

Western University

Scholarship@Western

Communication Sciences and Disorders
Publications

Communication Sciences and Disorders School

6-1-2017

Testing the accuracy of timing reports in visual timing tasks with a consumer-grade digital camera

Rachael E. Smyth

The University of Western Ontario

Janis Oram Cardy

The University of Western Ontario

David Purcell

The University of Western Ontario

Follow this and additional works at: <https://ir.lib.uwo.ca/scsdpub>

Citation of this paper:

Smyth, Rachael E.; Oram Cardy, Janis; and Purcell, David, "Testing the accuracy of timing reports in visual timing tasks with a consumer-grade digital camera" (2017). *Communication Sciences and Disorders Publications*. 22.

<https://ir.lib.uwo.ca/scsdpub/22>

Testing the accuracy of timing reports in visual timing tasks with a consumer-grade digital camera

Rachael E. Smyth^{1,2} · Janis Oram Cardy^{2,3} · David Purcell^{2,3}

Published online: 20 June 2016
© Psychonomic Society, Inc. 2016

Abstract This study tested the accuracy of a visual timing task using a readily available and relatively inexpensive consumer grade digital camera. A visual inspection time task was recorded using short high-speed video clips and the timing as reported by the task's program was compared to the timing as recorded in the video clips. Discrepancies in these two timing reports were investigated further and based on display refresh rate, a decision was made whether the discrepancy was large enough to affect the results as reported by the task. In this particular study, the errors in timing were not large enough to impact the results of the study. The procedure presented in this article offers an alternative method for performing a timing test, which uses readily available hardware and can be used to test the timing in any software program on any operating system and display.

Keywords Visual timing accuracy · Visual timing errors · Visual timing test · Consumer grade hardware

Introduction

This article describes an independent way of testing the accuracy of timing that is reported by computer programs used to

carry out visual timing tasks with any type of computer display. The example timing accuracy test described in this study was performed using The Benny Bee Inspection Time task (Williams, Turley, Nettelbeck, & Burns, 2009), which is run in Neurobehavioral Systems stimulus-response software, Presentation (<http://www.neurobs.com/>).

Vickers and Smith (1986, p. 609) defined inspection time (IT) as “the time required by a subject to make a single observation or inspection of the sensory input on which a discrimination of relative magnitude is based.” IT thresholds are calculated using the shortest amount of time the participant requires to make a correct decision about the stimuli in a specified percentage of trials. The percentage of correct trials required varies depending on the specific paradigm being used. For example, in the Benny Bee IT task, Benny the Bee appears on one of two identical flowers and, after a certain amount of time, eight identical bees appear on each flower, acting as a mask. The mask is an identical flower and bee image that overlays the stimulus and prevents further processing of the target stimuli (Nettelbeck & Young, 1989; Williams et al., 2009). In the case of the Benny Bee IT task, the threshold used in calculating inspection time is the point at which the participant is 79 % correct in making a decision (Williams et al., 2009). Understandably, the precision and accuracy of the timing in these paradigms are crucial to the accurate calculation of the participant's IT threshold.

The Benny Bee IT task uses a two alternative forced choice, adaptive staircase procedure (Wetherill & Levitt, 1965; Williams et al., 2009). This procedure changes the time between trials by a certain period based on whether the previous response was correct or incorrect. The period is determined by the display refresh rate that is estimated by the program running in Presentation. To date, the visual timing properties of various display programs have been tested for error rates and some solutions have been offered.

