GAMES FOR HEALTH JOURNAL: Research, Development, and Clinical Applications
Volume 4, Number 3, 2015

DOI: 10.1089/g4h.2014.0094

Original Article

Exploring the Potential of the iPad and Xbox Kinect for Cognitive Science Research

Camarin E. Rolle, BA, BS, Bradley Voytek, PhD, and Adam Gazzaley, MD, PhD

Abstract

Many studies have validated consumer-facing hardware platforms as efficient, cost-effective, and accessible data collection instruments. However, there are few reports that have assessed the reliability of these platforms as assessment tools compared with traditional data collection platforms. Here we evaluated performance on a spatial attention paradigm obtained by our standard in-lab data collection platform, the personal computer (PC), and compared performance with that of two widely adopted, consumer technology devices: the Apple (Cupertino, CA) iPad® 2 and Microsoft (Redmond, WA) Xbox® Kinect®. The task assessed spatial attention, a fundamental ability that we use to navigate the complex sensory input we face daily in order to effectively engage in goal-directed activities. Participants were presented with a central spatial cue indicating where on the screen a stimulus would appear. We manipulated spatial cueing such that, on a given trial, the cue presented one of four levels of information indicating the upcoming target location. Based on previous research, we hypothesized that as information of the cued spatial area decreased (i.e., larger area of possible target location) there would be a parametric decrease in performance, as revealed by slower response times and lower accuracies. Identical paradigm parameters were used for each of the three platforms, and testing was performed in a single session with a counterbalanced design. We found that performance on the Kinect and iPad showed a stronger parametric effect across the cued-information levels than that on the PC. Our results suggest that not only can the Kinect and iPad be reliably used as assessment tools to yield research-quality behavioral data, but that these platforms exploit mechanics that could be useful in building more interactive, and therefore effective, cognitive assessment and training designs. We include a discussion on the possible contributing factors to the differential effects between platforms, as well as potential confounds of the study.

Introduction

Recent physiological and physical health studies have compared consumer-facing devices to platforms that are more frequently used for research data collection (i.e., the personal computer [PC]), to assess the feasibility of using such devices as data collection tools.[1–3] However, the use of these new accessible technologies, such as mobile devices and gaming platforms, as cognitive performance data collection platforms to complement or replace more expensive/less mobile hardware is untested. Recently, neuroscientists have designed videogames to assess and train specific cognitive functions with the goal of developing rehabilitation programs for cognitive deficits in diverse populations.[4,5] This type of intervention often benefits from the use of mobile devices and gaming platforms to maximize task engagement (i.e., game interactivity) and accessibility (i.e., platform mobility) of the training environment. In a recent review, Boulos[6] highlighted the potential for improving cognitive performance through Kinect® (Microsoft, Redmond, WA) ''exergames.'' One study found that therapeutic games built on the Kinect platform, a gaming console with motion-detection interactivity, improved cognitive abilities beyond those attained by parallel training games using with sedentary platforms.[7] A more recent study demonstrated enhanced executive functioning in older adults following cognitive training on the Kinect platform.[8] Similarly, studies requiring at-home training or assessment would benefit from the availability of cognitive assessment paradigms on mobile tablets for ease in mobility and cost. We asked the question whether there was a difference between the data collected on different types of platforms used to assess participant performance on a cognitive task.

Materials and Methods

In the current study, we used a new cognitive paradigm that we developed as a modified Posner task to measure

Department of Neurology, Physiology & Psychiatry, Center for Integrative Neuroscience, University of California, San Francisco, San Francisco, California.

