third method is called the production method; after the experimenter specifies a target interval in temporal units, a participant produces this interval through some action (for example, pressing keys).

Many studies have investigated the differences between these principal measuring methods, especially RP and VE (e.g., Allan, 1979; Clausen, 1950; Danziger & Du Preez, 1963; Du Preez, 1963; Grondin, 2008; Hawkes, Bailey, & Warm, 1960; Hornstein & Rotter, 1969; Kruup, 1961; McConchie & Rutschmann, 1971; Ochberg, Pollack, & Meyer, 1965; Siegman, 1962; Wallace &

Rabin, 1960). These studies have compared the degree of accuracy or consistency for RP and VE in order to solve which is the most reliable measuring method. The term "accuracy" refers to the degree of closeness between the participants' time estimation and the actual length of the target duration, and the term "consistency" refers to the degree of variability within participants (Allan, 1979). Several previous studies examining the accuracy of the measuring methods have revealed that subjective duration is closer to objective duration in RP than in VE (Brown, 1985; Ochberg et al., 1965). It has also been demonstrated that the actual interval tends to be consistently overestimated using VE compared to RP, even for an equally objective duration (Brown, 1985; Hornstein & Rotter, 1969; McConchie & Rutschmann, 1971; Ochberg et al., 1965; Wallace & Rabin, 1960). Together, these findings suggest that RP has a higher degree of accuracy than VE. However, several studies examining the consistency of the two methods have reported incongruent results. Three studies found that RP has the worst degree of consistency of the three principal methods (Clausen, 1950; Kruup, 1961; Siegman, 1962). Nevertheless, one study provided evidence that the three principal methods have almost equal consistency (Hawkes et al., 1960) and another study demonstrated that VE yields a much lower level of inter-participant variability (Ochberg et al., 1965). The contradiction between these results leads to the conclusion that no single method can claim consistent superiority (Allan, 1979).

In addition to the issue of accuracy and consistency differences, several studies have questioned whether there are significant correlations between the participants' time estimations obtained by each method. Indeed, several studies have demonstrated that RP and VE are not correlated (Clausen, 1950; Danziger & Du Preez, 1963; Kruup, 1961), and these authors suggest that the methods involve different underlying functions. In contrast, one study has reported a significant and strong correlation between both methods (Du Preez, 1963). The author has argued that it is reasonable that there may be a common intra-individual basis for the different methods of time estimation.

The purpose of the current study was twofold. First, we examined not only the accuracy and consistency, but also the stability, of the two methods for a single experimental condition (where the apparatus, stimuli, duration length, and tasks were identical but not the measuring methods). While many studies examine the accuracy and consistency of the two methods, little is known about their stability. In this study, we used the coefficient of variation (CV) as an index of stability. The CV is a measure of the relative variation in distribution, which is independent of the units of measurement: it is the standard deviation divided by the mean or average. The higher the CV value, the greater the dispersion of the variable. We aimed to determine which measuring method is the most accurate, consistent, and stable.

Second, we aimed to characterize the relationship between time estimation and memory. The effects of memory on time estimation have been the subject of controversy since Ornstein (1969) proposed the storage-size model. This model assumed that memory processes and time estimation are closely related. We employed the dual task paradigm in order to examine the relationship between the effects of the measuring method on time estimation and memory

processes. In the dual task paradigm, participants have to do two tasks simultaneously: one is a memory task (memorizing nonsense words) and the other is the time estimation task. By examining the relationship between memory and time estimation, we may gain a better understanding of the measuring methods.

However, there is a problem of attention allocation to each task. When examining time estimation under dual task conditions, Macar, Grondin, and Casini (1994) found that attentional allocation to each task influenced time estimation. Thus, we instructed participants to focus their attention on the presented nonsense words and to ignore the passage of time, in order to control their selective attention. This instruction is also useful to prevent the use of counting strategy, which is defined as counting the number of seconds during target duration in order to intentionally perform accurate time estimation.

