Amsterdam University Press

COMPUTATIONAL COMMUNICATION RESEARCH 3.2 (2021) 133-151
HTTPS://DOI.ORG/10.5117/CCR2021.2.001.CALC

The Accuracy and Precision of Measurement

Tools for Validating Reaction Time Stimuli

Leandro Calcagnotto
Independent Researcher

Richard Huskey
Cognitive Communication Science Lab, Department of Communication, University of California Davis. Center for Mind and Brain, University of California Davis
rwhuskey@ucdavis.edu

Gerald M. Kosicki
School of Communication, The Ohio State University

Abstract

Measurement noise differs by instrument and limits the validity and reliability of findings. Researchers collecting reaction time data introduce noise in the form of response time latency from hardware and software, even when collecting data on standardized computer-based experimental equipment. Reaction time is a measure with broad application for studying cognitive processing in communication research that is vulnerable to response latency noise. In this study, we utilized an Arduino microcontroller to generate a ground truth value of average response time latency in *Asteroid Impact*, an open source, naturalistic, experimental video game stimulus. We tested if response time latency differed across computer operating system, software, and trial modality. Here we show that reaction time measurements collected using *Asteroid Impact* were susceptible to response latency variability on par with other response-latency measuring software tests. These results demonstrate that *Asteroid Impact* is a valid and reliable stimulus for measuring reaction time data. Moreover, we provide researchers with a low-cost and open-source tool for evaluating response time latency in their own labs. Our results highlight the importance of validating measurement tools and support the philosophy of contributing methodological improvements in communication science.

Keywords: reaction times, measurement accuracy, validation, open science

The Accuracy and Precision of Measurement: Tools for Validating Reaction Time Stimuli

A limitation encountered when conducting empirical research is related to the accuracy and precision of measurement. The noise associated with any measure is likely to differ by instrument and acts as a bottleneck for the level of depth and sophistication of scientific inquiry. The validity and reliability of an instrument's measure also constrains the insights and contributions capable of being drawn from it. Thus, in order to effectively test, falsify, and contribute to communication theory, valid and reliable measurement is required.

A related concern focuses on researcher's ability to extract meaningful signals from noise. Controlled experiments that use computer-based instruments are vulnerable to noise from various sources within their hardware and software. This noise can lead to inaccurate data and conclusions, which may result in inconsistent research findings. Such inconsistencies can potentially undermine replication efforts, and contribute to false-positive (or negative) findings in empirical science (Plant, 2016). One way of correcting for the potential noise in a measure is to validate an instrument's measurement accuracy.

Reaction times are commonly measured in communication research, and these measurements are subject to the concerns outlined above. Therefore, softwares used to measure reaction times must be carefully validated in order to ensure accurate results. A reaction time measurement tool has been recently implemented in *Asteroid Impact* (Huskey et al., 2018), an open source, naturalistic, experimental video game stimulus. In line with previous research (Neath et al., 2011; Schubert et al., 2013), this study used specialized hardware to establish a ground truth of measurement error for the reaction time measurement. In this brief report, we briefly discuss the importance and widespread use of reaction time methodology, highlight current issues in using reaction time measures, and present a tool to assess response time latency. We then apply this hardware to test reaction time accuracy in *Asteroid Impact* while systematically varying operating systems (macOS, Windows 10, Ubuntu 18.04), Python versions (Python 2, Python 3), and trial modalities (auditory, visual). Results show that each of these factors contributes to measurement accuracy. The results highlight the importance of validating measurement tools and are discussed in terms of how methodological improvements contribute to communication science.

Reaction Time Measures in Communication

Reaction time measures in communication have contributed to a vast breadth of research findings across various topics. Lang and colleagues (e.g., Lang et al., 2006; Lang & Basil, 1998) have conducted path breaking research to validate reaction times as a measure of cognitive processing during media use. These measures have been critical to testing theories on cognitive resource availability, including Lang's Limited Capacity Model of Motivated Mediated Message Processing (LC4MP; Lang, 2000, 2006b, 2017). To date, more than 50 studies (Huskey et al., 2020) have used reaction times to test various components of the LC4MP (for a review, see: Fisher, Huskey, et al., 2018; Fisher, Keene, et al., 2018).

Reaction times have also been used to measure levels of aggression after exposure to violent media (e.g., Anderson & Dill, 2000; but see: Ferguson et al., 2008), attentional dynamics during naturalistic video game play (Weber et al., 2018), perceptual and cognitive load during media use (Fisher et al., 2019), moral judgments when evaluating media characters (Matthews, 2019), cognitive elaboration during persuasive message processing (Wilcox et al., 2020), and attentional resource allocation during flow experiences (Huskey et al., 2018). Another line of research uses reaction times to study defensive message processing and message avoidance (Clayton et al., 2020; Clayton & Leshner, 2015; Liu & Bailey, 2019). Reaction times have a long history in memory research (Sternberg, 1969), and communication scholars have measured reaction time during memory tasks such as signal detection (Miller & Leshner, 2007). More recently, media researchers are using computational modeling and decision theory (Fisher & Hamilton, 2021) to study media selection (Gong, Huskey, Eden, & Ulusoy, 2021). Even in observational social science research such as random population surveys, timers are critical to monitor the steady progress of the respondents throughout the activity (Bassili & Fletcher, 1991; Lavrakas, et al., 2019). This brief (and certainly not exhaustive) summary of research showcases the versatility and expanse of reaction time measures as utilized in communication science. In fact, modern efforts increasingly situate reaction times in naturalistic behavioral tasks, such as watching media stimuli or playing a video game (Lang, 2006a; Mathiak & Weber, 2006).