✉ Rachael E. Smyth
rsmyth5@uwo.ca

¹ Health and Rehabilitation Sciences, Western University, London, ON, Canada N6A 3K7

² Communication Sciences and Disorders, Western University, London, ON, Canada

³ National Centre for Audiology, Western University, London, ON, Canada

Garaizar, Vadillo, López-de-Ipiña and Matute (2014) performed timing tests using three popular software platforms to determine which software platforms have the most accurate timing properties in visual computer tasks. They tested the timing of a commercially sold software (EPrime), a free software (DMDX), and a free, multiplatform software (PsychoPy). The test employed by Garaizar et al. (2014) used the Black Box Tool Kit (*BBTK*, Plant, Hammond, & Turner, 2004; Plant & Turner, 2009) to measure timing inaccuracies. The *BBTK* is a combination of software and hardware, namely the external stimulus sensors and response device, that uses a host computer and opto-detectors and/or microphones to compare the actual timing of visual and/or auditory stimuli against the timing that is requested by the program. Software programs record the timing of when stimuli are reported to have occurred, based on the specific parameters of the paradigm, whereas hardware-timing measures use external apparatus to record paradigms, which are then analyzed to measure timing. A photodiode apparatus operated by a program running on a dual-processor Athlon computer has also been used to test the timing accuracy of the aforementioned software programs (Garaizar et al., 2014; Stewart, 2006). Garaizar et al. (2014) found that there were certain times at which errors were more likely to occur within a test session, such as right before the monitor completed its vertical synchronization. They reported that errors in timing were related to the timing technology in the displays and the operating systems, but not necessarily related to the timing in the experimental paradigm software. The Garaizar et al. study supports testing the accuracy of timing in visual tasks because errors were related to the timing technology in the displays and the operating systems.

Schmidt (2001) investigated accuracy in the presentation of visual stimuli by internet-based psychophysics programs using a photo detector system. Internet-based programs perform with more variability than paradigms using local computer software because software and hardware are not as well controlled when remote software runs the experiment. Whereas local computer software can access the operating system and hardware relatively directly, internet experiments generally operate through mediating software (like web browsers) and are subject to effects from the Internet, such as random delays in information transmission (Schmidt, 2001). These factors make it more difficult to exert consistent control over the timing of the visual stimuli. Although Schmidt's (2001) main interest was to investigate the accuracy of internet-based experiments compared to experiments implemented using local software, the author reported that the accuracy of all but one of the programs tested was affected by the speed of the system used in the test. These results suggest that it is important to perform a timing test using the specific system that will be used for the experiment because the accuracy changes with different systems.

Plant and Quinlan (2013) have reviewed possible sources of timing errors in the hardware and software used in

neuroscience studies. While various studies have tested the timing accuracy using different combinations of hardware and software, the potential for timing errors to occur changes with each experimental procedure run. Conditions likely vary between studies, and may even vary within studies. The risk of these types of errors is that studies can be difficult to replicate. Different conditions lead to different timing errors and this can lead to inconsistencies in results (Plant & Quinlan, 2013). It is crucial that the accuracy of a study's timing is tested to ensure that the timing errors occurring are not creating differences so large that they could affect the results of a study.

As evidenced by the previous studies, the literature clearly identifies the need to test the accuracy of timing reports in computer-based tasks. As reviewed above, there are different testing methods available that generally use some type of photo/opto detector system (Garaizar et al., 2014; Plant & Turner, 2009; Schmidt, 2001). While photo/opto detector systems may be most commonly used, they are expensive. Photo/opto detector systems may not be feasible for research programs whose focus is not vision research, but who require the accuracy of a visual task's timing mechanism as an extension of another study.

The aims of the present study were to: (1) test the accuracy of the timing mechanism in a visual task by comparing the timing requested by the task to the actual timing that occurred, using readily available and inexpensive consumer hardware, such as a camera and monitor, and (2) present a step-by-step process that may be used when conducting a generalized timing accuracy test that may be used to test the timing accuracy of a visual computer task using any display program and monitor combination.

Method

Materials

We used a Lenovo T440 computer with an Intel® Core™ i7-4600U CPU @ 2.10 GHz 2.70 GHz running the 64-bit Windows 7 Professional operating system to perform the timing tests. The CRT display that was attached to the computer was an Elo Touchsystems monitor, model ET1725C-4UWE-3 (100-240 V, 1.5A, 60/50 Hz, P/N 454000-000). A consumer grade Fujifilm FinePix F550EXR camera was used to record high-speed video at 320 frames per second (fps) while mounted on a tripod with the display frame centered on the CRT display. Default settings for 320 fps were used and individual frames were 320 × 240 pixels. We felt it was prudent to verify this consumer grade camera operated at the 320 fps specified in its documentation. A test circuit was constructed to flash a white light emitting diode (LED) with a known pattern so that the number of camera frames showing the periodically lit LED could be counted. The circuit shown in Fig. 1 lights the white signal LED for any input voltage greater than approximately 200 mV with switching times faster