In short, the first purpose of the present study was to examine not only the accuracy and consistency, but also the stability, of the two methods in a single experimental condition. We hypothesized that time estimation obtained by RP will be more accurate than VE, based on previous research (Brown, 1985; Hornstein & Rotter, 1969; McConchie & Rutschmann, 1971; Ochberg et al., 1965). The second purpose was to characterize the relationship between time estimation and memory. In order to achieve these aims, we measured participants' time estimation using RP and VE in a dual task paradigm.

Method

Participants

Twenty-two Japanese students, 11 males and 11 females, voluntarily took part in this experiment. Their ages ranged from 20 to 26 years ($M = 22. 23$, $SD = 1.45$). The written consent of each participant was obtained prior to the experiment. All participants had normal or corrected-to-normal sight, and did not know the purpose of the present study.

Apparatus and stimuli

The experiment was conducted on a PC (Dell Dimension 8300). The experimental stimuli of the memory task consisted of Japanese dissyllables. The nonsense words were presented in the center of a 19-in. CRT monitor (Sony Trinitron Multiscan G420, pixel resolution 1024×768, with a refresh rate of approximately 60Hz) in black Mincho Tai 60-point font, on a gray background. Participants viewed the monitor binocularly from a distance of about 60 cm. The presentation of stimuli and recording of participants' responses were performed by the computer using E-Prime ver. 2.0 (Psychology software tools Inc.).

Procedure

Each participant was tested individually. At the beginning of the experiment, the participants were asked to leave their watch and cellphone with the experimenter. They were seated on a chair in a dimly lit room (without clocks) and instructed to perform the memory task and time estimation task. The memory task involved memorizing the successively

presented nonsense words. The time estimation task involved estimating the duration of performing the memory task. Participants were also informed that one trial contains four parts: stimuli presentation, time estimation, recognition test, and a questionnaire. The participants' time estimations were measured using both the RP and VE methods. Specific instructions were given regarding the two methods, and the participants were told not to count seconds mentally during the stimuli presentation and time estimation parts. After the experiment, we asked the participants whether they counted the number of seconds or nonsense words presented, and no participants reported doing so.

Trial design

Stimuli presentation

We chose longer durations for the target presentation following the distinction about length of target duration made by Block and Zakay (1997). Block and Zakay (1997) defined 5.0-14.9 s as "short," 15.0-59.9 s as "moderate," and 60.0 s or longer as "long." In the present study, three moderate or long target durations were used: 52 s, 67 s, and 82 s. The main reason for this choice is that adults have a stable sensation of the length of a second, and participants can stably and accurately estimate the short target durations regardless of the measuring method.

The presentation of the stimuli began when the participants pressed the "Enter" key on the PC keyboard. During each trial, participants were presented with a blank screen and one nonsense word alternately. The duration of each blank scene was 2 s, and the duration of each word was 3 s. One set was composed of one blank scene and one nonsense word (2 s + 3 s = 5 s) repeatedly presented. The number of repetitions varied according to the target duration. Finally, a blank scene lasting for 2 s was added at the end of the stimuli presentation.

The set of one blank scene and one nonsense word was presented repeatedly 10 times in the 52 s condition (5 × 10 + 2), 13 times in the 67 s condition (5 × 13 + 2), and 16 times in the 82 s condition (5 × 16 + 2). This means that the 52, 67, and 82 s conditions include 10, 13, and 16 nonsense words respectively. The nonsense words were all unique. In each trial, the total duration between the onset of the first blank scene and the offset of the last blank scene was the target of time estimation. Figure 1 illustrates the sequence of the stimuli.

Time estimation

After the stimuli presentation, participants estimated the length of the target duration using RP and VE. In RP, the participant's reproduction began when they pressed the spacebar on the keyboard. They then had to press the spacebar again when the time that had elapsed from the beginning of the reproduction was judged to correspond to the length of target duration. While participants were estimating the duration, the words "now measuring" were presented in the center of the display. In VE, the participants were instructed to give a verbal estimate of the length of the target duration in seconds, as accurately as possible. The order of the measuring methods was randomized among experimental blocks. Experimental blocks will

be explained in detail later.