Using Asteroid Impact to Measure Reaction Times

Asteroid Impact (Huskey et al., 2018) is an experimental, open source, and naturalistic video game stimulus for communication scientists. The software uses an embedded reaction time tool to capture secondary task reaction time measures (STRT; Lang et al., 2006; Lang & Basil, 1998). Briefly, STRTs are reaction times that are collected when participants are simultaneously engaged in a primary and secondary task (Lang, 2009). In *Asteroid Impact*, a participant's primary task is to collect targets (crystals) while preventing their spacecraft (mouse cursor) from being hit by flying asteroids. The secondary task is, by comparison, to press a keyboard key in response to an auditory or visual stimulus. The time between the onset of the stimulus and the participant's response constitutes the STRT measure. Theoretically, STRT measures resources available during message processing (Lang et al., 2006; Lang & Basil, 1998). However, response latency within computer-based measuring equipment potentially confounds the accuracy and precision of the STRT measure.

Sources of Response Latency Noise

Response latency remains a universal and pervasive problem in computer-based experiments. The widespread use of standardized experimental hardware and software overshadows its highly variable nature (Plant, 2016). Both computer hardware and software can contribute response latency noise, and random, uncontrolled response time latency is problematic because it constrains the precision of the measures being taken. The resolution for this problem heeds what Plant (2016) has recommended for all researchers: to self-validate computer-based experimental measuring instruments.

Microcontrollers as Validation Tools

In light of the potential issues with using computer-based measuring equipment, one fundamental question emerges: How can researchers account for response latency noise associated with their measuring tools? The answer can be found by looking to a specialized piece of hardware known as an Arduino (pronounced Ar du ween Oh) microcontroller.

The Arduino microcontroller (or simply 'Arduino') is an electronic computer device built with hardware and software from Arduino, a global open-source project based at the Interaction Design Institute Ivrea, in Italy. Arduinos contain a microprocessor capable of performing basic computing functions such as sending and receiving electrical inputs. The utility of the Arduino stems from its open-source programming that allows users to tailor its functionality to their specific needs. Arduinos are often combined with external sensory-modality equipment such as photoresistors, microphones, LEDs, and buzzers. A unique feature of the Arduino stems from its reliable and near-instantaneous processing speed that makes it an ideal tool to test for response latencies in computer equipment. In fact, previous research shows that Arduino boards can even be used to track and test the response latencies of equipment by emulating a keyboard or mouse (Neath et al., 2011; Schubert et al., 2013).

The Current Investigation

Asteroid Impact has been used as an experimental platform to support research in communication science, but its reaction time measure has not yet been extensively validated. Our study addresses this gap by testing how different software configurations contribute to response time latencies (which can be understood as a source of *measurement error*) in *Asteroid Impact*. To test this question, we used an Arduino Leonardo microcontroller to capture response time latencies in *Asteroid Impact's* reaction time measure across computer operating systems (macOS, Windows 10, Ubuntu 18.04), Python versions (Python 2, Python 3), and trial modalities (auditory, visual). Ideally, these response time latencies will have a small mean and variance, which would indicate that *Asteroid Impact* is accurate in measuring reaction times. We expect these response latencies to be in line with previous research (Neath et al., 2011; Schubert et al., 2013). The results from this study will help to establish a ground truth of measurement error for experimental configurations running *Asteroid Impact*.

Method

Open Science Practices

This study adopts open science practices (Bowman & Keene, 2018; Dienlin et al., 2020; Lewis, 2020) by making the materials, data, and code necessary to reproduce this study available on GitHub (https://github.com/cogcommscience-lab/ai_response_latency). Our stimulus is open source (https://github.com/cogcommscience-lab/asteroid_impact), as is our hardware (https://www.arduino.cc/).

Overview and Design

We evaluated the response time latency of *Asteroid Impact*'s STRT measure across twelve unique experimental configurations. We programmed an Arduino microcontroller to function as an emulated keyboard during data-collection with the following factorial design: 3 (operating system: macOS, Windows 10, Ubuntu 18.04) *x* 2 (Python version: Python 2, Python 3) *x* 2 (modality: auditory, visual). The total response latency measure was defined as the time between reaction time trial (auditory, visual) onset and response from the Arduino. We monitored the consistency of the Arduino's own response latency by programming it to start a timer when it detected a stimulus and stop it after it issued the keypress. This latency was measured at approximately 1/10 of a millisecond. Given that *Asteroid Impact* measures reaction times in milliseconds, any observed latency in reaction time recorded by *Asteroid Impact* reflects measurement error in the hardware and software configuration.