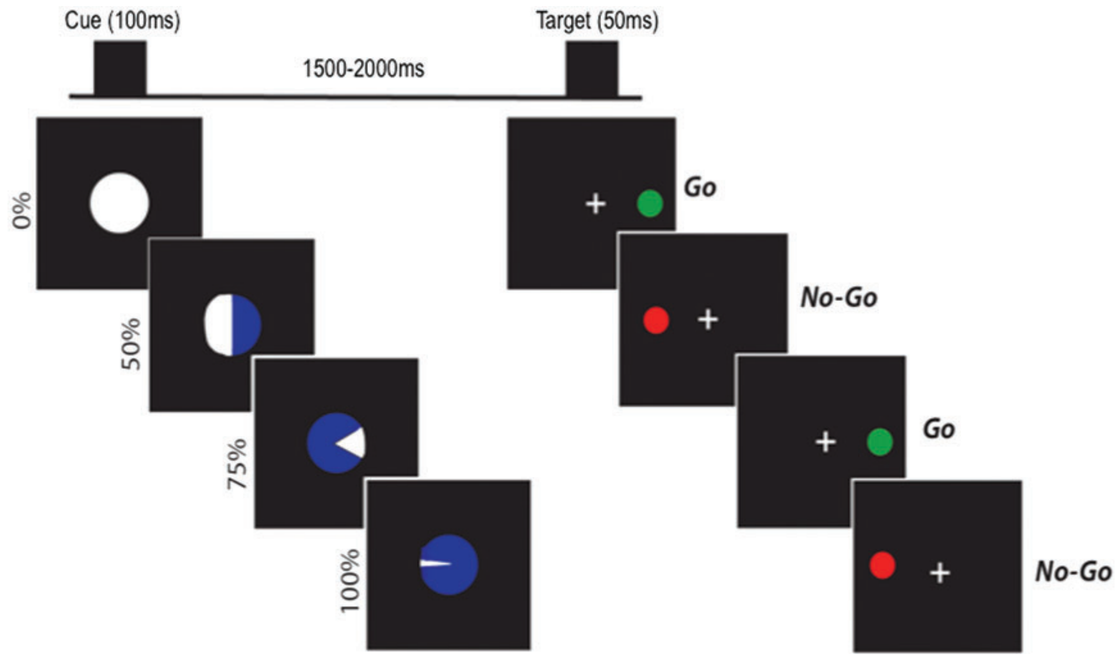

FIG. 1. Task paradigm. Color images available at www.liebertonline.com/g4h

spatial attention (Fig. 1) (see Supplementary Data available online at www.liebertonline.com/g4h).[9,10] The results from the seminal work by Posner[9] demonstrated that spatial cueing enhances performance on a detection task by reducing the area a participant has to distribute his or her attention. Each trial begins with a valid central cue that appears on the screen prior to the appearance of a go/no-go target. The cue presents four levels of information indicating the spatial area of target location: 0 percent, 50 percent, 75 percent, and 100 percent. On a given trial, the target that would appear (following the cue) in the cued space on the screen was either a go or no-go target. If it was a go target, participants were instructed to respond, and if a no-go target, participants were instructed to ignore. We hypothesized that performance on this task would parametrically increase with greater target location predictability provided by the prestimulus cues.

This task was initially developed in our lab for use on a personal computer (PC) platform with a 22-inch cathode ray tube monitor and keyboard for traditional, in-laboratory, cognitive performance data collection. We then adapted this paradigm for data collection on both an Xbox® (Microsoft) with a Kinect motion capture device displayed on a 46-inch monitor and a 12-inch Apple (Cupertino, CA) iPad® 2. The goal was to assess the utility of these consumer platforms, which offer benefits in terms of mobility, engagement, expense, and ease of data collection, compared with our PC data collection platform. We collected data from nine healthy participants (21–33 years old; four females). Participants performed the task twice on all three platforms during a single visit, for a total of 200 trials per platform, and the platform order was counterbalanced across participants.

Data were analyzed using repeated-measures analysis of variance with cue information (0 percent, 50 percent, 75 percent, and 100 percent) and platform (Xbox Kinect, Apple iPad, PC) as within-subjects factors. Pairwise comparisons were planned to examine effects of cue information and platform. Only trials in which participants accurately hit the target following their release from the home position were included in the response time (RT) analyses.

Results

The results for RT (Fig. 2a) showed main effects of platform [$F(2, 16) = 43.06$, $P < 0.001$, $\eta_p^2 = 0.84$] and information [$F(3,24) = 49.87$, $P < 0.001$, $\eta_p^2 = 0.86$], as well as an information × platform interaction [$F(6,48) = 7.509$, $P < 0.001$, $\eta_p^2 = 0.48$]. Pairwise comparisons of the main effects revealed that the Kinect RTs were significantly larger (i.e., slower) compared with both the PC and iPad and that RT increased as the amount of predictive information cued decreased.