Recognition test

Following the time estimation, the participants were asked to perform a recognition test of the presented nonsense words. The recognition test consisted of 10 presented words and 10 novel words (words that had not been presented) in the 52 s condition, 13 presented words and 13 novel words in the 67 s condition, and 16 presented words and 16 novel presented words in the 82 s condition. The participants were instructed to indicate which of the words they had seen in the stimuli presentation block. The number of words correctly recognized was corrected for guessing by subtracting the number of false alarms from the total number of recognized words (Baddeley, 1990). The rate of correct recognition in each condition served as a measure of recognition test.

Questionnaire

A questionnaire was used to check two points: expectation and attention to the memory task. First, the participants answered a question about expectation because Jones and Boltz (1989) demonstrated that perceived duration is shorter if the target duration ends earlier than their expectation about when it will end, whereas, perceived duration is longer if the target duration ends later than their expectation. The question about expectation used a 7-level Likert scale ranging 1 (*target duration seems to end much earlier than expected*) to 7 (*target duration seems to end much later than expected*). Next, they were asked to evaluate the percentage of attention allocated to the memory task, and the passage of time to check whether they were in accordance with our instruction.

In summary, one trial ran in the following order: stimuli presentation, measurement of time estimation (RP and VE), recognition test, a question about expectation, and reporting the percentage of allocated attention. One trial lasted for about 5 min.
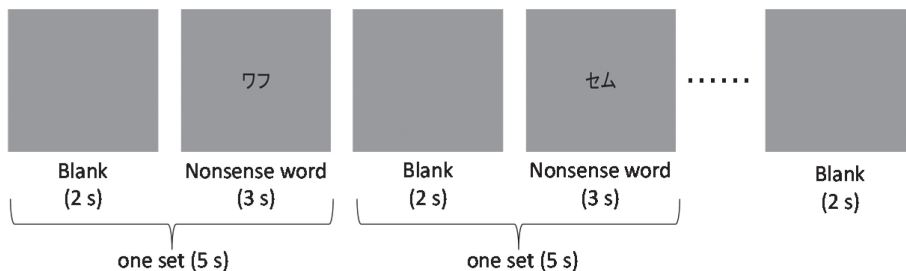

Figure 1. This figure illustrates the sequence of stimuli presentation. Blank scenes and nonsense words were presented for 2 s and 3 s, respectively. A set of one blank scene and one nonsense word was presented repeatedly. The degree of repetition of set presentation was dependent on target duration. In all target duration conditions, blank scenes lasting for 2 s were inserted at the end of stimuli presentation. s = seconds.

Experimental blocks

　　The experiment consisted of two experimental blocks. In one block, each target duration condition was repeated twice, resulting in six trials. As each target duration condition was repeated four times across two blocks, the participants completed 12 trials in total. The two blocks differed in the order of the measuring methods in the time estimation part: in one block, VE was followed by RP, while in the other block, RP was followed by VE. The order in which the trials were administered within blocks was randomized across participants. The order of the two blocks was counterbalanced.

Results

　　At first, two data were excluded from further analysis: one trial (67 s condition) in which the reproduced duration was shorter than 1 s, and one trial (82 s condition) in which the participant paid a great deal of attention to the passage of time (40%; see Table 1). Therefore, the valid data comprised 99.24% of all data (262/264). Table 1 shows the means and standard deviations of the estimated duration obtained by the two methods, performance on the recognition test, expectation, and attention to the memory task for each target duration condition ($n = 22$).

Table 1. Means and standard deviations of estimation duration obtained by the two methods, performance on the recognition test, expectation, and attention allocated to the memory task for the each target duration condition ($n = 22$).