Materials

For this study, we used an Arduino Leonardo Microcontroller (Figure 1). This microcontroller was selected as it is capable of functioning as an emulated USB keyboard. We programmed the Arduino using open-source software written in C/C++ functions using the Arduino Integrated Development Environment (IDE). The Arduino Leonardo has 20 digital input/output pins (12 of these pins can function as analog inputs). We connected the Arduino to an Elegoo MB-102 Breadboard by connecting jumper wires between the Arduino's ground and power pins to the breadboard's power rails. This gave us the capacity to connect multiple measurement devices (Anmbest Microphone Sensor, photoresistor) to the Arduino. The microphone supplied a digital signal (0 when sound amplitude was below a given threshold, 1 when sound amplitude exceeded the threshold) and the photoresistor supplied an analog signal in the form of a continuous numeric value representing moment-by-moment changes in luminance. These sensors were connected to the Arduino's analog inputs. The microphone sensor already contained an integrated circuit board, but the photoresistor did not. Therefore, we inserted a 10k ohm resistor to limit the electrical current distributed to the photoresistor and provide a valid baseline reference point for measuring changes in resistance. A 220k ohm resistor and LED was also connected to the breadboard. The LED was programmed to flash when the Arduino issued a keypress (see below) in response to changes in sound amplitude or luminosity. Therefore, the LED functioned as a visual indicator that the Arduino was working, and assisted in troubleshooting. The complete wiring schematics are available on the project's Github repository.

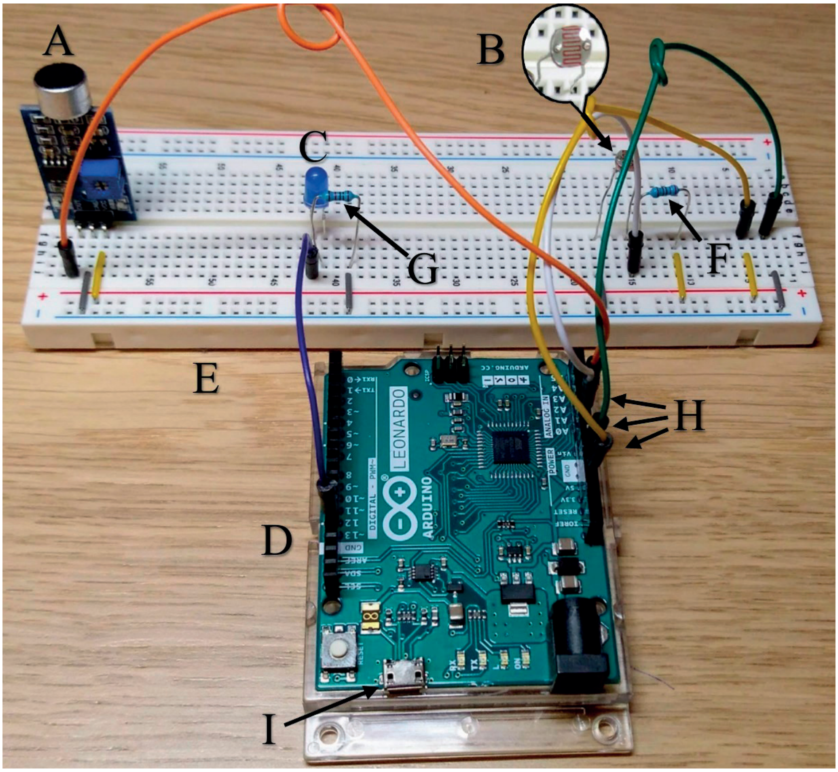

Figure 1 Arduino Microcontroller-Breadboard Circuit. A = Microphone sensor, B = Photoresistor, C = LED, D = Arduino Leonardo, E = Breadboard, F = 10k resistor, G = 220k resistor, H = Jumper wire, I = Micro USB port.

Procedure

The Arduino was programmed to scan the environment every millisecond for changes in either sound amplitude or luminance. During visual trials, the photoresistor was placed within an inch of the laptop display and all other surrounding light sources were turned off. When the photoresistor detected a change that met the luminance threshold, the Arduino issued a keypress response to the laptop. During auditory trials, the microphone sensor was placed within an inch above the laptop's speakers. When the microphone detected a change in sound amplitude that exceeded the decibel threshold, the Arduino issued a keypress response to the laptop. Upon receiving the keypress response, *Asteroid Impact* would close the STRT prompt, delay for 3 seconds, and repeat a new trial. We ran 100 trials for each of the twelve possible configurations (n = 1,200).

Data Analysis

We performed a 3 (OS) *x* 2 (Python version) *x* 2 (modality) ANOVA using the *stats* package in *R* (R Core Team, 2020). Pairwise comparisons were Tukey's honest significant difference corrected to maintain acceptable FWER rates.

Results

ANOVA results showed a significant three-way interaction $F(2, 1188) = 199.95$, $p < .001$, $\eta^2 = .337$ (Figure 2). Auditory trials using Python 3 on macOS had the lowest latency ($M = 44.6$, $SD = 8.4$). Auditory trials using Python 2 on Ubuntu 18.04 had the longest latency ($M = 121.1$, $SD = 18.4$; see Table 1). Raincloud plots of the main effects (all p's < .001) are included to further characterize the distribution of results (Figure 3).