Analyses of variance were also run to examine the effect of cue information for each platform independently. The PC, iPad, and Kinect all showed significant main effects of cue information [PC, $F(3, 24) = 5.69$, $P < 0.01$, $\eta_p^2 = 0.41$; iPad, $F(3, 24) = 16.46$, $P < 0.001$, $\eta_p^2 = 0.67$; Kinect, $F(3, 24) = 79.37$, $P < 0.001$, $\eta_p^2 = 0.91$]. Paired t tests were used to examine these main effects. On the iPad, RT differences between all information levels were significant ($P < 0.05$). On the Kinect, RT differed significantly between all information levels, except for the 0 percent and 50 percent information levels, where there was a strong trend [$t(8) = 2.03$, $P = 0.076$]. For the PC, performance on the 0 percent condition was faster than the 50 percent, 75 percent, and 100 percent information levels ($P < 0.05$), but no other comparisons were significant.

Given that RT was significantly faster on the 0 percent condition than the 100 percent condition for all platforms, we were interested in how that RT difference compared between platforms. We ran paired t tests between platforms using an RT cost (0 percent RT – 100 percent RT) to explore this relationship, and found that this cost was significantly larger for the Kinect than for the PC and iPad ($t > 2.50$, $P < 0.05$).

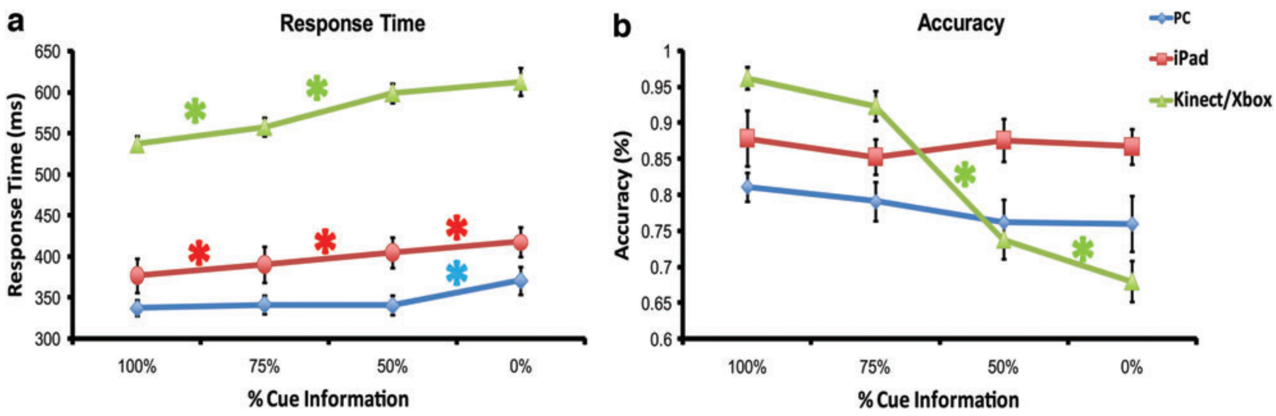

FIG. 2. Performance across platforms: **(a)** response time and **(b)** accuracy across information levels for each platform tested. **(a)** Response times on the Kinect and iPad showed significant between-level differences. **(b)** Accuracy on the Kinect platform showed between-level differences, whereas neither the iPad nor personal computer (PC) did. Data are mean ± standard error of the mean (error bars) values. $^*P < 0.05$. Color images available at www.liebertonline.com/g4h

These results suggest that both the iPad and Kinect showed the greatest sensitivity to cued-information-evoked RT differences compared with the PC.

Performing the same analysis for accuracy (Fig. 2b), we found main effects for information [$F(3, 24) = 32.41$, $P < 0.001$, $\eta_p^2 = 0.80$] and platform [$F(2, 16) = 4.00$, $P < 0.04$, $\eta_p^2 = 0.33$], as well as an interaction between information and platform [$F(6, 48) = 16.58$, $P < 0.001$, $\eta_p^2 = 0.68$]. The main effects indicated that overall accuracy was not comparable across all platforms or across all information levels. Analyses of variance were run to further explore the effect of cue information within each platform. The Kinect showed a main effect of cue information [$F(3, 24) = 5.69$, $P < 0.01$, $\eta_p^2 = 0.41$], whereas the PC and iPad did not [PC, $F(3, 24) = 1.44$, $P = 0.26$, $\eta_p^2 = 0.15$; iPad, $F(3, 24) = 0.60$, $P = 0.62$, $\eta_p^2 = 0.07$]. Using paired *t* tests to further explore these main effects, we assessed within-platform comparisons of information levels. Looking at both adjacent and nonadjacent information-level comparisons, accuracy significantly differed among all information levels for the Kinect, except for a near significant trend for the 100 percent versus 75 percent comparison ($P = 0.058$). The iPad and PC showed no significant between-level accuracy differences ($P > 0.05$). As far as accuracy, the Kinect was most sensitive to detecting between-information level differences.