Dependent variables	52 s condition Mean (*SD*)	67 s condition Mean (*SD*)	82 s condition Mean (*SD*)
Reproduction (s)	37.30 (12.19)	46.22 (13.10)	53.05 (14.76)
Verbal estimation (s)	31.08 (12.23)	40.70 (18.20)	48.99 (23.85)
Performance on recognition test (rate)	0.68 (0.186)	0.60 (0.944)	0.55 (0.87)
Expectation (1-7)	3.35 (0.94)	4.34 (0.97)	5.12 (0.72)
Attention to the memory task (%)	93.65 (6.7)	93.19 (6.76)	93.11 (7.45)

SD = standard deviation; s = seconds

　　The means and standard deviations (*SD*) of the RP and VE methods for each target condition are presented in Figure 2 and Table 1. We calculated the coefficient of variation (CV) for RP and VE to examine stability, which is the standard deviation divided by the mean. The mean level of the CV was 0.238 (*SD* = 0.05) in the RP condition, and 0.308 (*SD* = 0.13) in the VE condition. An advantage of using the CV is that it allows the direct comparison of data from the different target durations (Gamache & Grondin, 2010). Therefore, we conducted a paired *t*-test on each mean level of CV for measuring method. The analysis revealed that the CV was significantly higher in the VE condition than in the RP condition, $t(21) = 2.342$,

p < .05.

In order to assess accuracy, one sample *t*-tests were conducted on participants' time estimation. We compared each mean level of estimation duration and each target duration for all conditions. The analyses indicated that participants' time estimation was significantly shorter than the target durations (all *p* < .001). The results of these analyses are shown in Table 2. Then, in order to test the mean difference in participants' estimation duration among all conditions, we conducted a 3 (target duration) × 2 (measuring method) repeated measures ANOVA on both dependent variables. Prior to parametric analysis, all variables were analyzed using the Kolmogorov-Smirnov test and no significant deviation from normality was shown (*p* > .05 for all variables). The ANOVA revealed a significant main effect of target duration, *F* (2, 42) = 67.501, *p* < .001. There was no significant effect of measuring method, *F* (1, 21) = 1.928, *n. s.*, and no significant interaction, *F* (2, 42) = 0.169, *n. s.* This result showed that the measuring methods produced almost similar results for time estimation. Ryan's post hoc test was conducted to test the mean differences for target duration. The 82 s condition differed significantly from the 67 s and 52 s conditions (*p*s < .001), and the 67 s condition differed significantly from the 52 s condition (*p* < .001). The results indicated that the perceived duration increased as the target duration increased.

Correlation coefficients between participants' time estimation in experimental blocks 1 and 2 for the two measuring methods were calculated to assess consistency within participants. Such an analysis has been used in several previous studies (e.g., Clausen, 1950; Siegman, 1962). We calculated the correlations not on the mean of the dependent variables but on each all data (*n* = 262). There were two main reasons for this analysis. First, a sample of *n* = 22 seems to be insufficient for an appropriate calculation of correlation. Second, comparing correlations
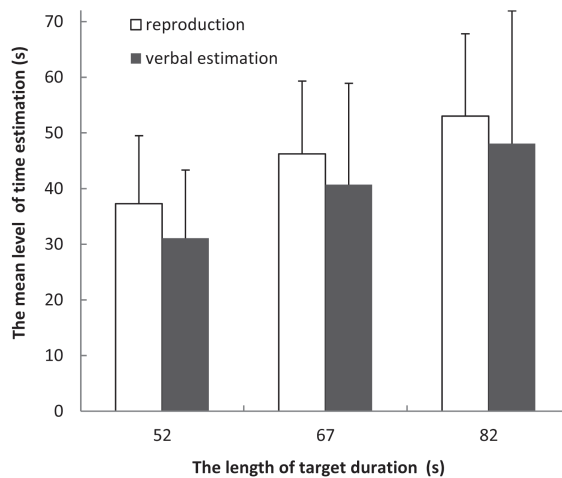

Figure 2. The mean level of time estimation for each condition (target duration × measuring method; *n* = 22). Error bars represent the standard deviations. s = seconds.

Table 2.   The results of one sample *t*-tests examining differences in target duration and the mean level of participants' time estimation for target duration.