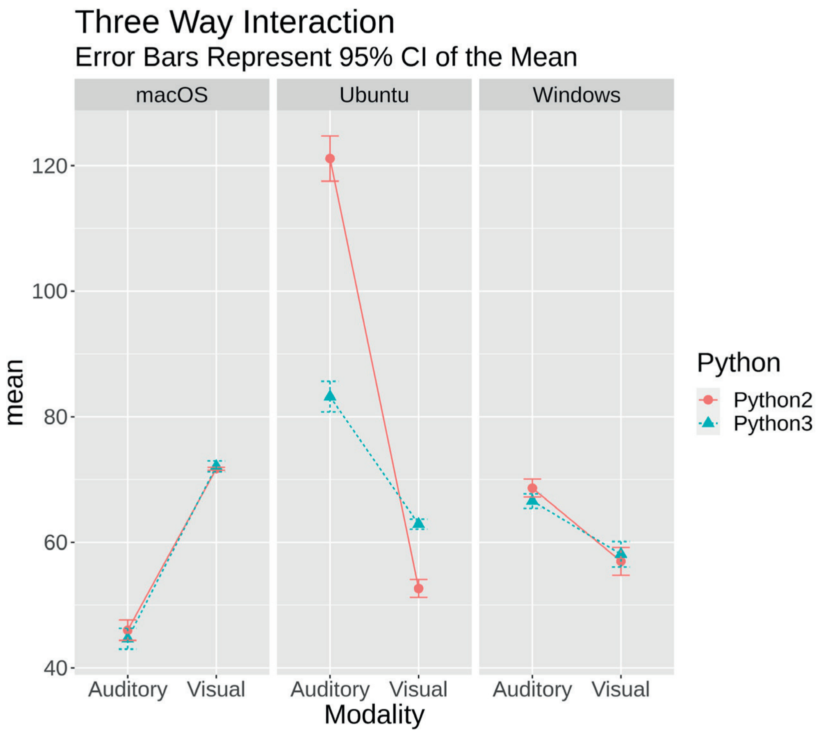

Figure 2 Three-way interaction. Error bars represent the 95% confidence interval of the mean response latency.

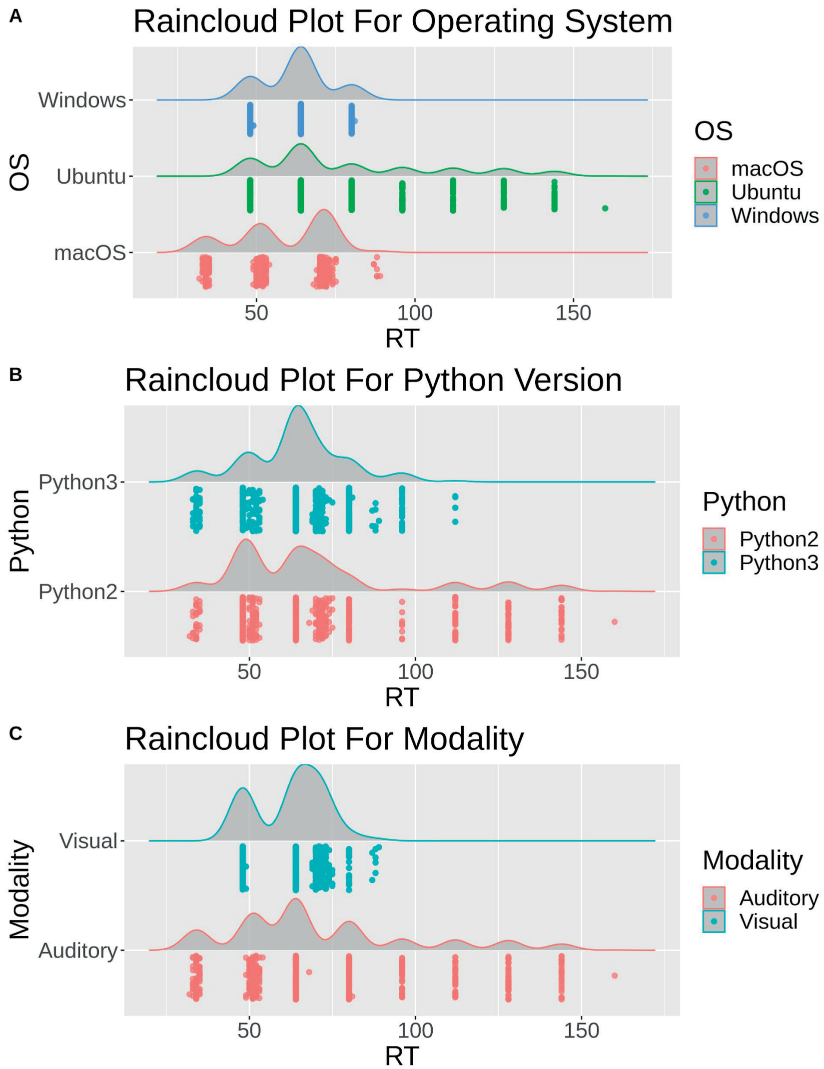

Figure 3 Raincloud plots for main effects. Each point represents an individual response latency measurement.

Table 1 Means, standard deviation, and 95% CI of the mean response latency.