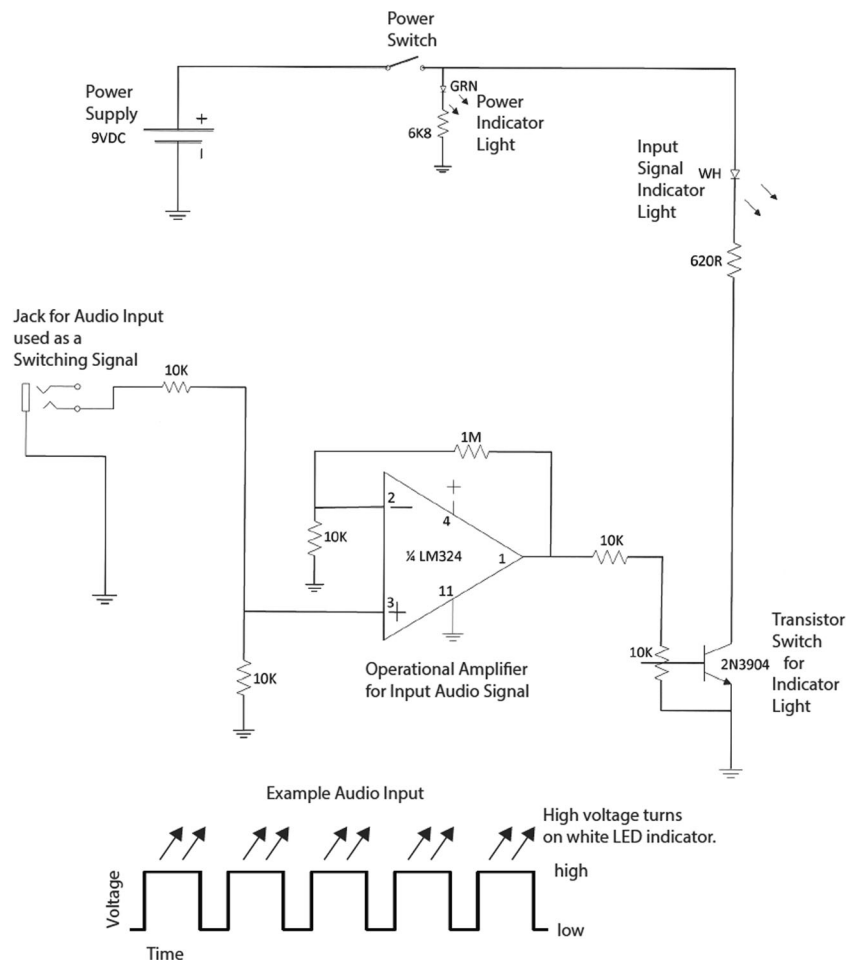


Fig. 1 Circuit (top) and audio to white-light test signal (bottom) used to evaluate frame rate of the high-speed video camera

than the camera frame rate of $1/320 = 3.125$ ms. Our test signal was generated by a computer sound card using MATLAB (Mathworks, MA, USA) such that the LED would be periodically lit for 20 frames (62.5 ms) and unlit for 10 frames (31.25 ms) if the camera operated as documented. The camera saves video as a “.mov” file that was first transferred to computer and then individual sequential frames were extracted using the program Zerano FFmpeg (version 2014-03- 20 git-19139d8, retrieved from <http://ffmpeg.zerano.com/builds/>). After viewing and counting lit and unlit frames, we concluded the camera frame rate was accurately 320 fps. The same manner of extracting and reviewing frames was also used to verify the timing of the Benny Bee IT task. The Benny Bee IT task was run using Neurobehavioral Systems Presentation (<http://www.neurobs.com/>) software, version 19.2.

