Computers in Human Behavior 51 (2015) 8–14

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Eye-hand coordination strategies during active video game playing: An eye-tracking study

Yuping Chen [a], Meng-Jung Tsai [b,*]

[a] Department of Physical Therapy, Georgia State University, USA
[b] Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan

ARTICLE INFO

Article history:

Keywords:
Children
Eye movements
Eye-hand coordination
Game-based learning
Visual strategy

ABSTRACT

The purpose of this exploratory study was to examine the eye-hand coordination patterns while playing two virtual-reality active video games in healthy children and adults. Eleven children (mean age 8.09 years) and ten adults participated in the study. Each participant played two digital games, Slap Stream and Kung Foo, from EyeToy Play software. Eye movements were recorded using Mobile Eye eye-tracker. Eye-hand coordination strategies and the time when virtual object appeared, the gaze shifted to the object, the reach started, the gaze shifted away, and the reach ended were coded from the video. The latencies between these events were computed and compared between adults and children and between games. The fixation duration, number of fixations, and number of gaze points were also computed for each game's areas of interests. Results showed that (1) all participants used multiple eye-hand strategies while playing active video games with some strategies more than others; (2) the Kung Foo game (with one target appearing on the screen) and the Slap Stream game (with potentially multiple targets appearing on the screen) induced different latencies and gaze points between children and adults; and (3) children had longer latencies and shorter fixation durations than adults. The study thus provides in-depth understanding of different patterns of eye-hand coordination in relations to active video game playing. The significant differences in coordinative control strategies found between adults and children as well as between game types provide a basis for further research in both child development and game-based learning fields.

1. Introduction

With the rapid development of game technology only the last decade, digital game playing has become one of the most popular activities for learning in school and for families at home. The cognitive process of computer game playing involves a complex process of multitasking: paying attention to a computer screen, catching the visual or verbal information stimulus, integrating the information stimulus with prior experience or prior knowledge in working memory, making a decision based on newly integrated knowledge, planning an action based on the decision, taking the action, anticipating responses and paying attention again to the computer screen. All of the elements of this process can simultaneously occur during only a few seconds while playing an interactive game. Recent studies in digital learning and education have begun to examine digital cognitive process using eye-tracking technology (Anderson, Love, & Tsai, 2014; Lai et al., 2013; Van Gog & Scheiter, 2010), but most of them focused on static image processing tasks such as text-and-graphic reading comprehension (e.g., Canham & Hegarty, 2010; Ho, Tsai, Wang, & Tsai, 2014; Liu, Lai, & Chuang, 2011), word-based or picture-based problem solving tasks (e.g., Tsai, Hou, Lai, Liu, & Yang, 2012) and dynamic images without interactivities (e.g., de Koning, Tabbers, Rikers, & Paas, 2010; Jarodzka, Scheiter, Gerjets, & van Gog, 2010). How do people view and interact with dynamic and highly interactive multimedia stimuli such as active video games has not been explored or discussed. How learners coordinate eyes and hands in video game playing may demonstrate the effective and efficient learning strategies critical for successful learning in game-based learning environments.

In our daily lives, our eyes, head, and hands are constantly in motion within different contexts (Pelz, Hayhoe, & Loeber, 2001). Gaze is used actively to gather information for the control of actions (Mennie, Hayhoe, & Sullivan, 2007): Eye-hand coordination is important to assure smooth and accurate hand movements, which include the shifting of gaze ahead of the hand, and attention moves ahead of gaze to an intended landing position (Deubel & Schneider, 1996; Hayhoe & Ballard, 2005). That is, we control gaze

* Corresponding author at: Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, 43, Sec. 4, Keelung Rd., Taipei 106, Taiwan. Tel./fax: +886 (2) 2737 6270.
E-mail address: mjtsai99@mail.ntust.edu.tw (M.-J. Tsai).

http://dx.doi.org/10.1016/j.chb.2015.04.045
0747-5632/

shifts and fixations proactively in order to gather sufficient visual information to guide smooth hand movements (Johansson, Westling, Backstrom, & Flanagan, 2001).

Anticipatory eye movements have been observed in natural tasks (e.g., (Land, Mennie, & Rusted, 1999; Mennie et al., 2007). Pelz et al. (2001) asked five healthy adults using their preferred hand, to perform a simple block-copy task, which involved gaze fixations to gather information about the block pattern, and hand movements to pick up and place blocks to copy the block pattern (Pelz et al., 2001). Generally, the eye always arrived first, followed by the hand for both pick-up and drop block actions. However, the patterns of eye and hand coordination differed between pick-up and drop events: When picking up the block, the eye departed from the block to be picked up 100–150 ms before the hand contacted the block. When dropping the block, the eye maintained fixation on the block until the block was in the desired place. Eye-hand latencies for initiation of the pick-up and drop actions were quite similar, regardless of the strategies used.