Modality	Operating System	Software	Mean(SD), [95% CI]
Auditory	macOS	Python 2	46.0(8.3), [44.4, 47.6]
		Python 3	44.6(8.4), [43.0, 46.3]
	Ubuntu 18.04	Python 2	121.1(18.4), [117.4, 124.6]
		Python 3	83.2(12.5), [80.8, 85.6]
	Windows 10	Python 2	68.6(7.3), [67.2, 70.0]
		Python 3	66.6(5.9), [65.4, 67.8]
Visual	macOS	Python 2	71.7(1.2), [71.5, 71.9]
		Python 3	72.1(4.47), [71.2, 73.1]
	Ubuntu 18.04	Python 2	56.6(7.3), [55.2, 58.0]
		Python 3	62.9(4.1), [62.1, 63.7]
	Windows 10	Python 2	57.0(11.2), [54.8, 59.2]
		Python 3	58.1(10.3), [56.1, 60.12]

Discussion

We used an Arduino microcontroller to generate a ground truth value of average response time latency for twelve configurations of *Asteroid Impact.* Specifically, we measured the latency between the onset of reaction time trials and a response from the Arduino across operating systems, Python versions, and trial modalities. Our results showed that interactions between the factors had a significant impact on mean response time latency. These findings yield actionable results that are relevant to future researchers. Specifically, if primarily concerned with optimizing performance in *Asteroid Impact* to test auditory responses, then macOS is the optimal choice and Ubuntu 18.04 is the worst, regardless of Python version. If primarily concerned with optimizing performance in *Asteroid Impact* to test visual data, then Ubuntu 18.04 running Python 2 software is the best. However, Windows 10 is not markedly slower than Ubuntu 18.04 in this respect and runs more consistently (across both Python and modality), so it is a more balanced choice. Thus, if selecting a configuration for overall performance, with no preference for either data type, then Windows 10 appears best.

More generally, what exactly do we mean by response latency? When we report a mean response latency, we are reporting the average time in milliseconds between the start and end of a reaction time trial in *Asteroid Impact*. Given that the Arduino microcontroller is capable of responding in 1ms or faster, any delay between the onset and offset of a trial is due to sources of noise within the system's features. Our study attributes variance to three possible sources of noise (operating system, Python version, trial type). A larger mean response latency indicates that a given configuration has greater measurement error. This response latency would exist both before the stimulus appeared on the screen (e.g. latency of display to present image depends on refresh rate) and after the subject initiated a keypress (e.g. latency of AI software to receive and interpret the keypress). The fundamental issue with these errors is that they cloud researcher's interpretations of the data and might mean the difference between attaining statistical significance or not. Of course, this is systematic error that applies equally across experimental conditions so long as the hardware and software remain constant. Nevertheless, the interpretation of our response time latency results can be boiled down to the oxymoron, 'less is more' meaning that small means and standard deviations indicate better performance across the configuration.

Broader Implications

More broadly, our study has important metascientific and theoretical implications, to which we now turn. Our capacity for testing communication theories is constrained by the accuracy of our measurement instruments. In this study, we show that measurement accuracy is contingent on factors related to computer operating system (macOS, Windows 10, Ubuntu 18.04), software version (Python 2, Python 3), and trial modality (visual, auditory). Researchers have the capacity to standardize operating systems and software versions, thereby eliminating these sources of noise. However, trial modality is often used to operationalize the measurement of theoretical variables of interest, such as visual or auditory attention, and it is not unusual to see studies that compare reaction times measured in different modalities. Here, the measurement accuracy of our experimental hardware and software can potentially have an impact on researchers' ability to resolve theoretical controversies.

Take, for example, the LC4MP (Lang, 2000, 2006b, 2017). The LC4MP assumes that there is one central pool of cognitive resources that is accessed by both visual and auditory processes. However, and as a recent systematic review pointed out, the evidence supporting this assumption is rather mixed (Fisher, Keene, et. al., 2018). In a follow-up article, Fisher, Huskey, and colleagues (2018) argued that the LC4MP would benefit by assuming that there are at least two cognitive resource pools, one related to visual information, the other related to auditory information. Lang (2020) has disputed this characterization of the literature. Resolving this controversy will require similar levels of measurement accuracy for reaction time trials presented in visual and auditory modalities. This is currently possible for studies using *Asteroid Impact* with Python 3 on Windows 10 where the mean difference between visual and auditory trials is just 8.5ms. However, there is a considerable difference in measurement accuracy between auditory and visual trials measured using Python 2 on Ubuntu (mean difference = 64.5ms). Differences are even larger when comparing between different operating systems and Python versions (largest mean difference = 76.4ms).

This difference potentially exaggerates modality-specific differences, regardless of any true underlying effect. Of course, it is impossible for our study to clarify if differences in measurement accuracy explain the modality controversy in the LC4MP, or if the modality specific results observed in the LC4MP represent a fundamental characteristic of the human information processing system. Recent work shows that, at least in research using *Asteroid Impact* (Fisher et al., 2019), modality differences are substantially (~215ms) larger than the measurement error observed in our current study (although some studies show considerably smaller differences, see e.g., Huskey et al., 2018). Therefore, it is reasonable to conclude that the modality differences observed in the LC4MP are not an artifact of measurement error. Nevertheless, our study underscores how important measurement accuracy is to answering pressing theoretical concerns, and future research studying more nuanced phenomena should be cautious in making sure that the most accurate software is used.

Our study also raises important meta-scientific implications. Previous research has shown that reaction time effect sizes vary depending on the number of trials measured and the way these trials are averaged (Brysbaert & Stevens, 2018). Our results show that computer hardware and software can potentially impact effect size. This has important implications for meta-analyses of reaction time data in that computer hardware and software potentially introduces a previously unknown source of heterogeneity. And, as others have noted (Carpenter, 2020; Levine & Weber, 2020), heterogeneity presents a threat to construct validity and the interpretation of effect sizes in meta-analyses.