Procedure

The Benny Bee IT task estimates the refresh rate prior to the start of each test session. In the Benny Bee IT task, the time between Benny the Bee appearing and his friends appearing changes with each trial. The minimum size of the change

between trials is equal to the period of the refresh rate. Prior to using this task in a study, we wanted to evaluate whether the times reported between Benny the Bee arriving and his friends arriving were accurate.

First, the camera was tested to ensure it was recording at a rate of 320 fps, as is described in the **Materials** section above. The camera was then set up using a tripod to record the Benny Bee Inspection Time task on the CRT monitor in videos of 30 s each. These videos were used to analyze the timing. The camera recorded 320 fps with a period of $[(1/320)*1,000=]$ 3.125 ms/frame. The monitor and Presentation software were set to the same refresh rate of 85 Hz. The refresh period of the monitor and software was $[(1/85)*1,000=]$ 11.76 ms.

To measure timing errors, the time between Benny the Bee appearing on the screen and the mask appearing on the screen as requested by the program was compared to the actual time between Benny appearing on the screen and the mask appearing on the screen by analyzing the high-speed video recordings of the task. Both the frame number in which Benny appeared and the frame in which his friends appeared were recorded for each trial. The difference in frame number between these two frames was multiplied by the camera’s frame period (3.125 ms) to give the

actual time between Benny arriving and his friends arriving. The actual time was then compared to the time difference requested by the program. We calculated the difference between the actual time and the requested time by subtracting the actual time measured from the time requested by the program. If the difference between the actual time and the requested time was larger than one monitor period, it was considered an error that likely impacted the results of that trial.

Results

The Benny Bee IT program's timing test estimated a refresh period of 11.76 ms by averaging the software reported delays of 100 test stimuli. This period was the estimated time required to complete one vertical refresh of the monitor and is the same as the actual refresh period rounded to two decimal places. In a test of 59 trials, 48 were recorded on video. Eleven trials were not recorded due to the 30-s maximum video length. Of these 48 recorded trials, there were no errors larger than 4 ms in this test. This suggests that for the hardware and software employed here, the Benny Bee IT paradigm can run with a high level of accuracy. Errors in timing were all smaller than the refresh rate (11.76 ms), as can be seen in Fig. 2 and, therefore, did not affect the accuracy of presentation.

Discussion

This study demonstrated that the Benny Bee IT paradigm run in Presentation on our Lenovo notebook computer connected to the Elo CRT monitor did not contain timing errors that affected the results. As described by Plant and Quinlan (2013), it is important to perform independent timing tests when using different hardware/software combinations to ensure timing errors are not creating differences in study outcomes. We verified the accuracy of timing reports independently using a combination of hardware and software (i.e., Fujifilm FinePix camera, the test circuit, and

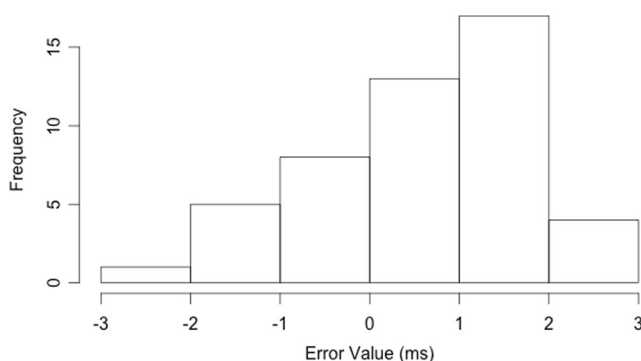


Fig. 2 Distribution of timing error values. This histogram shows the frequency of timing errors that occurred during the test. The horizontal axis shows errors with units of ms

Zerano FFmpeg software). This study describes an independent procedure that can be used to test the accuracy of visual timing paradigms. It can be used as a test of accuracy for paradigms run in any software program on any operating system using any type of display.