Mennie et al. (2007) asked 12 undergraduate students to assemble models using four wooden pieces in two sequences; in one sequence reaching and grasping were interrupted by an additional action to join the pieces together using wooden nuts and bolts (Mennie et al., 2007). In these two sequences, participants were asked to perform the same actions for the first portion (i.e., reach out to pieces 1 and 2 and join them together into a T-shape) but perform different actions afterwards. Results showed that 56.4% of all fixations were within the workspace and up to 94% of fixations were to the task-related areas. When reaching to a container to pick up a piece to assemble, the participants usually fixated on the container to guide movement prior to reaching for the piece. The latency between eye and hand was around 230 ms with guiding fixations, but they were much longer, around 353 ms, with look-ahead fixations. Look-ahead fixations were defined as fixations on objects not relevant to the immediate sub-task, but relevant to a future sub-task (Pelz & Canosa, 2001). That is, the participants fixated on objects that were manipulated several seconds later, showing advanced anticipatory use of gaze in acquiring information about objects for future manipulation.

Using eye-tracking technology to simultaneously examine gaze and movement-based indices enable researchers to capture eye-hand coordination skills which are essential for successful task performances or learning outcomes. Gaze and hand movement behavior have previously been examined in skill/procedural learning domains. For example, a recent study (Wilson et al., 2010) used a computer-simulated surgery training task which involved reaching for and touching a small target to compare the visual strategies used by expert and novice surgeons. The expert surgeons were found to have fewer errors, faster task time and longer final fixation on the target location, while the novice surgeons split their time fixating on the target and the tools. Another recent study (Causer, Harvey, Snelgrove, Arsenault, & Vickers, 2014) used a mobile eye tracker to examine the training effects on surgical residents' knot tying performances and visual strategies. Percentage of the final fixation duration located on a critical target before the initiation of a critical phase of the movement and the number of fixation recorded on the critical target were used along with total movement time and hand movement phase time, in order to compare performance effectiveness and movement efficiency between different training methods. They reported that higher performance scores are associated with a longer critical fixation duration, fewer fixations, faster total knot tying times, and faster movement phase times. In sum, these studies showed a support for using eye-tracking technology to better understand how people utilize visual information to plan and control movements in eye-hand coordinated learning tasks.

There were limited studies have investigated the coordination between gaze and hand movements in children in daily activities. Active video game consoles, such as the Nintendo Wii and Microsoft Xbox Kinect provide a virtual environment for users to interact with virtual objects in the digital world. Videogames represent an important part of children's leisure life (Marshall, Gorely, & Biddle, 2006). In the virtual environment, multiple objects might occur on the same screen simultaneously, however in relatively fixed locations, providing an opportunity to observe how users shift their gaze in a dynamic 2-D environment. To the best of our knowledge, there was no research investigating the eye-hand coordination when children playing active video games. The purpose of this exploratory study therefore was to examine the characteristics of unrestrained eye and hand movements in two real-world tasks (i.e. playing two virtual-reality games) in healthy children and adults, using adults as the reference group. We expected that children and adults would show different coordination of gaze and hand movements when playing in two active video games; and both children and adults would show different gaze and hand movements in the two video games.

2. Materials and methods

2.1. Participants

Eleven children with typical development (6 males and 5 females, age: 8.09 ± 1.87 years) and 10 healthy young adults (3 males and 7 females, age: 26.50 ± 3.57 years) volunteered to the study. All were right-handed without any known neurological deficits. All but 2 children had experience in playing video games.

2.2. Apparatus

The ASL Mobile Eye with a sampling rate of 60 Hz was used to track and record each participant's gaze data while playing the game. The lightweight eyeglasses consist of two digital high resolution cameras: one records the scene image and the other records the participant's eye. These images are then integrated into a single video recording representing the scene with a superimposed gaze cursor. This eye tracker can collect eye movements and gaze point information during the performance of natural tasks, allowing the use of unconstrained eye, head and hand movements.

The commercial EyeToy-Play System (Play Station 2, Sony) was used to provide the video games. The system consisted of a USB camera, EyeToy-Play software, a TV, and Playstation 2. In this system, the USB camera was used as a capturing and tracking device to place participants within the computerized VR environment so that they could interact with the virtual objects. A participant could see him/herself displayed in the middle of the TV screen. When the participant moved his/her arm, the corresponding image inside the TV also moved the arm. Among the games on EyeToy-Play software, Slap Stream and Kung Foo were selected. In the Slap Stream game (see Fig. 1), there were four clouds located at the four corners of the TV screen. The task for the participant was to smack the rats that appeared randomly in any of the four clouds, while avoiding the bunny girl. Each trial lasted about 3 min or ended after the participant hit three bunny girls. In the Kung Foo game (see Fig. 2), the opponents came from six locations on the TV screen: the upper, middle, and lower corner of each side. Each trial also lasted about 3 min or ended after the participant missed three opponents. Both games required eye-hand coordination with the chance to deal with multiple virtual objects simultaneously, the appearance of virtual objects was random. Although both games required good eye-hand coordination, Slap Stream also required participants to decide whether or