A related concern is about the extent to which an effect replicates. Mean differences between reaction time measurements are often small (Huskey et al., 2020). If two different labs use two different softwares with two different measurement accuracies, then it is possible that a failure to replicate an effect represents a type II (or false negative) error, rather than a true failed replication. This leads to a final meta-scientific concern. A recent meta-analysis of the LC4MP showed that the pooled effect size for STRT measurements was just $\eta^2 = .059$ ($r = .242$). This STRT effect size is 5.7 times smaller than the effect size associated with measurement error as observed in this study (η^2 = .337). Our study suggests that reaction time measurements are low signal, high noise. Generally speaking, low signal and high noise results in low reliability. This is important because measurement reliability constrains the maximum magnitude of an effect size (Lord et al., 1968). Specifically, the maximum correlation between two variables is bounded by the reliability of each variable, as is shown in equation one:

$$r_{ObservedA,ObservedB} = r_{A,B}\, x \sqrt{(reliability_A\, x reliability_B)}$$

What is the reliability of reaction time measures, such as STRT, as used in communication research? The short answer is that we do not know, no study has investigated this. But other test-retest investigations in healthy participants have shown reaction time reliabilities to be quite low (r = .38) for participants completing a simple reaction time task (Weafer et al., 2013). In clinical applications, reaction times used to measure concussions (ICC ranges .36 - .90; Eckner et al., 2011) and ADHD in children (ICC ranges .62 - .72; Soreni et al., 2009) show higher test-retest reliabilities. The implications here are twofold. First, there is an important need to characterize the test-retest reliability of reaction time measures as used in communication research. Second, communication researchers should expect small effect sizes when using reaction time measures in their research.

Limitations

One major design limitation was testing each operating system on a different laptop computer. While the hardware specifications across the three laptop computers were roughly comparable, the lack of experimental control for this introduced potential variation that we could not account for. Indeed, the three-way interaction accounts for 33.7% of the variance in reaction times, which indicates that as of yet unexplored sources of error remain. While it is possible to install all three operating systems onto one computer (this would require a Mac), in practice, our study is more likely to reflect results that are described in the empirical literature. Specifically, different research laboratories will use different machines with different hardware and software configurations. This treats different computers as interchangeable for collecting reaction time data, even when our results clearly show this is simply not true. If anything, our results should serve as a reminder to researchers that hardware and software standardization is of the utmost importance, and that they need to validate reaction time measurements on their own equipment and report the hardware and software characteristics of this equipment. Similar concerns can be raised as behavioral experimentation adopts web-based data collection tools (e.g., Bridges et al., 2020; Schubert et al., 2013). Here, the heterogeneity associated with different hardware and software configurations is vast, and presents a potentially large source of noise, which makes it all the more important that researchers are thoughtful about *a priori* effect sizes, and if their measurement tools are sufficiently accurate to detect these effects.

Another limitation was the lack of control regarding noise from background programs running on the laptops. For instance, pop-up blockers, firewalls, anti-virus programs, updates, and other software/firmware running in the background of each laptop computer could have added latency artifact into the measures (Plant, 2016). Lastly, we did not specifically measure differences in display luminance, refresh rate, or sound quality across each laptop, which may have differentially affected the Arduino's ability to detect auditory and visual trials. Here again, this likely reflects the differences in hardware and software configurations between different research laboratories.

Despite these limitations we are encouraged to see that *Asteroid Impact's* reaction time measures are comparable to other response-latency measuring software. Both ScriptingRT (Schubert et al., 2013), Matlab and Psychtoolbox (Neath et al., 2011) have mean response latencies ranging from 18-100 ms with small standard deviations below 10 ms. Several of our configurations produced response latencies within these limits (although all these tools are not as accurate as PsychoPy; Bridges et al., 2020). Moreover, these values are well below the speed at which the human visual system can detect and respond to complex visual perception tasks (Thorpe et al., 1996). Together, these results address two gaps in the literature. First, we provide an open source and easy to use tool for researchers to evaluate reaction time measurement latencies in their own lab. And second, we show that *Asteroid Impact* is an accurate and precise platform for measuring reaction times.

Conclusion

One final point we make is that method-driven contributions within communication science can be as rewarding and useful as theory-driven research. This idea is encapsulated by Greenwald (2012), who wrote: 'There is nothing so theoretical as a good method'. In this paper, Greenwald notes the relationship between method and theory by showing that more Nobel prizes were awarded for method-driven contributions than for theory-driven ones in the past half century. This suggests that by pursuing better methods we can advance the depth of scientific inquiry and uncover novel research findings, thereby producing valuable theoretical contributions. We believe that the present research supports method-driven contributions to theory testing in communication science by providing an open source, affordable, and easy-to-program device for measuring reaction time latencies. Moreover, our study shows that *Asteroid Impact* can be used as an accurate tool for measuring reaction time latencies in naturalistic (yet high-control) contexts.