Discussion

The results from our study show that the iPad and Kinect are both effective platforms for cognitive performance data collection. It is interesting that although all three platforms showed sensitivity to the cue information conditions for RT assessments, both the iPad and Kinect were more sensitive to detecting the hypothesized differences in cued attention than the traditional PC platform. Important differences in capability for task engagement and platform flexibility may explain the greater sensitivity of the iPad, the Kinect, or both.

Both the iPad and Kinect allowed participants to interact directly with the target (i.e., by hitting it), as opposed to the PC, which only allowed participants to report the target's appearance with a button push. The task interactivity that the Kinect and iPad elicited may have enhanced task engagement and consequently increased attention to task elements (i.e., spatial cue) via heightened motivation to maximize task performance.[11] Moreover, the Kinect, which demanded the most physical interaction in task play, also revealed the greatest sensitivity to performance (both RT and accuracy) as a function of cued information.

Another potential variable contributing to differences in performance curves seen across the platforms could be that the method of RT measurement for the PC (button push) is not as sensitive to the cue-information levels as it is for the iPad and Kinect (button/position release). It is also possible that participants were at ceiling across the easier, more predictive cued-information conditions for RTs on the PC and that this diluted the curve such that we could not identify between-level differences. Another consideration should be the difference in screen size between platforms and its potential contribution to the observed differences in results. Given that this study was conducted with healthy adults, it should be noted that it is unknown if clinical populations with either cognitive or physical disabilities might encounter usability issues with the Kinect and iPad (e.g., difficulty standing and moving to target). Future research will explore this important question. In any case, these results encourage the pursuit of innovative platforms for data collection that allow for more interactive task engagement, and may consequently offer the most sensitive means to address hypothesized behavioral patterns.

Author Disclosure Statement.

No competing financial interests exist.

References

1. Clark RA, Bryant AL, Pua Y, et al. Validity and reliability of the Nintendo Wii Balance Board for assessment of standing balance. Gait Posture 2010; 31:307–310.
2. Oresko JJ, Duschl H, Cheng AC. A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. IEEE Trans Inform Technol Biomed 2010; 14:734–740.
3. Black JM, Jacobs RJ, Phillips G, et al. An assessment of the iPad as a testing platform for distance visual acuity in

adults. BMJ Open 2013; 3(6). pii: e002730. doi: 10.1136/bmjopen-2013–002730.
4. Achtman RL, Green CS, Bavelier D. Video games as a tool to train visual skills. Restor Neurol Neurosci 2008; 26:435–446.
5. Anguera JA, Boccanfuso J, Rintoul JL, et al. Video game training enhances cognitive control in older adults. Nature 2013; 501:97–101.
6. Boulos MK. Xbox 360 Kinect exergames for health. Games Health J 2012; 1:326–330.
7. Maillot P, Perrot A, Hartley A. Effects of interactive physical-activity video-game training on physical and cognitive function in older adults. Psychol Aging 2012; 27:589–600.
8. Kayama H, Okamoto K, Nishiguchi S, et al. Effect of a Kinect-based exercise game on improving executive cognitive performance in community-dwelling elderly: Case control study. J Med Internet Res 2014; 16:e61.
9. Posner MI. Orienting of attention. Q J Exp Psychol 1980; 32:3–25.
10. Posner MI, Cohen Y. Components of visual orienting. Attention Perform 1984; 10:531–556.
11. Matthews G, Warm JS, Reinerman LE, et al. Task engagement, attention, and executive control. In: Gruszka A, Matthews G, Szymura B, eds. *Handbook of Individual Differences in Cognition: Attention, Memory, and Executive Control*. New York: Springer; 2012: 205–230.