Condition	N	*t* value	Test results
Reproduction, 52 s	22	6.59	*p* < .001
Verbal estimation, 52 s	22	9.34	*p* < .001
Reproduction, 67 s	22	9.76	*p* < .001
Verbal estimation, 67 s	22	7.99	*p* < .001
Reproduction, 87 s	22	12.06	*p* < .001
Verbal estimation, 87 s	22	9.54	*p* < .001

s = seconds.

between mean levels of participants' time estimations for the two methods seemed to be inadequate for discussing the degree of consistency. The participants' time estimations for target duration condition were standardized with a mean of 0 and standard deviation of 1 because the Kolmogorov–Smirnov tests for each variable did not result in normally distributed data ($p < .05$ for all variables). The correlations between participants' time judgments in the former and latter block were .574 ($p < .001$) and .503 ($p < .001$), for RP and VE respectively.

Next, correlations between time estimation obtained by RP, VE, and the other dependent variables (performances in the recognition test, expectation) were estimated using Pearson product-moment correlations. As with the consistency analysis, correlation coefficients between estimation duration and the other dependent variables were conducted not on the mean of the dependent variables but on each all data ($n = 262$). These dependent variables were also standardized with a mean of 0 and standard deviation of 1 because the Kolmogorov–Smirnov tests for each variable did not result in normally distributed data ($p < .05$ for all variables). Correlation analyses showed a significant negative relationship for RP and performance in the recognition test ($r = -.226$, $p < .01$), and significant positive relationship for RP and expectation ($r = .301$, $p < .01$), and VE and expectation ($r = .552$, $p < .001$). A correlation analysis was then employed to investigate the association between participants' estimations obtained by RP and VE, using Pearson product-moment correlations. This analysis showed a positive significant relationship ($r = .186$, $p < .001$). Table

Table 3.   Correlation coefficients for each dependent variable ($n = 262$).

	Verbal estimation	Performance on recognition test	Expectation
Reproduction	.186[*]	.-226*	.301*
Verbal estimation		.006	.552*

The asterisk (*) indicates that the coefficients are statistically different from zero at the 0.1 level.

3 shows the results of these correlation analyses.

Discussion

There were two primary aims of the present study. First, to examine not only the accuracy and consistency, but also the stability, of two methods of measuring time estimation in a single experimental condition. Second, to characterize the relationship between time estimations obtained by RP or VE and memory processes.

The CV, which is an index of stability, was significantly higher in the VE condition than in the RP condition, indicating that RP has a higher degree of stability than VE. To our knowledge, this study is the first to examine the stability of measuring methods for time estimation. The difference in stability may be attributed to the fact that time estimation tends to be given as a nice round number under VE, whereas in RP the participants try to reproduce the length of time as accurately as possible. Furthermore, the key pressing adds to this process, and participants could attend to the passage of time and compare their own time representation with that during reproduction. This checking process may lead to the higher stability of RP. On the other hand, in VE, the participants had to assign a number to the length of their time representation, which they subsequently had to verbalize. Such a verbalization process could lead to the tendency to giving a round number. This tendency could increase the range of the verbally estimated duration. For example, if the participant has a 33 s time representation, their verbally estimated duration is likely to be 25 s, 30 s, 35 s, or 40 s. Therefore, VE would have a lower degree of stability than RP.

With respect to accuracy, significant underestimations were observed in all conditions and there was no significant difference in the means between the RP and VE conditions. These results indicate that the two methods produce almost a similar pattern of time estimation, suggesting that these methods have almost the same accuracy. This finding is consistent with Zakay (1993), who found a similar pattern of results with RP and VE. However, it contrasts with the finding that VE durations were significantly longer than RP and target duration (Brown, 1985; Hornstein & Rotter, 1969; McConchie & Rutschmann, 1971; Wallace & Rabin, 1960). Hence, our hypothesis that time estimation obtained by RP is more accurate than that obtained by VE was not supported. It is possible to interpret this result as the effect of participants devoting their attention to the nontemporal task (memorizing nonsense words). Several previous studies have demonstrated that attention strongly influences time estimation regardless of the measuring method (Brown, 1985; Macar et al., 1994; Zakay, 1993). According to the attentional approach, subjective duration would be negatively correlated with the amount of attention given to nontemporal information processing (e.g., Zakay, 1993). For example, Macar et al. (1994) demonstrated that the ratio of subjective duration to target duration decreased as a larger proportion of attention was attributed to the nontemporal task. In this study, participants devoted a very large proportion of attention to memorizing the nonsense words, which is nontemporal information (see Table 1). For this reason, subjective

duration would be shorter than the target durations in both the RP and VE conditions.