References

Anderson, C. A., & Dill, K. E. (2000). Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. *Journal of Personality and Social Psychology*, 78(4), 772–790. https://doi.org/10.1037/0022-3514.78.4.772

Bassili, J. N., & Fletcher, J. F. (1991). Response-time measurement in survey research a method for CATI and a new look at nonattitudes. *Public Opinion Quarterly*, 55(3), 331–346. https://doi.org/10.1086/269265

Bowman, N. D., & Keene, J. R. (2018). A layered framework for considering open science practices. *Communication Research Reports*, 35(4), 363–372. https://doi.org/10.1080/08824096.2018.1513273

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, 8, e9414. https://doi.org/10.7717/peerj.9414

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, 1(1), 9. https://doi.org/10.5334/joc.10

Carpenter, C. J. (2020). Meta-analyzing apples and oranges: How to make applesauce instead of fruit salad. *Human Communication Research*, 46(2-3), 322–333. https://doi.org/10.1093/hcr/hqz018

Clayton, R. B., & Leshner, G. (2015). The Uncanny Valley: The Effects of Rotoscope Animation on Motivational Processing of Depression Drug Messages. *Journal of Broadcasting & Electronic Media*, 59(1), 57–75. https://doi.org/10.1080/08838151.2014.998227

Clayton, R. B., Leshner, G., Sanders-Jackson, A., & Hendrickse, J. (2020). When Counterarguing Becomes the Primary Task: Examination of Dogmatic Anti-Vaping Messages on Psychological Reactance, Available Cognitive Resources, and Memory. *Journal of Communication*, 70(4), 522–547. https://doi.org/10.1093/joc/jqaa01

Dienlin, T., Johannes, N., Bowman, N. D., Masur, P. K., Engesser, S., Kumpel, A. S., Lukito, J., Bier, L. M., Zhang, R., Johnson, B. K., Huskey, R., Schneider, F. M., Breuer, J., Parry, D. A., Vermeulen, I., Fisher, J. T., Banks, J., Weber, R., Ellis, D. A., … Vreese, C. D. (2020). An Agenda for Open Science in Communication. *Journal of Communication*. https://doi.org/10.1093/joc/jqz052

Eckner, J. T., Kutcher, J. S., & Richardson, J. K. (2011). Between-Seasons Test-Retest Reliability of Clinically Measured Reaction Time in National Collegiate Athletic Association Division I Athletes. *Journal of Athletic Training*, 46(4), 409–414. https://doi.org/10.4085/1062-6050-46.4.409

Ferguson, C. J., Rueda, Stephanie. M., Cruz, A. M., Ferguson, D. E., Fritz, S., & Smith, S. M. (2008). Violent video games and aggression: Causal relationship or byproduct of family violence and intrinsic violence motivation? *Criminal Justice and Behavior*, 35(3), 311–332. https://doi.org/10.1177/0093854807311719

Fisher, J. T. & Hamilton, K. (in press). Integrating media selection and media effects using decision theory. *Journal of Media Psychology. Theories, Methods, and Applications*.

Fisher, J. T., Hopp, F. R., & Weber, R. (2019). Modality-Specific Effects of Perceptual Load in Multimedia Processing. *Media and Communication*, 7(4), 149–165. https://doi.org/10.17645/mac.v7i4.2388

Fisher, J. T., Huskey, R., Keene, J. R., & Weber, R. (2018). The limited capacity model of motivated mediated message processing: Looking to the future. *Annals of the International Communication Association*, 42(4), 291–315. https://doi.org/10.1080/23808985.2018.1534551

Fisher, J. T., Keene, J. R., Huskey, R., & Weber, R. (2018). The limited capacity model of motivated mediated message processing: Taking stock of the past. *Annals of the International Communication Association*, 42(4), 270–290. https://doi.org/10.1080/23808985.2018.1534552

Gong, X., Huskey, R., Eden, A., & Ulusoy, E. (May, 2021). *People Prefer Negatively-Valenced Movies in a Two-Alternative Movie Decision Task: A Drift Diffusion Modeling Approach for Testing Mood Management Theory*. Annual Meeting of the International Communication Association, Virtual Conference.

Greenwald, A. G. (2012). There is nothing so theoretical as a good method. *Perspectives on Psychological Science*, 7(2), 99–108. https://doi.org/10.1177/1745691611434210

Huskey, R., Craighead, B., Miller, M. B., & Weber, R. (2018). Does intrinsic reward motivate cognitive control? A naturalistic-fMRI study based on the Synchronization Theory of Flow. *Cognitive, Affective & Behavioral Neuroscience*, 18(5), 902–924. https://doi.org/10.3758/s13415-018-0612-6

Huskey, R., Wilcox, S., Clayton, R., & Keene, J. R. (2020). The Limited Capacity Model of Motivated Mediated Message Processing: Meta-analytically summarizing two decades of research. *Annals of the International Communication Association*, 44(4), 322–349. https://doi.org/10.1080/23808985.2020.1839939

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, 50(1), 46–70. https://doi.org/10.1111/j.1460-2466.2000.tb02833.x

Lang, A. (2006a). Motivated cognition (LC4MP): The influence of appetitive and aversive activation on the processing of video games. In P. Messarsis & L. Humphries (Eds.), *Digital media: Transformation in human communication* (pp. 237–256). Peter Lang Publishing.

Lang, A. (2006b). Using the limited capacity model of motivated mediated message processing to design effective cancer communication messages. *Journal of Communication*, 56(1), S57–S80. https://doi.org/10.1111/j.1460-2466.2006.00283.x

Lang, A. (2009). The limited capacity model of motivated mediated message processing. In R. Nabi & M. B. Oliver (Eds.), *The SAGE handbook of media processes and effects* (pp. 193–204). Sage Publications.