Address correspondence to:
Camarin E. Rolle, BA, BS
Department of Neurology, Physiology & Psychiatry
University of California, San Francisco
675 Nelson Rising Lane, Box 0444
San Francisco, CA 94158

E-mail: cammie@gazzaleylab.ucsf.edu

Supplementary Data

Participants

All participants were from the San Francisco Bay Area and recruited through online and newspaper advertisement. All participants had normal or corrected-to-normal vision and did not have a history of stroke, traumatic brain injury, or psychiatric illness. Participants did not take psychotropic medication, and all participants reported playing less than 2 hours of videogames a month. All participants were paid for their participation and gave written informed consent prior to participation. The study was approved by the Committee on Human Research at the University of California, San Francisco.

Task details

Each trial begins with a valid central cue that appears on the screen prior to the appearance of a go/no-go target. The cue presents four levels of information indicating the spatial area of target location on an imaginary 360° annulus: 0 percent, 50 percent, 75 percent, and 100 percent. In the 0 percent condition, there is no information embedded in the cue, and thus the target could appear anywhere on the 360° annulus. In the 50 percent cued-information level, half of the cue is colored, thus directing attention to half of the screen (180° arc) as a potential target location. The 75 percent information cue narrows the target location to 25 percent of the screen (90° arc). Similarly, the 100 percent information cue gives exact information about the specific location of the target. The predictive cue was on the screen for 100 milliseconds, and the target was on the screen for 50 milliseconds. The stimulus onset asynchrony on all platforms was 1750 ± 250 milliseconds. The four cue information levels were randomized at 25 percent appearance rates.

On a given trial, the target that would appear in the cued space on the screen was either a go or no-go target. If it was a go target, participants were instructed to respond, and if a no-go target, participants were instructed to ignore. A fixation cross on the center of the screen served as a gaze-fixation point during the trial prior to target appearance and was also used to provide immediate posttrial feedback to the participant. The fixation cross flashes green if the subject correctly responded to the go target within a 2-second window or correctly ignored the no-go target, and flashes red if he or she incorrectly responded to a no-go target, or if he or she didn't respond to a go target within the 2-second window. All participants were instructed to respond to go targets as quickly as possible, without sacrificing accuracy, and to withhold responses to no-go targets. Participants were instructed to keep their gaze fixed on the central fixation cross throughout each trial until the appearance of the target.

All participants were oriented to each device and given practice trials and demonstration to ensure equivalent comfort with the three platforms. On the PC version, participants were instructed to press a single button on a keyboard when they saw a go target appear on the screen in the space predicted by the cue. RT was indexed as the time between the appearance of the go target and the time of the response button press. For the Kinect and iPad, participants were instructed to respond to the target by physically ''hitting'' it at the actual location it appeared, capitalizing on the respective advantages offered by motion capture and touch screen technology compared with the keyboard of the PC. Therefore, RT for the Kinect and iPad was indexed as the time difference between the appearance of the go target and the participant's initial movement from the required holding state at a home position. On the iPad, we implemented a ''home'' position on the bottom of the screen where participants were instructed to maintain their dominant index finger throughout the trial prior to target appearance. The Kinect used a similar method; participants were instructed to keep their hands at their sides, and the Kinect was programmed to verify this position before starting each trial. Because the time to hit the target is affected by its physical distance from the home position, the actual hit time was not a reliable reflection of cognitive RT.

For the Kinect version, participants were standing in an open exam room approximately 4 feet from the screen. Participants played the iPad version of the task sitting down, with the iPad lying flat on a table in front of them.

RT variance

To investigate whether there is an effect of within-subject variance in RT between platforms, we used the same analysis as above for RT standard deviation. We found a two-way interaction between information level and platform [$F(6, 48) = 2.60$, $P < 0.05$]. We also found main effects for information level [$F(3, 24) = 3.48$, $P < 0.05$] but not for platform [$F(2, 16) = 3.06$, $P = 0.075$]. Using paired t tests to further explore the two-way interaction, we ran paired t tests within information levels between platforms. We only saw differences in variance on the 100 percent information level, with the Kinect reporting lower variances than both the PC [$t(8) = 3.45$, $P < 0.01$] and the iPad [$t(8) = 2.29$, $P < 0.05$]. For within-platform comparisons between information levels, the Kinect showed significant differences between all information levels, as well as trending significance for the 100 percent versus 75 percent comparison [$t(8) = 2.10$, $P = 0.069$]. Neither the PC nor the iPad showed any differences in variance between information levels. It is interesting that these results show that the Kinect has the most sensitivity to detect within-subject RT variance between information levels and is more or less comparable to the other platforms in measures of within-level variance.