The question arises of why the attentional effect would decrease the degree of accuracy, but not stability? Since the attentional effects on time estimation would occur at the encoding stage, attention is assumed to affect the perceived duration directly (Brown, 1985; Macar et al., 1994; Zakay, 1993). Therefore, since perceived duration is connected directly with the degree of accuracy, we consider that attention plays an important role for the degree of accuracy, but not stability. Our data suggested that devoting attention to the nonsense words would lead to shorter perceived duration. On the other hand, the measuring method is assumed to affect the expression of the time representation rather than the perceived duration. A typical example, as mentioned above, is that VE has the tendency to be a nice round number. Thus, we assume that attention would influence accuracy but not stability.

In terms of the consistency of the measuring methods analysis, we demonstrated that the correlations of estimated duration between the former and latter experimental blocks were significantly high for both measuring conditions. This result indicates that RP and VE have almost the same degree of consistency. The correlation coefficient for RP, .57, is within the range of .39 to .58 that was found in a previous study (Siegman, 1962). In contrast, the consistency of the correlation coefficient for VE, .50, was lower than that found in previous studies. For example, the consistency found by Siegman (1962) ranged from .82 to .84. This consistency difference in VE between the present and previous study may be due to the previous study using a target duration of less than 20 s. In general, it is known that the degree of variation in time estimation increases as the length of target duration increases. The main reason that the previous study found a higher consistency correlation coefficient for VE would be that the shorter target durations were estimated to be more invariable than longer target durations. However, we obtained the same consistency for RP as the previous study that used short target durations. Our data showed that the SD was higher with longer target durations in both RP and VE, although the SD differences between the target durations were smaller in the RP condition than in the VE condition (see Figure 2). This finding suggests that it is possible that the SD of the RP slightly increases with the length of target duration, which might cause no difference in consistency between the present and previous study.

We will now discuss the correlation analysis. First, we found a significant positive relationship between time estimation obtained by both methods and expectation about when the target duration would end. This result indicates that expectation is related to the estimation duration obtained by both methods. The expectation/contrast model of judged duration (Jones & Boltz, 1989) may account for this result. That is, as the target duration seems to end earlier than expected, the perceived duration would shorten; in contrast, as the target duration seems to end later than expected, the perceived duration would lengthen.

Second, performance on the recognition test was correlated significantly with RP, but was not correlated with VE. A possible explanation for this finding is that RP is susceptible to the effects of memory processing whereas VE may be independent of it. At the beginning of the reproduction, the participants would consolidate the duration during the memory task. It may

be that participants automatically used something stored in their memory, such as the number of the presented nonsense word, as cues for reproduction. However, in VE, expectation was shown to have a stronger relationship with estimation time than performance on the memory task, which suggests that expectation/contrast about target duration work as a cue in VE.

Third, a significant positive relationship was observed between RP and VE, which replicates previous research, suggesting that there is a common intra-individual basis for different methods of time estimation (Du Preez, 1963). However, it is noteworthy that the correlation coefficient of .186 does not indicate a strong association. This weak association between RP and VE is probably due, at least in part, to the stability or variability difference between RP and VE. As we showed RP to have a higher stability than VE, it is predicted that time estimation is more variable in the VE condition than in the RP condition. We considered that this stability difference between RP and VE may be responsible for the weak association between these methods. However, we could not determine the precise mechanisms underlying the weak association between RP and VE.