Lang, A. (2017). Limited capacity model of motivated mediated message processing (LC4MP). In P. Roessler, C. Hoffner, & L. van Zoonen (Eds.), *The International Encyclopedia of Media Effects* (pp. 1–9). Wiley & Sons. https://doi.org/10.1002/9781118783764.wbieme0077

Lang, A. (2020). How the LC4MP became the DHCCST: An epistemological fairy tal. In K. Floyd & R. Weber (Eds.), *The Handbook of Communication Science and Biology* (pp. 397–408). Routledge.

Lang, A., & Basil, M. D. (1998). Attention, resource allocation, and secondary task reaction times in communication research: What do secondary reaction task reaction times measure, anyway? In M. E. Roloff (Ed.), *Communication Yearbook* (21st ed., pp. 443–473). Sage Publications.

Lang, A., Bradley, S. D., Park, B., Shin, M., & Chung, Y. (2006). Parsing the resource pie: Using STRTs to measure attention to mediated messages. *Media Psychology*, 8(4), 369–394. https://doi.org/10.1207/s1532785xmep0804_3

Lavrakas, P. J., Kennedy, C., Leeuw, E. D. de, West, B. T., Holbrook, A. L., & Traugott, M. W. (2019). Probability Survey-Based Experimentation and the Balancing of Internal and External Validity Concerns. In P. Lavrakas, M. Traugott, C. Kennedy, A. Holbrook, E. de Leeuw, & B. West (Eds.), *Experimental Methods in Survey Research* (pp. 1–18). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119083771.ch1

Levine, T. R., & Weber, R. (2020). Unresolved Heterogeneity in Meta-Analysis: Combined Construct Invalidity, Confounding, and Other Challenges to Understanding Mean Effect Sizes. *Human Communication Research*, 46(2–3), 343–354. https://doi.org/10.1093/hcr/hqz019

Lewis, N. A. Jr. (2020). Open Communication Science: A Primer on Why and Some Recommendations for How. *Communication Methods and Measures*, 14(2), 71–82. https://doi.org/10.1080/19312458.2019.1685660

Liu, J., & Bailey, R. L. (2019). Effects of Substance Cues in Negative Public Service Announcements on Cognitive Processing. *Health Communication*, 34(9), 964–974. https://doi.org/10.1080/10410236.2018.1446251

Lord, F. M., Novick, M. R., & Birnbaum, A. (1968). *Statistical theories of mental test scores*. Addison-Wesley.

Mathiak, K., & Weber, R. (2006). Toward brain correlates of natural behavior: FMRI during violent video games. *Human Brain Mapping*, 27(12), 948–956. https://doi.org/10.1002/hbm.20234

Matthews, N. L. (2019). Detecting the Boundaries of Disposition Bias on Moral Judgments of Media Characters' Behaviors using Social Judgment Theory. *Journal of Communication*, 69(4), 418–441. https://doi.org/10.1093/joc/jqz021

Miller, A., & Leshner, G. (2007). How Viewers Process Live, Breaking, and Emotional Television News. *Media Psychology*, 10(1), 23–40. https://doi.org/10.1080/15213260701300915

Neath, I., Earle, A., Hallett, D., & Surprenant, A. M. (2011). Response time accuracy in Apple Macintosh computers. *Behavior Research Methods*, 43(2), 353. https://doi.org/10.3758/s13428-011-0069-9

Plant, R. R. (2016). A reminder on millisecond timing accuracy and potential replication failure in computer-based psychology experiments: An open letter. *Behavior Research Methods*, 48(1), 408–411. https://doi.org/10.3758/s13428-015-0577-0

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Schubert, T. W., Murteira, C., Collins, E. C., & Lopes, D. (2013). ScriptingRT: A software library for collecting response latencies in online studies of cognition. *PloS One*, 8(6), 1–12. https://doi.org/10.1371/journal.pone.0067769

Soreni, N., Crosbie, J., Ickowicz, A., & Schachar, R. (2009). Stop Signal and Conners' Continuous Performance Tasks: Test—Retest Reliability of Two Inhibition Measures in ADHD Children. *Journal of Attention Disorders*, 13(2), 137–143. https://doi.org/10.1177/1087054708326110

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 57(4), 421–457.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381(6582), 520–522. https://doi.org/10.1038/381520a0

Weafer, J., Baggott, M. J., & de Wit, H. (2013). Test–retest reliability of behavioral measures of impulsive choice, impulsive action, and inattention. *Experimental and Clinical Psychopharmacology*, 21(6), 475–481.

Weber, R., Alicea, B., Huskey, R., & Mathiak, K. (2018). Network dynamics of attention during a naturalistic behavioral paradigm. *Frontiers in Human Neuroscience*, 12(182), 1–14. https://doi.org/10.3389/fnhum.2018.00182

Wilcox, S., Huskey, R., & DeAndrea, D. C. (in press). Attitude consistent health messages about electronic cigarettes increase processing time: Perceiving message senders as socially close increases message recall. *Journal of Media Psychology. Theories, Methods, and Applications.*.

About the authors

Leandro Calcagnotto, Independent researcher.

Richard Huskey, University of California Davis, Department of Communication, Cognitive Science Program, Center for Mind and Brain, Cognitive Communication Science Lab. ORCiD: https://orcid.org/0000-0002-4559-2439
Correspondence address: 377 Kerr Hall, UC Davis, Davis, CA 95616

Gerald M. Kosicki, The Ohio State University, School of Communication. ORCiD: https://orcid.org/0000-0003-3679-4087