The CRT monitor was selected for a number of reasons. Traditionally, Liquid Crystal Displays (LCDs) have been problematic in their timing, although progress has been made. Certain LCD models perform comparably to CRTs on various aspects of timing. There are also certain characteristics, which are preferable using one type of monitor or the other (Elze & Tanner, 2012; Wang & Nikolić, 2011). In this study, a CRT was selected because of the historical timing challenges with LCD displays. This model of CRT monitor was selected because it represents a monitor that was manufactured near the most mature and technologically advanced period in the history of CRT popularity. The availability of CRT monitors may be an ongoing concern in this line of research, but there may still be lightly used monitors from the end of CRT popularity that are available due to the decreasing demand for CRT monitors.

This approach is particularly useful for researchers whose program of research is not directly related to vision research, but who have use for testing the accuracy of timing reports in a visual task. The equipment used in this method is readily available to most and, financially, the equipment is less of an investment. Moving forward, this method of testing the accuracy of visual timing tests should be applied to more hardware/software combinations. Using this method to test the accuracy of timing reports using a variety of monitors, hard drives, programs to run the stimuli, and paradigms would provide a larger base of support for its use in neuroscience testing. While this method offers an accessible and inexpensive alternative to previously reported timing tests, it is manually intensive. It would be improved by developing a more automatic method of analyzing the timing reports.

This study offers a readily available, inexpensive, consumer hardware-based alternative to previously reported timing tests that employ photo detectors or other less accessible hardware. The method can be used to test the timing accuracy using any monitor, hard drive, program, and paradigm combination.

Acknowledgments We would like to thank Dr. Andrew Johnson for his feedback on the manuscript. We would also like to thank David Grainger at the National Centre for Audiology for his assistance with the timing circuit shown in Fig. 1.

References

- Elze, T., & Tanner, T. G. (2012). Temporal Properties of Liquid Crystal Displays: Implications for Vision Science Experiments. *PLoS ONE*, 7(9). doi:10.1371/journal.pone.0044048
- Garaizar, P., Vadillo, M. A., López-de-Ipiña, D., & Matute, H. (2014). Measuring software timing errors in the presentation of visual stimuli

- in cognitive neuroscience experiments. *PLoS One*, 9(1), e85108. doi:10.1371/journal.pone.0085108
- Nettelbeck, T., & Young, R. (1989). Inspection time and intelligence in 6-year-old children. *Personality and Individual Differences*, 10, 605–614.
- Plant, R. R., Hammond, N., & Turner, G. (2004). Self-validating presentation and response timing in cognitive paradigms: How and why? *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc*, 36(2), 291–303. doi:10.3758/BF03195575
- Plant, R. R., & Quinlan, P. T. (2013). Could millisecond timing errors in commonly used equipment be a cause of replication failure in some neuroscience studies? *Cognitive, Affective & Behavioral Neuroscience*, 13, 598–614. doi:10.3758/s13415-013-0166-6
- Plant, R. R., & Turner, G. (2009). Millisecond precision psychological research in a world of commodity computers: New hardware, new problems? *Behavior Research Methods*, 41(3), 598–614. doi:10.3758/BRM.41.3.598
- Schmidt, W. C. (2001). Presentation accuracy of Web animation methods. *Behavior Research Methods, Instruments, & Computers*, 33(2), 187–200. doi:10.3758/BF03195365
- Stewart, N. (2006). Millisecond accuracy video display using OpenGL under Linux. *Behavior Research Methods*, 38(1), 142–145.
- Vickers, D., & Smith, P. L. (1986). The rationale for the inspection time index. *Personality and Individual Differences*, 7(5), 609–623.
- Wang, P., & Nikolić, D. (2011). An LCD monitor with sufficiently precise timing for research in vision. *Frontiers in Human Neuroscience*, 5(85), 1–10. doi:10.3389/fnhum.2011.00085
- Wetherill, G. B., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *The British Journal of Mathematical and Statistical Psychology*, 18(1), 1–10.
- Williams, S. E., Turley, C., Nettelbeck, T., & Burns, N. R. (2009). A measure of inspection time in 4-year-old children: The Benny Bee IT task. *British Journal of Developmental Psychology*, 27(3), 669–680. doi:10.1348/026151008X354573