It is noted that the findings of the present study are limited under the dual task paradigm. It is possible that by using a single task paradigm (only the time estimation task), the effects of the measuring method on time estimation would be altered because participants may devote their attention to the passage of time. Previous studies that have shown that VE involves an overall tendency for overestimation all used the single task paradigm (Clausen, 1950; Hornstein & Rotter, 1969; McConchie & Rutschmann, 1971; Ochberg et al., 1965). There remains the possibility that the accuracy difference between the two methods may be observed only when a single task paradigm is used (i.e., all of the participant's attention is directed to the passage of time). Therefore, further research using a single task paradigm is necessary to examine the accuracy difference between the two methods, and to expand our understanding of the effects of measuring method on time estimation.

In conclusion, our data showed that RP and VE have equal accuracy and consistency, but RP has a higher degree of stability than VE. Therefore, we can conclude that RP is superior to VE in stability. Moreover, we found that RP was correlated with memory task performance, whereas VE was not. These findings suggest that RP is the better method for time estimation and has a closer relation to memory processes than the VE method.

References

Allan, L. G. (1979). The perception of time. *Perception & Psychophysics*, **26**, 340-354.

Baddeley, A. (1990). Human memory, theory and practice. Boston: Allyn & Bacon.

Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review*, **4**, 184-197.

Brown, S. W. (1985). Time perception and attention: The effects of prospective versus retrospective paradigms and task demands on perceived duration. *Perception & Psychophysics*, **38**, 115-124.

Clausen, J. (1950). An evaluation of experimental methods of time judgment. *Journal of Experimental Psychology*, **40**, 756-761.

Danziger, K., & Du Preez, P. D. (1963). Reliability of time estimation by the method of reproduction. *Perceptual & Motor Skills*, **16**, 879-884.

Du Preez, P. D. (1963). Relation between verbal estimation and reproduction of a short time interval: Preliminary study. *Perceptual & Motor Skills,* **17**, 45-46.

Gamache, P. L., & Grondin, S. (2010). Sensory-specific clock components and memory mechanisms: investigation with parallel timing. *Europian Journal of Neuroscience*, **31**, 1908-1914.

Grondin, S. (2008). Methods for studying psychological time. In S. Grondin (Ed.), *Psychology of time*, 51-74. Bingley, U.K.: Emerald Group.

Hawkes, G. R., Bailey, R. W., & Warm, J. S. (1960). Method and modality in judgments of brief stimulus duration. *Journal of Auditory Research*, **1**, 133-144.

Hicks, R. E., Miller, G. W., & Kinsbourne, M. (1976). Prospective and retrospective judgments of time as a function of amount of information processed. *The American Journal of Psychology*, **89**, 719-730.

Hornstein, A. D., & Rotter, G. S. (1969). Research methodology in temporal perception. *Journal of Experimental Psychology*, **79**, 561-564.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, **96**, 459-491.

Kruup, K. (1961). The influence of method on time judgments. *Australian Journal of Psychology,* **13**, 44-54.

Macar, F., Grondin, S., & Casini, L. (1994). Controlled attention sharing influences time estimation. *Memory & Cognition*. **22**, 673-686.

McConchie, R. D., & Rutschmann, J. (1971). Human time estimation: On differences between methods. *Perceptual & Motor Skills*, **32**, 319-336.

Ochberg, F. M., Pollack, I. W., & Meyer, E. (1965). Reproduction and estimation methods of time judgment. *Perceptual & Motor Skills,* **20**, 653-656.

Ornstein, R. E. (1969). *On the experience of time.* London : Penguin Books Ltd.

Siegman, A. W. (1962). Intercorrelation of some measures of time estimation. *Perceptual & Motor Skills*, **14**, 381-382.

Wallace, M., & Rabin, A. I. (1960). Temporal experience. *Psychological Bulletin*, **57**, 213-233.

Zakay, D. (1993) .Time estimation methods – do they influence prospective duration estimation. *Perception*, **22**, 91-101.

Zakay, D., & Block, R. A. (1997). Temporal cognition. *Current Directions in Psychological Science,* **6**, 12-16.